

Российская академия наук
Самарский научный центр
Институт проблем управления сложными системами

Труды XII Международной конференции
Proceedings of the XII International Conference

ПРОБЛЕМЫ УПРАВЛЕНИЯ И МОДЕЛИРОВАНИЯ В СЛОЖНЫХ СИСТЕМАХ

COMPLEX SYSTEMS: CONTROL AND MODELING PROBLEMS

21-23 июня 2010, Самара, Россия

June 21-23, 2010, Samara, Russia

Редакторы:

академик **Е.А. Федосов**

академик **Н.А. Кузнецов**

профессор **В.А. Виттих**

Editors:

academician **E.A. Fedosov**

academician **N.A. Kuznetsov**

professor **V.A. Vittikh**

Проблемы управления и моделирования в сложных системах:

Труды XII Международной конференции (21-23 июня 2010 г. Самара, Россия) /

Под ред.: акад. **Е.А. Федосова**, акад. **Н.А. Кузнецова**, проф. **В.А. Виттиха**. -

Самара: Самарский научный центр РАН, 2010. - 712 с.

ISBN 978-5-93424-476-8

В предлагаемом сборнике содержатся труды XII Международной конференции «Проблемы управления и моделирования в сложных системах», распределенные по следующим направлениям: системный анализ и теория управления, управление и измерения в сложных технических системах, реформирование государственного управления в регионе, новые информационные технологии в управлении, мультиагентные модели и системы.

ISBN 978-5-93424-476-8

- © Учреждение Российской академии наук Институт проблем управления сложными системами РАН, составление, оформление, 2010
- © Все права принадлежат авторам публикуемых работ, 2010

The XII International Conference “Complex Systems: Control and Modeling Problems” (CSCMP’2010), was held on June 21–23, 2010 in Samara (Russia) by the International Association for Mathematics and Computers in Simulation (IMACS), National Committee of Automatic Control of Russia, Samara Scientific Centre of RAS, Institute for the Control of Complex Systems (ICCS) of RAS with the participation of the Volga Region University of Telecommunications and Informatics and Knowledge Genesis Company (Samara, RF).

The Conference was funded according to the project № 10-08-06036 of Russian Foundation for Basic Research.

The conference reports were presented on the following sections:

- System Analysis and Control Theory;
- Control and Measurement in Complex Technical Systems;
- Reform of Regional Governance;
- New Information Technologies in Control and Management;
- Multi-Agent Models and Systems

The scientists from universities and research institutes of Great Britain, Germany, Netherlands, Italy, Turkey participated at the conference. The Russian participants represented the Russian Academy of Sciences, universities, research and development centers, enterprises and governmental authorities.

XII Международная конференция «Проблемы управления и моделирования в сложных системах» (ПУМСС-2010) проводилась в г. Самаре (Россия) 21-23 июня 2010 года Международной ассоциацией по математическому и компьютерному моделированию (IMACS), Российским Национальным комитетом по автоматическому управлению, Самарским научным центром РАН, Институтом проблем управления сложными системами (ИПУСС) РАН при участии Поволжского государственного университета телекоммуникаций и информатики, а также научно-производственной компании «Генезис знаний» (г. Самара).

Конференция получила финансовую поддержку Российского фонда фундаментальных исследований (проект № 10-08-06036).

На Конференции были представлены доклады по следующим секциям:

- системный анализ и теория управления;
- управление и измерения в сложных технических системах;
- реформирование государственного управления в регионе;
- новые информационные технологии в управлении;
- мультиагентные модели и системы.

В Конференции приняли участие ученые из университетов и научных учреждений Великобритании, Германии, Нидерландов, Италии, Турции. Отечественные ученые представляли Российскую академию наук, вузы, научно-производственные объединения и промышленные предприятия страны, органы государственной власти.

ПРОГРАММНЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

1. *Федосов Е.А.* - академик РАН, председатель Программного комитета
2. *Виттих В.А.* - д.т.н., заместитель председателя, ИПУСС РАН, г. Самара
3. *Боровик С.Ю.* - к.т.н., ученый секретарь, ИПУСС РАН, г. Самара
4. *Анишаков Г.П.* - член-корреспондент РАН
5. *Васильев С.Н.* - академик РАН
6. *Вишиневецкий Р.* - президент IMACS, профессор Университета Радгерса, г. Нью-Йорк, США
7. *Инден У.* - профессор Кельнского университета, г. Кельн, Германия
8. *Клименко А.В.* - к.э.н., проректор Государственного университета – Высшей школы экономики, г. Москва
9. *Кузнецов Н.А.* - академик РАН
10. *Нариньяни А.С.* - к.ф.-м.н., генеральный директор ЗАО «ИнтелиТек», г. Москва
11. *Новиков Д.А.* - член-корреспондент РАН
12. *Рапопорт Э.Я.* - д.т.н., ИПУСС РАН, г. Самара
13. *Ржевский Г.А.* - профессор Открытого университета, г. Лондон, Великобритания
14. *Себряков Г.Г.* - член-корреспондент РАН
15. *Силберт Н.* - директор корпорации «Национальные ИТ-исследования», г. Чикаго, США
16. *Теряев Е.Д.* - член-корреспондент РАН
17. *Шорин В.П.* - академик РАН

ОРГАНИЗАЦИОННЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

1. *Виттих В.А.* - д.т.н., председатель Организационного комитета, ИПУСС РАН, г. Самара
2. *Смирнов С.В.* - д.т.н., заместитель председателя, ИПУСС РАН, г. Самара
3. *Боровик С.Ю.* - к.т.н., ученый секретарь, ИПУСС РАН, г. Самара
4. *Андреев В.А.* - д.т.н., Поволжский государственный университет телекоммуникаций и информатики, г. Самара
5. *Гриценко Е.А.* - заместитель министра - руководитель департамента организации социальных выплат и развития информационных технологий Министерства здравоохранения и социального развития Самарской области
6. *Жеребин А.М.* - д.т.н., ГосНИИ авиационных систем, г. Москва
7. *Задирако И.Н.* - к.э.н., заместитель генерального директора ФГУ «Объединение «Росинформресурс» Минпромэнерго России, г. Москва
8. *Золотухин Ю.Н.* - д.т.н., Институт автоматизации и электротехники СО РАН, г. Новосибирск
9. *Ильясов Б.Г.* - д.т.н., Уфимский государственный авиационный технический университет
10. *Клецев А.С.* - д.ф.-м.н., Институт автоматизации и процессов управления ДВО РАН, г. Владивосток
11. *Козырев О.Р.* - к.ф.-м.н., директор Нижегородского филиала Государственного университета – Высшей школы экономики
12. *Кузнецов С.В.* - зам. директора ИПУСС РАН, г. Самара
13. *Кульба В.В.* - д.т.н., Институт проблем управления РАН, г. Москва
14. *Лазарев Ю.Н.* - д.т.н., Самарский научный центр РАН
15. *Скобелев П.О.* - д.т.н., генеральный директор НПК «Генезис знаний», г. Самара

CONTENTS СОДЕРЖАНИЕ

Plenary Papers

Пленарные доклады

1. *G. Rzevski* 3
Product life as a complex system
2. *U. Inden* 5
Real-time Management and Complexity.
A strategy model for real-time management in complex service systems
3. *Е.Д. Теряев, К.В. Петрин, А.Б. Филимонов, Н.Б. Филимонов* 18
Проблема интеллектуальной поддержки оператора
в эргатических системах управления
4. *A.C. Bijl* 28
Environment care in the Russian Federation: an analysis from the Dutch perspective
5. *В.А. Виттих, А.Н. Гинзбург, О.Р. Козырев, С.В. Смирнов* 33
Демократизация процессов управления ВУЗом
6. *В.Ф. Меньшиков, А.Г. Назаров, О.Е. Головихина* 40
Состояние экологической безопасности в атомной отрасли

System Analysis and Control Theory

Системный анализ и теория управления

1. *B. Nacke* 51
Optimization techniques for design and control of induction heating systems
2. *E. Baake* 59
Analysis of metallurgical melting processes using numerical simulation
3. *F. Dughiero, M. Forzan, S. Lupi, F. Nicoletti* 67
Aluminium billets heated by rotating magnets: comparison of
analytical and numerical results
4. *B. Nacke, Yu.E. Pleshivtseva* 74
The review of common activity of ETP and SamSTU in the field of
simulation and optimisation of electrotechnological processes
5. *Yu.E. Pleshivtseva* 81
Optimisation of production complex for metal pressing with induction pre-heating
6. *E. Baake, K. Blinov, S. Korshikov, O. Sharapova* 88
Numerical simulation of multi-physics dynamic processes in induction heating
systems granted by German Academic Exchange Service (DAAD)
7. *Yu.V. Shemyakin* 94
Mathematical simulation of pressing stage
in the “heater – hot forming equipment” line

8.	<i>А.Н. Шатунов, И.В. Позняк, А.Ю. Печенков, А.И. Максимов, В.В. Кичигин</i> Исследование энергетических характеристик индукционной системы с холодным тиглем	100
9.	<i>В.В. Кичигин, И.В. Позняк, А.Ю. Печенков, А.И. Максимов, А.Н. Шатунов</i> Технология получения мультикристаллического кремния на основе индукционной плавки в холодном тигле	105
10.	<i>Э.Я. Рапопорт</i> Элементы синергетической теории управления динамическими системами с распределенными параметрами	110
11.	<i>О.С. Гавриленко, В.С. Елсуков, В.И. Лачин</i> Синтез нелинейных систем комбинированного управления с бинарно-операторными связями	120
12.	<i>Н.В. Дилигенский, М.В. Цапенко, А.Н. Давыдов</i> Сравнительный многокритериальный анализ эффективности операционной деятельности промышленных предприятий	126
13.	<i>Ф.Д. Михайлов</i> Прямые и обратные задачи распространения тепла в почве	137
14.	<i>А.М. Прокофьев</i> Устойчивое управление знаниями в компании: проблемы и задачи	141
15.	<i>Д.И. Пунда</i> Когнитивная технология управления. Технология со-управления, модель для управления	146
16.	<i>И.Д. Виноградов</i> Онтологический подход к построению систем сбалансированных показателей	158
17.	<i>А.Н. Дилигенская</i> Синтез и анализ энергоэффективных систем управления на основе наблюдателей состояния	165
18.	<i>М.А. Чефранова</i> Применение методологии системного анализа при разработке когнитивных моделей кредитного риска	171
19.	<i>Б.Г. Ильясов, И.В. Дегтярева, Е.А. Макарова, А.Н. Павлова</i> Моделирование неравновесных состояний рынка труда и его влияния на динамику кругооборота финансовых потоков	176
20.	<i>Б.Г. Ильясов, А.Ш. Назаров, Г.А. Саитова</i> Алгоритм синтеза регулятора адаптивной многосвязной системы автоматического управления на основе анализа характеристических функций	187
21.	<i>В.Е. Гвоздев, Д.В. Блинова</i> Комплексный анализ состояния территориальных систем на основе решения задачи типологической классификации	195
22.	<i>О.В. Скачков, И.А. Ларин</i> Методы менеджмента качества и бережливого производства в системе управления перевозочным процессом на железнодорожном транспорте	204

23.	<i>А.М. Кистанов</i> Системный анализ восприятия наглядных алгебраических структур	211
24.	<i>А.З. Асанов, Д.Н. Демьянов</i> Аналитическое конструирование систем управления многосвязными динамическими объектами с учётом передаточных нулей	218
25.	<i>А.З. Асанов, В.С. Каримов</i> Синтез системы автоматического управления для многосвязного объекта с запаздываниями по управлению на основе компенсационно-наблюдательной схемы	224
26.	<i>А.З. Асанов, И.Ю. Мышкина</i> Методика корректировки учебного плана профессиональной подготовки с учетом требований заинтересованных сторон	232
27.	<i>Р.Н. Курманалиева</i> Синтез линейных систем управления с использованием эталонной модели	240
28.	<i>М.Н. Филиппов</i> Метод оценки надежности программного обеспечения по информации, сохраняемой системой управления версиями	244
29.	<i>Л.Л. Чекалов</i> Описание, анализ и преобразование структуры сложных многоуровневых процессов в автоматизированных системах с использованием графов с оболочкой	250
30.	<i>Л.Л. Чекалов</i> Информационная модель сознания и описание эволюции её функций от животного к человеку	256
31.	<i>В.И. Батищев, Н.Г. Губанов</i> Категорные методы комплексного представления и структуризации разнородных данных в информационных системах анализа сложных объектов	263
32.	<i>П.А. Зиновьев, А.В. Мейко</i> Применение марковских метрик «накопленного дохода» в задачах анализа живучести корпоративных информационных систем	268
33.	<i>И.В. Аникин</i> Методика управления внутренними рисками информационной безопасности в корпоративных информационных сетях	274
34.	<i>И.В. Аникин, В.С. Лукоянов</i> Методология выявления внутренних нарушителей в корпоративных информационных сетях	281

Control and Measurement in Complex Technical Systems
Управление и измерения в сложных технических системах

1.	<i>Ю.Н. Золотухин, А.А. Нестеров</i> Управление перевернутым маятником с подвижной точкой подвеса	291
----	--	-----

2.	<i>И.С. Бабер, А.Н. Гинзбург, Н.И. Карпов, С.П. Марьев, О.Г. Митрофанов</i> Система управления установкой магнетронного напыления	298
3.	<i>А.Н. Гинзбург, В.М. Дёмкин</i> Алгоритмы селекции спектрограмм в системах автоматизации атмосферных микроволновых наблюдений	304
4.	<i>В.А. Белослудцев, Ю.Н. Секисов</i> Система измерений координат смещений приводного вала топливного насоса силовой установки	310
5.	<i>В.А. Белослудцев, С.Ю. Боровик, М.М. Кутейникова, Б.К. Райков, Ю.Н. Секисов, О.П. Скобелев, В.В. Тулунова</i> Экспериментальная проверка работоспособности кластерных методов измерения координат смещения торцов лопаток и деформаций статора в лабораторных условиях	316
6.	<i>М.М. Кутейникова, Ю.Н. Секисов</i> Алгоритм вычисления радиальных и осевых смещений торцов лопаток	323
7.	<i>Л.Б. Беленький, С.Ю. Боровик, Б.К. Райков, О.П. Скобелев</i> Систематизация и классификация групповых измерительных цепей одновитковых вихретоковых датчиков	328
8.	<i>Л.Б. Беленький, С.Ю. Боровик, О.П. Скобелев</i> Анализ основных характеристик групповых измерительных цепей одновитковых вихретоковых датчиков	336
9.	<i>В.В. Тулунова</i> Об адекватности аналитической модели электромагнитного взаимодействия чувствительного элемента одновиткового вихретокового датчика с торцом лопатки и её совершенствование	343
10.	<i>А.Н. Галютдинова, Ю.С. Масленникова, А.А. Пороунов</i> Результаты моделирования каналов аэрметрической системы измерения воздушных сигналов вертолета	348
11.	<i>П.Ю. Вильвер, А.В. Протасов</i> Имитационное моделирование процесса определения нагрузки от колесных пар локомотива	364
12.	<i>Н.В. Фалдин, А.Ю. Липатов</i> Синтез оптимального по быстродействию следящего электропривода постоянного тока	370
13.	<i>В.А. Яшков</i> Струйный измеритель кинематических параметров системы управления подвижным объектом	376
14.	<i>Г.К. Чигарьков, А.В. Кочемасов</i> Калибровка акустических антенных решеток	382
15.	<i>А.П. Колеватов, А.В. Куликов</i> Метод определения юстировочных углов при отсутствии точных данных о параметрах ориентации объекта	387

16.	<i>С.Г. Николаев, Ю.В. Ившина</i> Моделирование калибровки бесплатформенной инерциальной навигационной системы по выходным сигналам	392
17.	<i>С.В. Моржова, Н.В. Фалдин</i> Методы исследования чувствительности автоколебаний в релейных системах управления к изменению параметров объекта управления	398
18.	<i>S.A. Kizhaev, A.K. Bulkhin, V.F. Klyuchnikov</i> Basic principles of intellectual computer calculators construction for real-time simulation of manufacturing processes and control	404
19.	<i>С.А. Кижжаев, А.К. Бульхин, В.Ф. Ключников</i> Разработка метода оценки адекватности синтезированных моделей реальным объектам в сложных технических системах, использующих разные виды энергии	412
20.	<i>С.А. Кижжаев, А.К. Бульхин, В.Ф. Ключников</i> Практические примеры реализации методов декомпозиции и децентрализации в сложных технических системах	420
21.	<i>М.А. Габидулин</i> Математическая модель четырехфазного растрового СКП с кольцевой считывающей системой и аксиальным излучателем	429
22.	<i>В.А. Буцких, Н.Н. Васин</i> Механизм нечеткого логического вывода в управлении железнодорожным переездом	435

Reform of Regional Governance

Реформирование государственного управления в регионе

1.	<i>В.Д. Фурасов, Д.В. Фурасов</i> Сравнительный анализ развития федеральных округов России. Численность населения, рождаемость и смертность	443
2.	<i>В.А. Виттих</i> Феноменологическое время в процессах принятия решений	452
3.	<i>А.Н. Визгунов, М.Л. Зубов, А.В. Шишин</i> Системы оперативной обработки и анализа обращений граждан как эффективное средство поддержки решений в системе государственного и муниципального управления	458
4.	<i>Н.В. Асеева, Э.А. Бабкин, А.Н. Визгунов, В.П. Князькин, О.Р. Набиуллин, М.С. Шиткова</i> Использование имитационных моделей для проверки выполнимости административных регламентов	464
5.	<i>В.Ю. Волков, В.В. Батышклина, Б.А. Самаха</i> О проблемах принятия управленческих решений в автоматизированной системе экологического мониторинга и управления г. Новомосковска Тульской области	470

- | | | |
|----|--|-----|
| 6. | <i>И.В. Дрынова, С.В. Дегтярев</i>
Имитационный подход к моделированию пожароопасных ситуаций в ВУЗах | 478 |
| 7. | <i>Д.В. Абарин</i>
Выявление отличительных признаков моделей отраслей жилищно-коммунального хозяйства и прогнозирование их развития | 483 |
| 8. | <i>Д.П. Маевский, В.А. Филимонов</i>
Концепция проекта создания системы коллективного управления образовательным учреждением с использованием технологий ситуационного центра | 488 |

New information technologies in Control and Management
Новые информационные технологии в управлении

- | | | |
|-----|--|-----|
| 1. | <i>Babkin E.A.</i>
Integration of telecommunication management systems using ontology-based methods | 495 |
| 2. | <i>А.Н. Гинзбург, В.М. Дёмкин</i>
Компьютерный контроль знаний – миф или реальность | 501 |
| 3. | <i>О.И. Бедняк, М.А. Кораблин, А.А. Салмин, С.В. Смирнов</i>
Категориальный анализ в задачах принятия управленческих решений | 505 |
| 4. | <i>Ю.А. Загорулько, Г.Б. Загорулько, С.В. Булгаков</i>
Подход к разработке системы поддержки принятия решений для добывающего предприятия нефтегазового комплекса | 512 |
| 5. | <i>З.В. Апанович, Т.А. Кислицина</i>
Расширение подсистемы визуализации наполнения информационного портала средствами визуальной аналитики | 518 |
| 6. | <i>С.В. Смирнов, Е.Г. Суворова</i>
Адекватная реализация когнитивного потенциала обозревателя онтологических моделей | 526 |
| 7. | <i>Е.А. Сидорова</i>
Обзор задач ЕЯ-сервисов в информационных системах под управлением онтологии | 534 |
| 8. | <i>И.С. Кононенко, Е.А. Сидорова</i>
Применение лингвистических технологий в справочно-информационной системе по катализу | 540 |
| 9. | <i>А.Н. Гинзбург, М.В. Игнатьев</i>
Концептуальная модель контроллера обеспеченности ресурсами в СППР на основе онтологий | 548 |
| 10. | <i>П.В. Ситников, Е.Г. Суворова</i>
Способ исследования онтологических моделей производственных систем на основе применения графовых методик | 553 |
| 11. | <i>Е.Г. Серова</i>
Основные аспекты формирования внешней информационной инфраструктуры современной компании | 562 |

12. *И.Н. Мараканов* 570
Корпоративная автоматизированная система оперативного обмена информацией на основе сервисориентированного подхода
13. *Д.А. Ёлкин, И.А. Минаков, С.И. Вольман* 577
Использование методов статистической оптимизации для улучшения эффективности показа Интернет рекламы
14. *А.В. Якушин, А.В. Кочуров, С.И. Вольман, И.А. Минаков* 586
Разработка системы прогнозирования поведенческих характеристик пользователя в сети Интернет
15. *А.Е. Колоденкова* 593
Задачи программного инжиниринга сложных систем на основе критерия жизнеспособности проекта
16. *Г.И. Горемыкина, М.А. Пономарева* 599
К вопросу о прогнозировании количественных параметров сложных систем при использовании нечёткой информации
17. *А.С. Потапов, И.В. Аникин* 606
Подход к управлению рисками информационной безопасности на основе продукционно-фреймовой модели представления знаний
18. *А.С. Катасёв, Ч.Ф. Ахатова* 615
Нейронечеткая модель формирования баз знаний экспертных систем с генетическим алгоритмом обучения
19. *В.И. Шелехов* 622
Верификация и синтез эффективных программ стандартных функций floor, isqrt и ilog2 в технологии предикатного программирования

Multi-Agent Models and Systems

Мультиагентные модели и системы

1. *G. Rzevski* 633
Using tools of complexity science to diagnose the current financial crisis
2. *А.И. Павленко* 644
Использование агентно-ориентированного подхода к реализации методологии системного анализа
3. *А.Д. Тарасов* 650
Система физической защиты на основе агентно-ориентированного подхода и нечеткой логики
4. *А.В. Иващенко, У. Инден, П.О. Скобелев, И.А. Сюсин, А.В. Царев* 657
Метод адаптивного планирования в мультиагентной системе моделирования работы наземных служб аэропорта
5. *М.В. Андреев, А.В. Иващенко, Д.М. Мартышкин, П.О. Скобелев, Л.В. Уланова, А.В. Царев* 670
Применение конструктора онтологии в мультиагентной системе динамического планирования персональных задач для мобильных пользователей

6. *А.В. Вайсблат, А.Р. Дязитдинова, А.В. Иващенко, П.О. Скобелев, А.В. Царев* 678
Организация интерактивного взаимодействия в мультиагентной системе
управления транспортно-экспедиционной компанией
7. *А.Р. Дязитдинова, А.В. Иващенко, Е.В. Симонова, П.О. Скобелев, М.В. Сычева,
И.И. Хамиц, А.В. Царев* 686
Концепция мультиагентной системы интерактивного построения программы
полета и планирования грузопотока Международной космической станции



Plenary Papers
Пленарные доклады

PRODUCT LIFE AS A COMPLEX SYSTEM

G. Rzevski

Professor, Complexity Science and Design Group, The Open University, UK
Founder, Emergent intelligence Technology, Inc. FL, USA
george@rzevski.net; www.rzevski.net

Keywords: *complexity, self-organisation, emergence, evolution, adaptation, multi-agent technology, product life, product reliability, product availability*

Abstract:

На основе понятий и средств науки о сложности предложен новый метод для исследования пригодности таких продуктов, как самолет и гоночные автомобили. Модель принимает во внимание взаимодействие неисправности элемента, воздействие определенных особенностей окружающей среды на работоспособность индивидуальных продуктов и различные процессы старения составляющих компонентов.

Introduction

Component failures are caused by:

- Design/manufacturing errors
- External impacts
- Wear and tear
- Fatigue

There are several aspects of failures that are not properly understood at present and not taken into account in reliability assessment.

Perhaps the most important aspect is that failures interact with each other and should not be analysed in isolation. The importance of failure interaction is exemplified by the recent crash of an Air France Concorde airliner: allegedly, during the take-off, the plane hit a small piece of metal left on the runway, its tyres disintegrated and damaged the fuel reservoir; leaked fuel caught fire.

It is also important to take into account that product life critically depends upon environmental conditions in which the product operates and therefore every individual product has its own specific life cycle. Airliners that fly regularly to Middle East and are exposed to effects of sand will have different failure history from airliners that fly mostly to the North and are exposed to severe frost. It is not meaningful to calculate reliability of a product type; it must be assessed for each product.

Further failure aspect ignored by current reliability assessment methods is that wear and tear, as well as fatigue, differ for different types of product components and consequently the overall product reliability changes in time. Product life is subject to evolution. Reliability is a time-variant property.

A brief outline of the proposed conference presentation is given below. The emphasis will be placed on developing complex adaptive product life models and on real-time servicing and spare parts supply scheduling.

Complex Adaptive Product Life Model

To increase product availability there is a need to link

- Product design
- Failure prognostics
- Failure diagnostics

- Servicing and repair scheduling and
 - Spare parts supply
- to a single Product Life Model.

This is particularly important for aircrafts and vehicles for which it may be far more cost-effective to provide preventive maintenance and repair at locations where failures occur, or are anticipated to occur, rather than at their home base.

A comprehensive Product Life Model takes into account different life spans of different groups of product components, their failure-related interactions and salient features of operating environments in which products operate. This implies that the proposed Product Life Model has the following characteristics:

- Each individual product has a different life model.
- Product life models are time variant.
- Product failure is an emergent property – it emerges as a result of unpredictable events and failure interactions. Although failures are unpredictable, failure patterns can be discovered by dynamic data mining of failure data.
- Product availability evolves in time reflecting different aging characteristics of different groups of constituent components.

The Product Life Model is a basis for simulation aimed at using discovered Failure Patterns to estimate capacity of required Servicing and Spare Part Resources for the given product.

Real-Time Servicing and Spare Parts Supply Scheduling

During the product operation (e.g., for aircrafts during flights, and for cars when they are on the road) it is possible to monitor Reliability Sensors and broadcast the signals to the nearest maintenance and repair shop for signals to be analysed for indications of possible imminent failures.

Based on long-term plans and on results of real time reliability monitoring, it is necessary to perform Real-Time Scheduling of Servicing and Spare Parts Supply with the aim to arrange for servicing personnel to meet the aircrafts/cars at their destination and execute required maintenance or repair, if required. Spare parts will be supplied where and when required.

Under this scheme servicing intervals are dynamic and depend on operating conditions and the conditions of the product; they change as the requirements for maintenance and repair change. Costs of storing spare parts may be reduced dramatically.

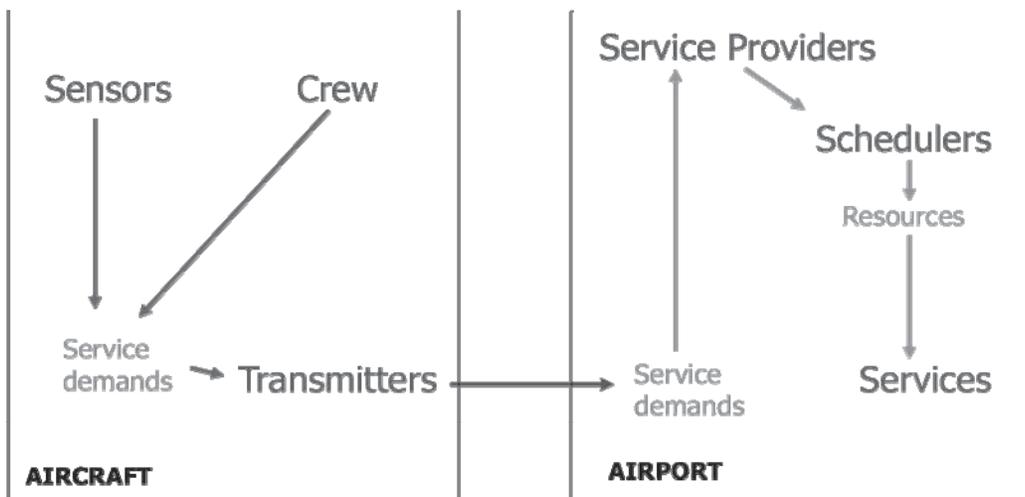


Figure 1 – A Real-Time Aircraft-Airport Servicing and Spare Parts Supply System

**REAL-TIME MANAGEMENT AND COMPLEXITY.
A STRATEGY MODEL FOR REAL-TIME MANAGEMENT
IN COMPLEX SERVICE SYSTEMS**

U. Inden

Cologne University of Applied Sciences
Centre for Knowledge Management and Applications of Intelligent Systems
Claudiusstrasse 1, 50678 Köln, Germany
udo-inden@fh-koeln.de

Key words: *real-time management, complexity, strategy model, service system, self-organization*

Abstract

Рассматриваются проблемы оперативного менеджмента в сложных системах сервисного обслуживания. Описываются особенности и составляющие стратегии «оперативного критического менеджмента», приводящей к появлению в новых системах эффекта самоорганизации.

Any big airport in the world is to be understood as an open complex dynamic system formed by more than hundred different services such as passenger check-in, baggage claim, cargo handling, apron security, aircraft fuelling, or apron control, on its side consisting of many sub-services. Each of them acts with some autonomy and all of them are massively depending on each other. Airports show the emergence of critical incidents within minutes amplifying the normal floor of interferences and requiring managements to answer in real-time.

Obviously there are non-linear dynamics (positive feed-backs, butterfly effects): Local interferences have the potential of growing into system-wide failure, driven by interaction along service supply-chains and flight trajectories, by coincidence with other small local failures etc. And developments at airports are sensitive to history: They do not repeat (or as an experienced COO of an airport said: 'Every day is different and there are many one-of cases').

Actually aviation systems are well planned and the objective of operations management is to maintain these plans – if necessary immediately adapting to unexpected events by unplanned but nevertheless coordinated action. I.e. are airports also to be understood as self-organising systems? Talking to managers and dispatchers supports the notion that their self-understanding differs from practical doing in daily business: There is hardly another system but airports (or aviation in total) more marked by extensive planning and extensive "improvisation" at the same time. I.e. experienced agents in airport business define their proficiency as 'ability achieving objectives irrespective of the unexpected'.

From this point of view the challenge is 'domesticating the beast'. In fact 'noise floor' and 'critical incident' mean that there is a continuous risk that the system gets out of hand. Criticality is a property emerging from current operations hardly to be traced back to a concrete event. In this case the current plan is ruined and dispatchers have to provide a new one - now. Thus criticality emerges counteraction. Planners or adapters, hen or egg – that's a philosophical discussion: Dynamics forces them to act and new plans feed back into dynamics – hopefully in a 'domesticating' way. In fact proficient operations systems are distinguished by the proficiency of self-regulation.

But as mentioned it is a 'social system' meaning that coordinated counteraction in real-time requires the exchange of relevant information in real-time and on any appropriate level of detail and acting i.e. down to subservices and individual resources included. For mitigating a critical situation

one particular resource may have to be re-allocated. But for making this decision local knowledge is to be explored across all relevant operations. Self-regulation relies on the exchange of local information – based on a common understanding of principles and scene. They need to effectively share reality.

Given that many providers of services collaborate and compete simultaneously, a kind of prisoners' dilemma appears: Should information be shared and if, which information? What is the balance between disclosing and non-disclosing? How is it determined? Actually it is a flux-balance; a dynamic equilibrium, being re-defined for the individual case (a particular collaboration) or the operations pattern (e.g. a stress-level of the system). I.e. the balance changes in daily business but also because of changes of general operational patterns, finally of break-through changes. Mostly new technologies and new ways of thinking, or external events affect the balance. E.g. 9/11 or SARS changed rules, behaviour, and culture of exchange, as or currently the Icelandic volcano Eyjafjalla does.

However dispatchers learn that they can trust in the behaviour, the performance etc. of other agents – or not. They may learn that exchanging information is advantageous for both sides and thus may agree in new standards. And they may learn that sharing knowledge also may pay out by new ideas, new options to differentiate from competition – i.e. that the prisoners' dilemma is just a fiction and that competition even requires to collaborate. In the following a strategy featuring real-time criticality management (RTCM) will be described:

- identification respectively anticipation as well as evaluation of critical incidents, i.e. evaluate risks, likelihoods of occurrence and expectation values accordingly management thresholds,
- 'improvisation', i.e. finding and proposing B-plans able to meet objectives despite of incidents, search knowledge bases for solutions matching patterns of operations,
- control of the execution of non-standard activity (improvisation), i.e. avoid sparking,
- learning from cases of success or failure, adapting defaults in planning to current operations' patterns, and able of supporting self-improving organisational structures,
- growing a "treasury" of exemptions by feeding back into online and offline knowledge bases and human-resource development,
- exploiting network effects of knowledge¹ from collaboration, likely the most powerful but alas a very controversial driver of improvement.

The Situation: Congested Infrastructure and Action Taken to Improve

Driven by the dynamics of the global economy since decades air traffic shows a stable growth of about 3.5 % per year, tucking away events like 9/11, or Sars. Sooner than current global economic crisis will be overcome, travelling and shipping may return to this path². But since by this traffic is expected to double faster than it takes to develop new infrastructures the ATS is going to stall. Under stress even small failures may widely propagate causing subsequent ones, air-streets, airports, or traffic management capacity will be blocked and finally also impair European Economies.

In Europe problems concentrate in the area between London, Paris, Berlin, Rome, with limited airspace and a concentration of major hubs. Here – because of impact on the natural environment or the live in urban areas – new infrastructure developments meet increasing social and political opposition changing side-conditions of operations and turning air-traffic into a multi-context environment with contradictive objectives. Obviously overcoming inefficiencies in air-traffic management systems

¹ The focus here is on fostering the development of new knowledge rather than just disseminating existing knowledge. Prominent examples are competence networks in medicine. Because of their competitive impact they need particular focus as described in chapter 1.1.4 and WP 300. [e.g.: www.knowledgecommunities.org/cops.htm].

² On the other side the current crisis drives structural reductions of airline revenues and by this an accelerated shake-out in the market and a potential concentration of business likely influencing the management of air-traffic systems in general.

is the most promising approach. This is the focus of the European lead program SESAR and of a large number of related projects, addressing improvements of ATM from all relevant angles.³ Objectives can be summarised as follows:

- The chief-directive is to significantly increase quantity and quality of transportation services as well as utilisation of air-space, airports, aircrafts or any other resource committed to the system the minimise and reduce costs and time expenditure as well as minimise external effects of air-traffic per flight, passenger or shipment.
- Execute flight-plans effective – the major means of coordination – as reliable respectively in case of failing recover them as fast as possible.
- Improve information quality on the level of tactical planning e.g. by taking advantage from new technologies or improved forecasting methodology.
- Maintain consistency of information used for tactical decisions by implementing a system-wide information backbone for all operators in the aviation system.
- Improve local decision making and use of information.
- Enable and encourage collaboration for handling the systemic interdependence of planning and acting on tactical level (avoid problems) and in-situ during execution (react on non-avoided problems) since the reliance of all activity across aviation systems does not allow any isolated approach.
- Establish a culture of continuous collaborative improvement and learning.
- Automize wherever reasonable, improve human-machine interaction and employ intelligent technology for effective real-time management support.

The Challenge: Improve Responsiveness to Critical Incidents

Flight-plans effective are *the* means of coordination since air-traffic is simultaneously managed at many nodes: flight-decks of thousands of aircrafts and their airline operations centres, dozens of air-traffic control centres as well as airport control centres and at each airport by dozens of ground service providers. For maintaining flight-plans the supply chains of all services need to be coordinated “just in time and sequence”. In this system “critical incident” is defined by failing of ATM to avoid

- nonconformity to a flight-plan respectively the inability to recover it,
- propagation of failing across other flight-plans because of insufficient recovery.

Within regulatory frameworks all agents locally act autonomously – while 4D flight-trajectories make them globally interdependent either directly (along the same one) or indirectly (if trajectories cross). Actually the aviation system incorporates two sources of disturbance: They (1) are sensitive to external events like weather or irregular behaviour etc. and (2) emerge internal frictions, disabling response to external events and “staling” air traffic, the more the ATS operates at its limits.

Figure 1 shows how frictions emerge from uncertainty. It describes the point of view of an individual dispatcher: Points ①, ②, ③ represent actual and expected future points in times along the trajectory of a flight. In order to be ready for service at the arrival of a flight the agent uses two parameters:

- the estimated in-block time⁴ ③ which may be updated by the actual time-stamp ①,
- and a time-span, the default response-time (tR) of resources required for the respective job: To be ready for service in ③ the operation has to be started in ②.

³ SESAR = Single European Sky Air traffic management Research.

⁴ Aircrafts are accessible for ground-services when brake-blocks are attached at a parking position.

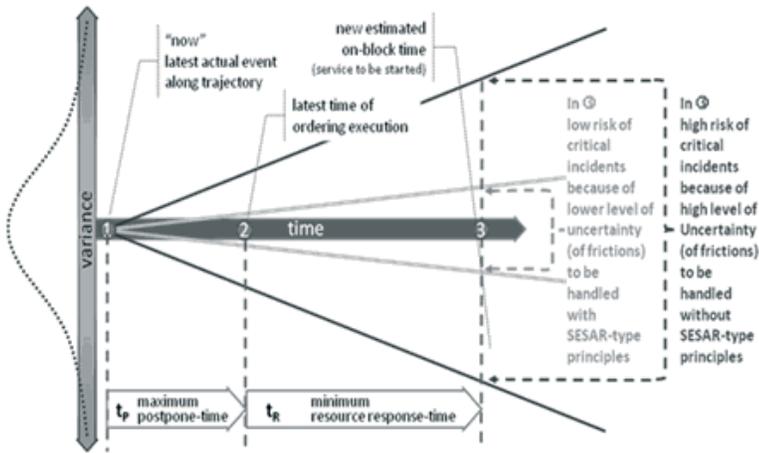


Figure 1 – Information Quality and Uncertainty

Uncertainty for a given t_R is represented by the distribution of possible events exemplified by two scenarios: The green funnel with improved telemetric and information management systems and the red one without. Assumed the update of ③ is immediately transferred at time ① the operator may postpone his order for the time t_P e.g. in order to wait for improved updates of ③.

Chances of meeting objectives across conflicting contexts (security, costs, service quality, resource utilisation, emissions ...) depend on the uncertainty to be ‘digested’ in ③. I.e. the more agents decide on the base of poor information, suboptimal routines and just scanty cooperation and the more the system is under stress, the risk of subsequent mismatches (ripple effects) is high and finally the whole system may fail. The more SESAR-type principles are implemented (accuracy of time-stamps, quality of calculating subsequent estimates, capacity of re-planning in case of updated estimates, but also on resources’ response-time) the more likely the system will be stabilised.

Critical incidents are failures to comply the active flight-plans and regulations. They may occur at any time but particularly when stress amplifies the noise-floor, driven by external events, failing internal resources, or if traffic consumes buffers, in the extreme even disables adaptiveness to small problems so that criticality grows: First there is only a risk observant and experienced dispatchers and service agents can avoid. If this fails it may be at least contained locally. Otherwise the problem will “infect” other services, becoming a problem for a flight, impair other flights etc. I.e. robust service systems require effectively avoiding or breaking chains of propagation. The fundament is that everybody works with integer actual and estimated data on flight trajectories and that everybody makes a proficient use of them.

The Approach to Manage Critical Incidents

Most critical incidents derive from interactions of different sources. E.g. a delayed flight may run into a peak which consumed the flexibility of a service provider driving now the risk of building up even more delays in the rotation of the aircraft. For real-time management a multi-agent system (MAS) is to be implemented as an intelligent add-on to the dispatching system of the respective service and continuously up to date about flights and states of all running and subsequently planned service resources and operations. If information on an estimated delay of a flight arrives the system will analyse the impact of this delay to the efficiency of the service answering questions on two levels:

- Is there an immediate impact, i.e. will the current plan be corrupted by the delay?
- Are there potential risks deriving from the delay?

The second step is called anticipatory logistics⁵: The announced delay it is still an estimated one – i.e. until the aircraft actually will be on-block a lot can happen. Assumed statistics tells that an additional delay will be the most likely incident and that real-time pattern-recognition (context-awareness) and case-based learning tell that in peak-times the particular flight will likely be deviated to an apron stand etc. the MAS will run parallel simulations analysing impacts of potential events filtered by threshold regarding the likelihood of occurrence. Given some likelihood of immediate impact the system will do next steps:

- Economic information is used to analyse expectation value of the impact.
- Above user's threshold the MAS identifies options to avoid or mitigate the problem and
- control execution of active ones to avoid later problems ("controlled improvisation").

To a degree the MAS will act like experienced dispatchers: Improvise, i.e. finding a way to meet objectives "in spite of ...", finding B-plans produced and activated in different ways:

- Low-level action (e.g. "hurry a little") is taken by any committed service agent – if he / she is aware of the situation. So the MAS may advise such action automatically.
- In other cases the system may find a B-plan and propose a new schedule or other action.
- By case-based learning, the ontology will become a "treasury of exception plans".

In the regulated aviation environment even ingenious ideas whether from machines or humans need to be approved. With the time the MAS will dispose in its ontology of an extremely rich information base on operations, problem, and potential critical incidents and how they become real, on finding exception plans and controlling them etc. pattern recognition and if-then simulations will support the development of situation-based exception management plans, subsequently stored in the ontology. I.e. with the time the MAS will be able of choosing from thousands of plans the one answering to the problem to be solved.

In systems of interdependent services critical incidents also derive from defective collaboration and it is likely as well that the efforts to mitigate a problem by one service require the support of another. Instead implementing common centralised planning collaboration on-demand should be enabled depending on the operational context and on rules of collaboration agreed by service providers. Particularly in critical situations face-to-face collaboration is as paramount as capabilities of decision support systems of anticipating and evaluating critical incidents, of providing and controlling valid B-plans and of learning on the fly – e.g. identifying changes in the operational context probably sooner than human supervisors.

The example above may pertain to a caterer suddenly short of highloaders since the two ones waiting for the delayed flight have run out of their buffers, taking the catering of subsequent flights is under risk. So the B-plan may be that operators of at least two subsequent flights are asked whether they agree that their flights are served by one truck only instead of two – with the consequence that the highloader has to reposition from the front to the rear-galley, increasing the time for catering. To catch up with this delay the plan also needs asking gate-agents to agree that the loading of the rear-galley continues while passengers are boarding. This decision among others depends on the type of the stand: At the terminal passengers use gangways and the front door, not interfering with activity in the rear galley. On apron-stands also rear-doors may be used. I.e. there likely are occasions MAS and users can learn from. This induces the following considerations:

- Anticipatory logistics translates into significant advantages of reaction: Degrees of freedom answering to requests depend on time. I.e. the caterer-MAS should approach the MAS of flight gate operations already when the problem has been identified.

⁵ e.g. Noel P. Greis, Jack G. Olin, and John D. Kasarda: The Intelligent Future, page 18 ff, Supply-Chain Management Review, May / June 2004.

- There is a large potential to improve the performance of the whole airport by learning from collaboration.
- Barriers to collaborate are likely to be lowered: Service providers have reasons to not collaborate, at least to not disclose information about internal problems. However as systemic means of collaboration are controlled by agreements and rules and all sides are likely interested in the lessons learned from collaborative exception management.
- There are also loser-/winner scenarios of collaboration and the MAS will not only use but also make suggestions to improve economics balancing the efforts of providers contributing to the overall efficiency of operations.

Finally we have to be aware of the strategic and managerial requirements to take advantage from MAS-based intelligent, anticipatory, collaborative, and learning technology. Improvements of robustness and particularly of competitiveness will not be provided by just implementing a new technology but requires a sound understanding of economics and of appropriate operational and organisational managerial strategies to effectively take advantage from them – not only in terms of improved efficiency (costs, time, utilisation), but also in understanding impacts of intelligent technology to existing and potential new business models.

Basic Proceedings of a RTCM Concept

A Starting-Point: Cause and Reason Analysis

Particularising the RTCM concept and connecting this with the work-breakdown structure and deliverables needs to start with an understanding of the interrelationship of business processes. The following picture gives an overlook how the trajectory of an aircraft respectively the flights it operates are connected to ground service processes. Fig. 2 illustrates a high-level model of service structures on the land-side of the airport for the processes of outbound passengers and baggage, maintenance, outbound cargo and mail.

Via the block-time of the aircraft and the standards of turnaround management all services are interconnected. Each service has a slot to be executed and is required not to interfere with other services around. Within these limitations each service organises itself autonomously.

Each service consists of a number of sub-services representing the level of execution. E.g. passengers need to drop their luggage, each piece has to pass security control, probably in some steps of escalation, will be allocated for the booked flight, containerised, transported, and finally loaded as well as crosschecked with the boarding of the respective passenger who at the same time has made his way through the terminal. Thus the objective is managing trajectories of all objects (passengers, baggage, spare-parts cargo, mail and many more) in a way that they meet the trajectory of the aircraft within its on-block time.

Critical incidents derive from the interaction of uncertainties along all trajectories. Taken flight-trajectories and with this the utilisation of air-space, aircrafts and airport infrastructures to be priorities, ground processes have to adapt – i.e. to manage the uncertainty inherent to their operations to avoid or mitigate critical incidents.



Figure 2 – High-level Model with Exemplary Ground Services

Figure 3 gives an idea of the risk-structure of such a system. The latest actual time measured of the incoming flight ① is translated into an update of the estimate of the in-block time ③. If time ① is in schedule the value of the update is 0. However because of current conditions of air-traffic or taxiing on the airport etc. there is a remaining uncertainty in the estimate, being higher the longer the time between ① and ③.

Dispatchers of services have to call resources for starting service operations in ②, a default span of time before ③. This span depends on the minimum default time of resources to respond to the call starting their job. Besides this it includes some buffer time enabling to react on unplanned events. Such buffers e.g. are included in the time passengers are asked to arrive at the airport prior to the time of boarding. Because of the remaining uncertainty in the estimated time ③ dispatchers are likely to postpone this call until ② in order to wait for another update of ③.

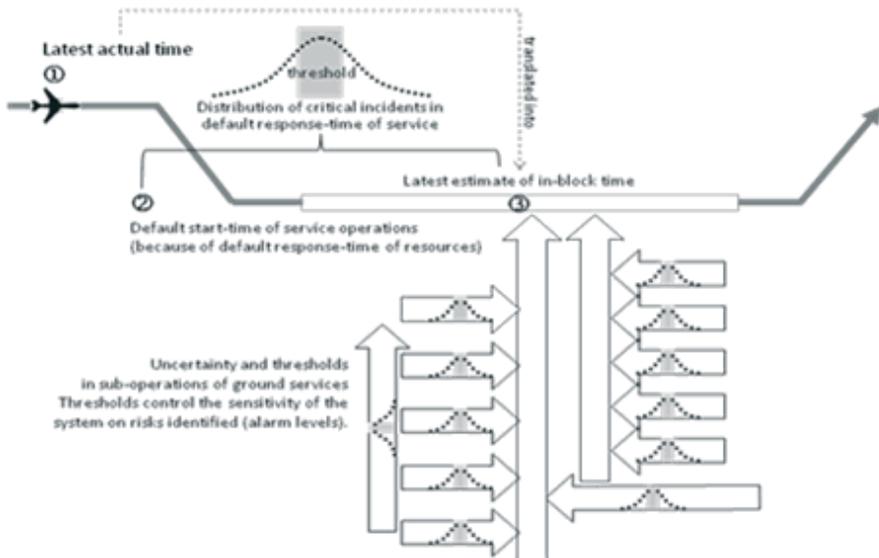


Figure 3 –1 Coupled Risks of Flight- and Service Trajectories

Thresholds (alarm levels) are required for managing the sensitivity of dispatching to risks identified. I.e. dispatchers (respectively automatic systems) may not react on risks with a likelihood of occurrence lower than 30 %. I.e. everything below is considered to be manageable with planned buffers. Such thresholds derive from experience and may be different depending on the “stress” (high noise-floor) in the service system requiring a higher alertness (lower thresholds).

The default response time is calculated back along the chains of services these i.e. translated into latest start-times of subservices. And on that level now more uncertainty and further thresholds re introduced: Transfer passengers may arrive later, queues at check-in, security check, border control may be longer than expected, problems with baggage sorting may appear, crews for gate-services may be blocked by problem of previous flights etc. The following chapters will explain basic procedures.

The Concept of Controlled Improvisation

Avoiding uncertainties of flights and services to multiply (that’s actually the way of stalling airports) is the core idea of the concept. Necessarily this starts with understanding inherent uncertainty (noise). In an initial approach and based on this model the technical implementation of the MAS for “controlled improvisation” may work as follows.

The RTCM will continuously analyse the situation in order to identify potential critical incidents learning from historic data respectively from parallel simulation investigating alternative futures.

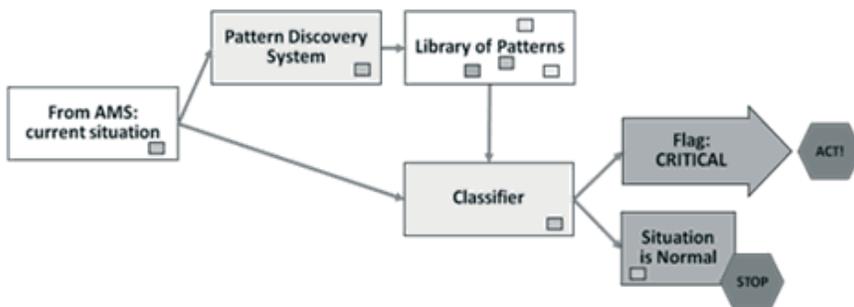


Figure 4 – Identification of Potential Critical Incidents in ART-IMS

There are many scenarios of future. The bold Line represents the plan as active in a particular situation at the airport, i.e. it likely already includes previous updates of the tactical plan of the day. Circles are new unexpected events as provided from airport and service management systems. These events continuously trigger the system to analyse risks with the goal to recognize potential critical incidents. New events may cancel previous and start new threads of concurrent simulation for investigating, identifying, and evaluating critical incidents (colours represent alarm levels: the redder, the more risk).

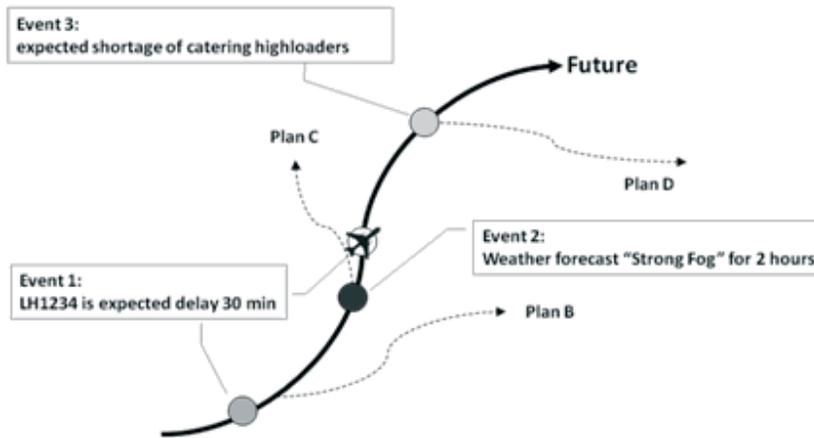


Figure 5 – Investigating Future Flight- / Service-Trajectories (Massive Simulation)

The procedure includes the evaluation of potential critical incidents and the deliver of B-plan(s). E.g. expectation values (likelihood of occurrence times economic or if given also other impact and their trade-offs) of critical incidents identified Ip1 and Ip2 may differ ($Ip1 \gg Ip2$). For calculation the system relies on default operations parameters (time, costs, fuel consumption etc.) if given corrected by information about current operations which may be provided by the pattern recognition module. Finally expectation values will go into policies (e.g. if resources are short only measures are taken mitigating problems with high EV. This filter influences the economics of the system. In a second step and depending on alarm levels investigations for Plan B are started resulting into a portfolio of options to answer to the risk. The system makes proposal to dispatcher or may decide automatically to switch to new scenario. The response time of resources is part of mitigation planning. Dispatchers can cancel alarm, ask for more detailed investigation, or dispose search for B-plan.

On this stage the RTCM will read exception management procedures from the ontology. Such procedures can be based on recommendations or may appear as a result of dynamic scheduling with overriding constraints and managers will have a chance to tell system what constraints it can override.

- Finding: B-plans in terms of valid exceptional operations defined for particular cases and context are stored in the ontology of the system. If the ontology is lacking appropriate concepts of action (particularly in the beginning the system will not have many B-plans prepared) this lack will be closed by learning (online but primarily offline). Another strategy may be that the future ontologies with a higher degree of sophistication and particularly with a larger richness of details will support finding solutions which are not fully pre-structured. A simple example may be that the system may have access to the information that a loader working in the baggage hall also holds a licence for driving a motorised conveyor belt etc.
- Activating: Also B-plans come with default response times. They just are shorter as those of standard procedures. I.e. there is also a state \odot for starting it at the latest. This start may either be prompted by the dispatcher or started by the system automatically.
- Control: B-plans involve non-standard activity and for this have to be controlled to avoid subsequent problems. This control will employ a particular thread of the MAS and continue until the B-plan is successfully finished. For control the MAS will use the same principles as before: identify and evaluate critical incidents deriving from executing B-plans in the current environment etc.

The Concept of Case-Based Learning

Case-based Learning can be provided on Individual and Collaborative Levels. The proposed technology will provide a wide spectrum of opportunities for machine-based as well as for human and organisational learning (the learning / intelligent enterprise).

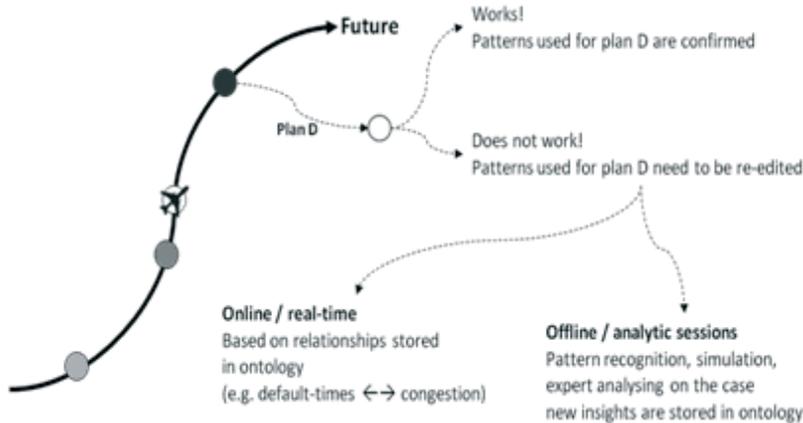


Figure 6 – Online and Offline Learning from Successful and Failing B-plans

In the first instance a rich database of recorded operations and mitigating action will grow rapidly. This data will feed into the pattern recognition module of the MAS real-time (online) e.g. for learning about appropriate values of the defaults used for analysis and B-plan development.

It is crucial to see that there is no reason to learn as long as everything is in plan (more likely it will cause illusions to be safe). Thus critical incidents represent the chance of learning about the dynamics of the operations system. Therefore systemic measures of learning from mitigation as well as from failure are paramount. E.g. without large efforts the MAS automatically will learn online (“on the fly”) from controlling the execution of a B-plan that default times assumed in the plan are too short given the currently very busy airport environment. I.e. in the next plan the system will calculate some more time e.g. for opening another counter for dropping oversized baggage. Such obvious relationships (default times \leftrightarrow level of activity) will be pre-defined in ontologies or the MAS may learn from parallel pattern recognition.

In a second instance records are used offline for analysing historic operations data by means of pattern recognition, combined with simulation and human reasoning:

- The daily de-briefing meeting on yesterday’s operations is a first step of offline analysis. Here managers from operations departments of the airport and likely most service providers around discuss whether there is reason to learn from recent incidents and to take action on them. An example of action may be that managers ask for a deeper analysis of the data in order to find hidden patterns behind problems. Or they may decide on results from analysis e.g. by changing rules of operations or arrange for training of staff.
- Another step is that e.g. managers from operations and quality management department develop a new B-Plan enriching the ontology with a new option of mitigating critical incidents. Actually there is a large potential to grow experience
- Finally operation standards derived from such analysis will go into quality manuals and knowledge management systems. And in a similar way the discipline of learning from mitigation and failure, from continuous analysis of data, from developing and re-editing B-plan ontologies etc. also will be documented – as case-based simulation and pattern analysis will become a means of training of management and not at least also of staff.

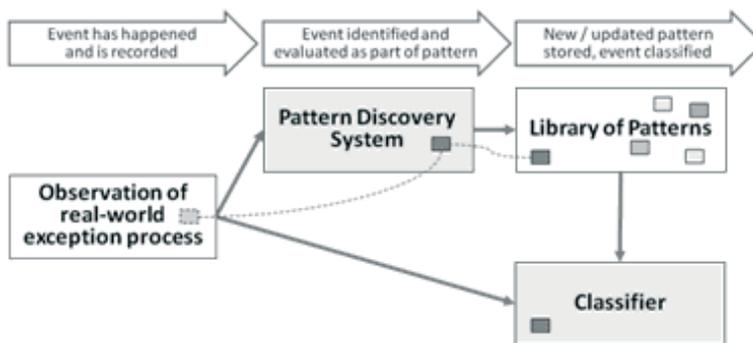


Figure 7 – Discovering new Patterns and Classification of Situations

For automatic discovering of problems pattern discovery and classification methods and other toolsets are employed and if need be complemented by offline-analysis sessions. The situation at airport is constantly under consideration and is analysed by pattern discovery toolset which generate new patterns when possible (significant difference with existing patterns) and add new patterns in library of patterns or makes changes in structure of patterns (3). In parallel new situation is analysed by classifier which detects what pattern is most close to current situation and if situation is close to pattern of dangerous situation it will generate special signal (flag) and activate procedure of exceptional management (usage or generation of Plan B).

The Concept of Autonomous Escalation (Collaboration on Demand)

Mitigating critical incidents along flight trajectories like significant delays of incoming flights may be not possible by involving just one service and on the other side an individual service may run into problems which cannot be mitigated internally i.e. will at least require to inform other services about the situation respectively to ask them for support.

For the further description an example may be useful, which has been developed in the context of air-catering supply chains. It is based on knowledge stored in the ontology the system about typical business operations sequences: catering, passengers loading to aircraft (embankment), etc. (standard = all services in cabin should be finished before embarking). But there is an option to do catering in parallel with embankment of passengers (in ontology it should be described as possible in critical situation)

Example: Assumed that catering by default employs two highloaders per flight in order to cut turnaround time. Now two of them are waiting at a stand for a delayed aircraft of flight 1 (F1). In the moment preparation for service for this flight was started the expected delay was not in conflict with the buffer times of the highloaders but because of another delay in taxiing they run out of their thresholds and now are lacking for serving subsequent flights. Returning the highloaders and abandoning service for the delayed aircraft is no valid option – but because of peak-time and more delays there now the trucks for a flight scheduled 40 minutes later (F2) are lacking and any replacement is available.

Now plan B is to serve this F2 and a third flight F3 with only one instead of two high-loaders – with the consequences that: the operation will

- require an OK from the operating airline,
- take more time because either the loader has to move trolleys for the second galley through the aisle or to reposition the truck to the second galley after having finished the first one. Therefore action is needed to avoid a delay of the flight.

Assumed that the Airline will agree in the plan – either because it is also the operator of F1 or because they know from a professional perspective that they also may have to rely on the OK of another airline in similar future cases. However the OK is not given.

The B-plan tackling the second problem is to parallelise the operations of galley loading and passenger boarding: Therefore the dispatcher calls the supervisor of gate operations (or e.g. the ramp-agent of the airline) to agree that the rear-galley is loaded while passengers are already embarking. The OK given also from this side the caterer may be able.

- doing the jobs for F1, as well as in time for F2 and F3
- i.e. has compensated the lack of 2 trucks

Actually the OK from gate services may be harder to get than that of the Airline: On the gate, during boarding etc. operational uncertainty may deliver reason to abandon the OK. This may mean that B-planning of gate-services interferes with the B-planning of the caterer. Theoretically in the given scenario of autonomously planned services with private spheres there are two solutions:

- The caterer has to accept a No from one of the gates and thus to search for another opportunity.
- Both sides will have to search for a compromise i.e. MAS of both sides have to negotiate their trade-offs and finally to find a balance.

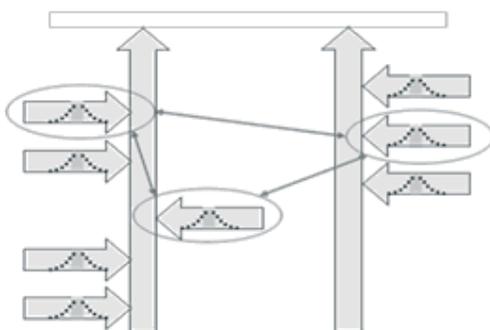


Figure 8 – Interacting Sub-Services of Different Providers

The example shows how interdependencies between services and sub-service level (i.e. the transportation department of the caterer needs to understand boarding processes, v.v.) can be exploited for increasing the robustness of plans. It relies on factors and levels of professional relationships between service providers, rules of conduct, or applicable strategies compensating each other for efforts supporting another services – to be investigated in the economic workpackages and integrated into the decision-making logic of the MAS.

In a wider view a comprehensive understanding of the architecture of collaboration on the level of airport operations is required – as it is now and as it should be in future (picture 14). Currently the pre-tactical and tactical flight plan is the major means of coordination across services. I.e. each provider plans and executes individually and coordination only takes place when services physically meet at a/c-stands, primarily prompted the ramp-agents of the airline if a problem occurs. There is no pro-active collaboration in the phase of planning before.

If collaborative controlled improvisation shall work as many services as possible should be equipped with ART IMS solutions, connected by the respective service bus system. At least it will be a topic of research to understand how the overall effectiveness (meeting objectives) respectively efficiency (in terms of cost, time, utilisation, pollution ...) depends on the number or also the type of providers equipped and participating in collaboration.

There are reasons to be equipped: The common goal of all stakeholders (customers, airlines, airport, ground service providers, air traffic control, social and natural environment) is to not delay flights since all will suffer from subsequent problems. Flight-plans are ‘holy’ because they are results of enormous efforts to insure the operability of the whole system in the air and on all grounds. I.e. there are substantial common interests of all parties in the system, also among competitors.



Figure 9 – Equipped and Non-equipped Operators in a Collaborative Environment

Best “candidates” for automatic exchange are data on B-plan operations based on collaboration. In the case described in the chapter on collaboration the caterer, the airlines as well as passenger gate agents are very much interested in the results of their coordinated activity, although they may draw different conclusions from it. On a deeper level managements may be interested in learning about the interdependency of service operations in critical situations. In standard operations all services together may form a good disciplined team. But if this team has to collaborate in adapting to critical situations more than “just” discipline is required. From these considerations and the example above primary contributions of “autonomous collaboration” can be derived:

- It enables significantly faster reaction because information is already exchanged before critical incidents have materialised. In the example the gate team wins time to prepare for boarding in parallel with the loading of the rear galley. E.g. this like is of particular advantage at apron positions where passengers also may board through the rear door.
- To the profit of all stakeholders service plans clearly become more robust because more degrees of freedom can be exploited – or more precisely: because there is a system improving the effectiveness of service specific dispatchers widening their horizons.
- It professionalizes relationships between stakeholders in the service network, also between competitors by introducing clear rules of collaboration which can be learned from experience in simulation as well as from real operations.
- Proficiency also increases because effects of collaboration in terms stability, costs, utilisation ... of operations, i.e. that synergies as well as dys-synergies become transparent and because “mechanism” allow for compensation for downsides.
- Actually it is a serious managerial issue to decide on the disclosure of operational data, but the system not only allows to control the exchange but also to link it to benefits of collaboration.
- Finally RTCM helps to overcome the human factor and instead to come to the economic terms of collaboration.

ПРОБЛЕМА ИНТЕЛЛЕКТУАЛЬНОЙ ПОДДЕРЖКИ ОПЕРАТОРА В ЭРГАТИЧЕСКИХ СИСТЕМАХ УПРАВЛЕНИЯ

Е.Д. Теряев, К.В. Петрин, А.Б. Филимонов, Н.Б. Филимонов

Институт машиноведения им. А.А. Благонравова РАН
101990, Москва, Малый Харитоньевский пер., 4, Россия
kpetrin@yandex.ru

тел: +7 (499) 135-35-34, факс: +7 (499) 135-32-56

***Ключевые слова:** эргатические системы управления, интеллектуальная поддержка оператора, когнитивная психология, гибридный интеллект, агентный подход, интеллектуальное планирование, сценарии, человеко-машинный интеллект*

Abstract

The problem of intellectual support of operator work in complex ergatic systems is considered. The cognitive aspect of the given paper is presented. The urgency of the hybrid intellect principle is justified. The perspectives of agent approach use shown. The methodology of the intellectual designing is stated. The conception of scenario designing of the operator's actions is formulated. The questions of design of the intellectual human machine interface are discussed.

Введение

В сфере автоматизации технических и производственных систем исключительно важную роль играют эргатические системы управления (ЭСУ). Общую тенденцию развития ЭСУ характеризуют следующие факторы: возрастание степени сложности объектов управления; повышение степени потенциальной опасности внешней среды; расширение состава и уровня сложности решаемых функциональных задач; рост объема обрабатываемой информации; сокращение времени на принятие решений; возрастание цены ошибок, правильности и качества принимаемых оперативных решений и др. В условиях действия данных факторов работа человека-оператора (ЧО) чревата сенсорными, эмоциональными и интеллектуальными перегрузками. В связи с этим одна из важнейших проблем разработки высокоэффективных ЭСУ заключается в оптимальном распределении функций между оператором и техническими средствами с учетом психофизиологических возможностей человека. Однако решение данной проблемы выходит за рамки возможностей традиционных систем автоматизации. Единственный выход из данного положения состоит в обеспечении интеллектуальной поддержки действий оператора средствами вычислительной техники.

В докладе обсуждаются научно-методические основы создания систем интеллектуальной поддержки операторов (СИПО) сложных эргатических систем. Поскольку в данном классе систем центральную роль играет эргономический аспект, то их исследование требует комплексного подхода на базе современных достижений в ряде научных дисциплин - в психологии, эргономике, информатике и теории искусственного интеллекта.

1 Психологическая теория деятельности

Для задач автоматизации эргатических систем важное значение имеют инженерно-психологические исследования деятельности оператора. В отличие от механической схемы «человек - машина», согласно которой человек и машина трактуются как однопорядковые элементы системы, в инженерной психологии базовое значение придается деятельности опе-

ратора - функционирование эргатической системы подчиняется той цели, которую он перед собой ставит [6].

Деятельность является формой активного взаимодействия, в ходе которого человек целесообразно воздействует на объекты окружающего мира. Компонентами деятельности являются [5]:

- мотивы, побуждающие субъект к деятельности;
- цели, как прогнозируемые результаты этой деятельности;
- действие, как часть деятельности, имеющая вполне самостоятельную, осознанную человеком цель;
- операция, как исполнительная единица деятельности, способ выполнения действия.

Психологическая структура операторской деятельности включает следующие составляющие: восприятие и оценка условий, в которых действует оператор; планирование последовательности действий; поддержание активного внимания; применение волевых усилий; извлечение из памяти необходимой информации; наблюдение за ходом деятельности, контроль степени результативности действий.

Деятельность человека-оператора состоит из следующих этапов:

- 1) Прием и восприятие поступающей информации, где выполняются следующие основные действия: обнаружение сигнала; выделение наиболее важных из них; расшифровка и декодирование информации; построение предварительного образа ситуации.
- 2) Оценка и переработка информации: запоминание информации; извлечение из памяти нормативных информационных образцов; декодирование информации.
- 3) Принятие решения (во многом зависит от «энтропии множества решений», т.е. от имеющихся альтернатив).
- 4) Реализация принятого решения путем воздействия на исполнительные органы, которая во многом зависит от готовности оператора быстро, на уровне автоматизма выполнять сложные действия в экстремальных условиях.
- 5) Проверка решения и его коррекция.

В процессе деятельности оператор выполняет следующие функции: сенсорные - прием информации, интеллектуальные - восприятие, оценка и переработка информации, волевые - принятие решения, эффекторные - реализация решения. Психика, психические процессы являются регуляторами деятельности, посредством которых оператор направляет, контролирует и корректирует свои действия в соответствии с поставленной целью. Реализация и контроль осуществляются циклически.

Задача проектирования деятельности оператора является главной задачей инженерной психологии. Поэтому в проектируемой СИПО существенную роль должен играть инженерно-психо-логический аспект. Но не только - искомые проектные решения должны также базироваться на современных компьютерных технологиях, методологии и инструментах искусственного интеллекта. Таким образом, проблема интеллектуальной поддержки ЧО неизбежно порождает широкий круг вопросов из различных отраслей знаний, которые необходимо рассматривать на единой теоретической платформе. Такой платформой является когнитивная парадигма.

2 Когнитивный аспект исследования эргатических систем

Когнитология (когнитивная наука) - междисциплинарное научное направление, объединяющее теорию познания, когнитивную психологию, нейрофизиологию, когнитивную лингвистику и теорию искусственного интеллекта. *Когнитивная психология* - это наука о человеческом разуме: мыслях, умозаключениях, языке, памяти и роли сенсорных стимулов. Она возникла в 60-е годы 20-го века в противовес бихевиоризму и возвратила в предмет психологии субъективный аспект.

Когнитивная психология во многом основывается на аналогии между преобразованием информации в компьютере и осуществлением познавательных процессов у человека. Когнитивная система человека рассматривается как система, имеющая устройства ввода, хранения, вывода информации с учетом ее пропускной способности. Такой подход называют *компьютерной метафорой* в когнитивной психологии.

В когнитологии в целом и, в особенности, в когнитивной психологии ключевыми являются понятия репрезентации и схемы.

Ментальные *репрезентации* есть внутренняя, субъективная модель внешнего мира. Информация, полученная от органов чувств индивида, хранится в его памяти в виде абстрактных репрезентаций. *Схема* - это такая категория ментальных структур, которая организует конфигурацию данных, т.е. хранит и организует предшествующий опыт индивида и управляет дальнейшим его восприятием и опытом.

И физический мир, и мышление структурированы. Ментальная система репрезентирует структуру мира, но не во взаимно однозначном соответствии.

Сценарий - это упорядоченная последовательность действий, разворачивающихся в некотором пространственно-временном контексте и подчиненная некоторой цели. Сценарии определяют агентов действий (actors), сами действия (actions), сопряженные с достижением целей в определенных обстоятельствах. Сценарий состоит из слотов и требований, чем эти слоты могут быть заполнены. Иными словами, сценарий точно определяет роли, а также обязательные и факультативные действия. Для каждого из слотов имеются «значения по умолчанию» (default values), которые подразумеваются, если агенты, действия и объекты действий не определены для данного конкретного контекста.

Сценарии репрезентируют события, протекающие во времени и пространстве, одни акты следуют за другими, поэтому отдельные элементы сценария - акты - связаны друг с другом пространственно-временными и каузальными связями.

Подобно самим событиям реальности, сценарии структурированы, т.е. они состоят из подсценариев (subscripts) или сцен (scenes). Каждая сцена распадается на ряд последовательных действий. В каждой сцене есть центральные, самые важные акты. Главные акты каждой сцены сопряжены с подцелями.

Итак, к когнитивным процессам относят память, внимание, восприятие, действие, принятие решений и воображение. Ключевую роль в их понимании играют понятия репрезентации, схемы и сценария. Когнитология позволяет описывать поведение ЧО в терминах информационного потока или функционирования и поэтому является мощным инструментальным средством исследования эргатических систем.

3 Адаптация и принцип гибридного интеллекта

В простых случаях машина обеспечивает ЧО информацией о реальной обстановке и осуществляет трансформацию его действий в сигналы управления исполнительными механизмами. Для многих видов ЭСУ основными являются ручной и полуавтоматический режимы управления. При этом предъявляются высокие требования к операторам: к их сенсомоторным реакциям, тренированности и координации движений, способности мгновенно ориентироваться и принимать решения в критических ситуациях и автоматически выполнять эти решения. На более высоких уровнях автоматизации ЭСУ компьютер замещает человека при выполнении ряда функций. Более гибкое взаимодействие человека и машины обеспечивается в адаптивных ЭСУ.

Свойство *адаптации* ЭСУ заключается в приспособлении к изменяющимся условиям работы как внутри самой системы, так и по отношению к внешней среде, в изменении режима функционирования в соответствии с новыми условиями. В адаптивных эргатических системах

осуществляется не жесткое, а гибкое распределение информационно-управляющих функций между человеком и машиной.

До недавнего времени свойство адаптации эргатических систем реализовывалось благодаря приспособительным возможностям человека, гибкости и пластичности его поведения, возможности его изменения в зависимости от конкретной обстановки. В настоящее время на повестку дня ставится вопрос о реализации механизмов адаптации средствами компьютерного интеллекта: интеллектуальная информационно-управляющая среда ЭСУ должна изменять свои параметры и функции в зависимости от текущего конкретного психофизиологического состояния человека и показателей эффективности его деятельности.

Однако, машина не может конкурировать с человеком в слабо структурированных ситуациях, поэтому в сложной и изменчивой среде функционирования ЭСУ оператор неизбежно будет являться ведущим звеном в управляющем тандеме «человек - компьютер». Но одних лишь сенсомоторных и психофизических кондиций и навыков оператора недостаточно для эффективного управления - необходимо также учитывать и использовать его интеллектуальные ресурсы. Поэтому эволюционирование ЭСУ неизбежно будет основываться на принципе *гибридного интеллекта* [1] - симбиотической интеграции функциональностей искусственного (компьютера) и естественного (оператора) интеллектов.

4 Агентная парадигма интеллектуализации эргатических систем управления

В общем арсенале средств интеллектуализации автоматизированных систем все большее признание получает агентный подход [10, 2, 14, 17], который представляет несомненный интерес и для эргатических систем. В докладе обсуждаются перспективы применения мультиагентных технологий к данному классу систем.

Многоагентные системы интегрируют в себе самые передовые достижения в области программного обеспечения, систем искусственного интеллекта и распределенных информационных систем, демонстрируя принципиально новые качества самоорганизации и интеллектуального поведения. В рамках мультиагентного подхода система строится как сеть агентов. Под *агентом* понимается программно-аппаратная или программная сущность, которая может принимать информацию из внешней среды и реагировать на внешние возмущения. Агент способен автономно решать возложенные на него задачи, адаптироваться к изменениям во внешнем окружении, а также общаться с другими агентами для достижения глобальных целей.

Программные агенты ЭСУ - это компьютерные программы, которые исполняются асинхронно в соответствии с предписанным им целенаправленным поведением. В системах с агентной архитектурой задачи разбиваются на типовые подзадачи, выполняемые командой программных агентов. Каждая из них инициируется либо периодически, либо при возникновении определенных ситуаций. В ЭСУ с агентно-ориентированным программным обеспечением осуществляется гибкое распределение ролей в процессе управления между оператором и агентами.

Агентный подход органично сочетается с идеей гибридного интеллекта. Воплощением его является мультиагентная архитектура компьютерной компоненты гибридного интеллекта. В итоге эргатические функции рационально распределяются между ЧО и сетью специализированных агентов, которые дополняют сенсорные и моторные возможности человека, выполняют роль его ассистентов при решении задач фильтрации, селекции и интеграции информационных потоков, принятии оперативных решений и реализации управляющих воздействий.

Агентные решения не укладываются в рамки традиционных проектных решений. В связи с этим необходима разработка агентно-ориентированной методологии проектирования ЭСУ. Ее основу должны составлять новые системотехнические принципы:

- 1) виртуализация функциональной структуры системы;
- 2) делегирование функциональных задач агентам;

- 3) ролевая специализация агентов;
- 4) структуризация ролей агентов, гарантирующая достижение целей;
- 5) агентная организация параллельных процессов реального времени и др.

Агентный подход применим ко всем базовым функциям ЭСУ. В принципе, каждая из функциональных задач автоматизируемой системы может решаться на агентной основе. При этом комплекс таких задач распределяется между организованными в сеть агентами-исполнителями, а поток задач управляется агентами-менеджерами. Сенсорные и эффекторные агенты обеспечивают взаимодействие системы с внешним миром. Другие агенты осуществляют управление информационно-измерительными каналами, процессами, ресурсами, работами, выполняют адаптивную обработку информации. Интерфейсные агенты поддерживают человеко-машинное взаимодействие.

В отличие от традиционной методологии проектирования «сверху вниз» для мультиагентных систем применим подход проектирования «снизу вверх»: вследствие взаимодействия агентов образуются логически связанные устойчивые структуры, которые могут стать компонентами более сложных возникающих структур и приводить к иерархии возникающих структур и их масштабируемости. При этом агентная архитектура расширяет функциональность проектируемой системы: появляются новые функции, к примеру, функция саморганизации/реорганизации самой системы.

5 Интеллектуальное планирование

Интеллектуальное планирование (AI Planning) - это область искусственного интеллекта, которая в настоящее время привлекает значительный интерес [10]. Основные области его применения - интеллектуальные агенты, автономные роботы и беспилотные аппараты.

Планированием называется поиск последовательности действий, ведущих к намеченной цели. Формально, интеллектуальное планирование основано на аппарате математической логики, а рассуждения, необходимые для формирования плана, сводятся к логическому выводу.

Выполнение всякого плана окружено контекстом, в котором этот план выполняется. Модель контекста, в котором выполняются действия, называется *моделью мира* (world model).

Модель мира может включать статическую и динамическую составляющую. К статическим элементам относят модели объектов мира и отношения между ними. Динамическим элементом является модель изменения мира по своим внутренним законам, независимо от воздействия со стороны субъекта, исполняющего план. Действия плана и динамика мира осуществляют преобразование модели мира (как правило, только статической составляющей). Про изменение модели мира говорят, что модель мира переходит в новое *состояние* (state). Состояние модели мира в момент, непосредственно предшествующий моменту начала выполнения плана, называется *начальным состоянием* (initial state) задачи планирования.

Конструктивным элементом для процесса планирования является *действие* (action). Оно определяется как кортеж:

<Name, Precondition, Effect>.

Здесь Name, Precondition и Effect - соответственно имя, предусловие и эффект действия. Предусловие описывает условия, при которых возможно выполнение действия, а эффект действия - описание изменений, которое действие производит.

Задача планирования заключается в поиске *плана* - последовательности действий, применение которой в начальном состоянии модели мира, приведет к целевому состоянию. Программное средство, осуществляющее решение задачи планирования, называется *планировщиком* (planner).

В зависимости от условий, в которых происходит процесс планирования, выделяют несколько разновидностей *сред планирования* (planning environment). Они определяются рядом свойств исполнителя и мира, в котором он оперирует:

- world observability - наблюдаемость мира (полностью или не полностью);
- world static character - статичность мира (статичен или динамичен),
- action determinancy - детерминированность действий исполнителя (детерминированные или нет),
- action durability - длительность действий (длительные или дискретные),
- action parallelism - параллелизм действий (допустим или нет).

В *классической среде планирования* мир полностью наблюдаем и статичен, а действия детерминированы, дискретны и недопустимо их параллельное исполнение. Планирование в такой среде называется *классическим планированием*. Все остальные виды сред называются неклассическими, а планирование, соответственно, неклассическим.

Цель в классическом планировании задается при помощи одной формулы исчисления предикатов первого порядка. Как правило, для простоты цель задают в виде конъюнкции литералов. Предусловие и эффект также описываются формулами исчисления предикатов первого порядка. Пример описания действия «погрузить объект objA на грузовик truckB, если объект и грузовик находятся в складском помещении depC»:

action: LOAD-TRUCK

precondition: InDepot(truckB, depC) & InDepot(objA, depC)

effect: \neg InDepot(objA, depC) & InTruck(objA, truckB)

Параметризованное описание действий осуществляется посредством *схемы действия* (action schema) - абстракции действия, в которой ряд объектов модели мира, использованных в описании действия, заменяется параметрами. В процессе планирования схема действия конкретизируется. То есть вместо параметров схемы подставляются объекты модели мира.

Первый язык, который создавался для классических планировщиков, - STRIPS (Stanford Research Institute Problem Solver), - разработан в 1971 г. Р.Файксом и Н.Нильсоном. Поиск решения в этой системе был основан на методологии поиска системы GPS (General Problem Solver - универсального решателя задач), созданной А.Ньюэлом, Дж.Шоу и Г.Саймоном, а для проверки истинности предусловий действия в текущем состоянии применялась методология доказательства теорем методом резолюций.

С точки зрения автоматизации ЭСУ интеллектуальное планирование представляет двоякий интерес. Во-первых, в сложных эргатических системах с динамичной и потенциально опасной внешней средой, в условиях жестких требований к правильности и качеству принимаемых оперативных решений велика роль процессов планирования действий оператора. Интеллектуальная поддержка данных процессов может базироваться на машинных методах интеллектуального планирования. Во-вторых, в случае применения агентных технологий для автоматизации эргатических функций возникает проблема планирования действий самих агентов. В качестве примера можно привести задачу локального планирования для делиберативных агентов с InteRRaP-архитектурой [2].

6 Ситуационно-целевое управление

Ситуационный подход, берущий начало от аппарата ситуационного исчисления Дж.Маккарти [19] и модельной теории мышления Д.А.Поспелова и В.Н.Пушкина [8], находит широкое применение в сфере организационного управления, а также в системах искусственного интеллекта.

Ситуация - конкретный набор обстоятельств, которые влияют на систему в данное конкретное время. *Ситуационная модель* дает формальное описание мира с помощью ситуаций, в которых предстоит действовать рассматриваемой системе. В модели учитываются те динамические факторы в структуре объекта и внешнего мира, которые действительно влияют на последующее развитие ситуаций.

Классическое ситуационное управление подчиняется двухступенчатой схеме «ситуация - действие». Алгоритм управления явно не задается, а реализуется на *ситуационной сети*, под которой понимается ориентированный взвешенный граф переходов по эталонным ситуациям, определяемый исходной и целевой ситуациями. Дуги графа отражают смену ситуаций и вызывающие ее информационно-управляющие процессы.

Компьютерная поддержка работы оператора требует разработки формальной модели функционирования ЭСУ. Деятельность ЧО носит сознательный характер - оператор осознает цель своих действий в каждой конкретной ситуации, способы достижения цели и предвидит результат. Следовательно, для описания работы оператора сложной эргатической системы необходимо применять четырехступенчатую схему «ситуация - цель - план - действие». Поскольку каждая ситуация порождает определенную локальную цель действий, то в итоге глобальная цель функционирования ЭСУ достигается посредством отработки перманентно возникающих локальных целей.

Компьютеризация процесса ситуационно-целевого управления сводится к решению цепочки следующих задач:

- мониторинг ситуаций;
- генерация локальной цели;
- формирование оперативного плана действий;
- выполнение запланированных действий.

В рамках агентного подхода ситуационно-целевое управление воплощается двухуровневой иерархической структурой: каждая из выделенных функциональных задач делегируется соответствующему агенту-исполнителю, а действия последних координируются вышестоящим агентом. В данную группу агентов включаются также интерфейсные агенты, реализующие интерактивную оболочку ситуационно-целевого подхода.

Термины «ситуация» и «ситуационное управление» достаточно широко распространены в области автоматического и автоматизированного управления, получая различные истолкования (см., например, [4, 9, 15]). В связи с этим необходимо подчеркнуть особенности развиваемого авторами ситуационно-целевого подхода:

- 1) Ситуация не является совокупностью переменных состояния объекта и окружающей среды, а идентифицирует некоторую область локализации динамических переменных. К примеру, при малых возмущениях текущего динамического режима состояние системы изменяется, однако ситуация остается неизменной.
- 2) Каждая ситуация порождает локальную (рабочую) цель деятельности, для достижения которой формируется свой план действий, который принимается к исполнению. В итоге процесс ситуационно-целевого управления носит дискретно-временной характер.
- 3) Важнейшими компонентами ситуационно-целевого подхода являются классификация ситуаций (в частности, выделение критических - предаварийных и аварийных ситуаций) и базирующаяся на данной классификации задача ситуационного мониторинга.
- 4) Предлагаемый ситуационно-целевой подход органично сочетается с многоцелевым принципом функционирования ЭСУ: главная цель системы разбивается на подцели, порождаемые сменяющимися ситуациями.
- 5) Излагаемая трактовка ситуации позволяет характеризовать и анализировать динамические процессы в объекте с учетом различных факторов неопределенности в динамике объекта и свойствах окружающей среды.
- 6) Ситуационное пространство динамических объектов является конечным и представляется ситуационным графом. Тем самым открывается возможность построения аппарата ситуационного исчисления и разработки логико-динамического формализма ситуационно-целевого управления.

Описанные особенности механизма ситуационно-целевого управления придают ему свойства гибкости и наращиваемости, что весьма ценно для реализации функций обучения и само-

организации в ЭСУ. Действительно, расширения ситуационного пространства в конечном итоге воплощается в конструирование новых цепочек «ситуация - цель - план - действие», причем в процессе функционирования системы возможен поиск наиболее эффективного плана действий.

7 Сценарное планирование

Понятие сценарного планирования широко используется в когнитивной психологии [13] и, в частности, в когнитивной теории сценариев Р.Шенка и Р.Абельсона [20, 21]. Оно также является одним из наиболее эффективных системных инструментов стратегического менеджмента, которые позволяет хозяйствующим субъектам управлять неопределенностью будущего [11].

Сценарий является упорядоченной последовательностью сцен. Сцена определяется стереотипной ситуацией, сложившейся вследствие наступления некоторого набора событий, и действиями в данной ситуации. Сценарий разворачивается в процессе функционирования системы. Сценарные модели, по сути, являются одним из способов представления знаний [7].

Сценарное планирование заключается в разработке альтернативных сценариев будущего и создании на их основе гибких планов действий. Все альтернативные сценарии формируются на основе исследования окружающей среды на наличие предопределенных элементов и ключевых неопределенностей, и комбинировании исходов последних.

Сценарное планирование рассматривает все сценарии как одинаково возможные в будущем. Сценарный подход придает гибкость стратегиям управления, помогает сократить время реагирования на изменение ситуаций.

В рассмотренной выше цепи эргатических функций «ситуация - цель - план - действие» крайне проблематичным оказывается механизм оперативного планирования действий. Такое планирование должно предписывать выполнение последовательности операций над изменением состояния системы «машина - среда» в текущей ситуации и должно учитывать особенность психологии ЧО: восприятие им рабочей среды, структуру информационных и мыслительных процессов. Таким образом, здесь налицо весьма специфический тип планирования, которое логично трактовать как сценарное планирование.

Задачи сценарного планирования рабочих процессов в эргатических системах, в сущности, являются новым классом задач исследования операций, порожденным проблемой интеллектуальной поддержки деятельности операторов. Данный класс задач требует проведения надлежащих теоретических исследований и разработки адекватной алгоритмической базы с учетом современных достижений в области программных и интеллектуальных технологий.

8 Интеллектуализация человеко-машинного интерфейса

Важнейшим компонентом архитектуры перспективных ЭСУ является человеко-машинный интерфейс (ЧМИ), в максимальной степени учитывающий человеческий фактор - психологические, физиологические и анатомические аспекты человеко-машинного взаимодействия с учетом всего комплекса решаемых функциональных задач.

Оператор лишен возможности непосредственно наблюдать за управляемыми объектами и вынужден пользоваться информацией, которая поступает к нему по каналам связи. В этом случае его деятельность совершается не с реальными объектами, а с информационными моделями реальных объектов.

Информационная модель - совокупность информации о состоянии и функционировании объекта управления и внешней среды. Она отражает все существенно важные для управления свойства реальных объектов, т.е. тех источников информации, на основе которого оператор формирует образ реальной обстановки, производит анализ и оценку сложившейся ситуации, планирует управляющие воздействия, принимает решения, обеспечивающие правильную ра-

боту системы и выполнение возложенных на нее задач, а также наблюдает и оценивает результаты их реализации.

Объем информации, включенной в модель, и правила ее организации должны соответствовать задачам и способам управления. Физически информационная модель реализуется с помощью устройств отображения информации, на основе которой строится вся деятельность оператора.

Процесс автоматизированного управления должен осуществляться на основе информационной модели. Однако существуют границы возможностей оператора, за которыми восприятие им информации и его действия затруднены или становятся невозможными [16], вследствие чего необходимо рассматривать проблему ЧМИ через фильтр ограничений.

ЧМИ должен обеспечивать оптимальный информационный баланс и не приводить к таким нежелательным явлениям, как дефицит или избыток информации. Он должен осуществлять адаптивную фильтрацию информационных потоков: в зону внимания ЧО должна вводиться только лишь релевантная оперативная информация, отвечающая текущей цели управления.

Современный ЧМИ включает компьютерный пользовательский интерфейс. С увеличением сложности операторской работы возрастают требования к функциональным и эргономическим аспектам пользовательского интерфейса. Многообразие всех этих требований возможно учесть лишь на пути создания интеллектуального многомодального интерфейса.

Многомодальный интерфейс естествен для межчеловеческого общения. Человек сам выбирает, какой канал для передачи какого типа информации ему наиболее удобно использовать в данный момент. Многомодальные интерфейсы могут параллельно обрабатывать несколько потоков информации. Различные формы вывода визуальной и звуковой информации обеспечивают устойчивый интерактивный диалог с пользователем [12]. Современный диалоговый интерфейс организуется с применением речевого интерфейса, компьютерных технологий 3D визуализации и виртуальной реальности.

Интеллектуализация интерфейса направлена на обеспечение гибкого и адаптивного взаимодействия оператора с компьютером, управление активностью и формами ведения диалога. В его функцию входит также обеспечение диалога на ограниченном естественном языке. Считается, что естественный интерфейс между человеком и компьютером будет достигнут тогда, когда они смогут говорить друг с другом.

Основным информационным каналом для оператора является визуальный. В связи с этим самого пристального внимания для понимания процессов человеко-машинного взаимодействия заслуживает концепция экологической оптики, разработанная Дж.Гибсоном [3].

Голосовой интерфейс является наиболее естественным и простым способом взаимодействия человека с компьютером [18]. Он дополняет (а в ряде случаев исключает) зрительный канал передачи информации, разгружая нервную систему ЧО. Базовые функции голосовых сервисов: командное управление и информирование о состоянии системы и внешней среды, подтверждение выполнения команд, выдача инструкций-подсказок. Голосовой интерфейс играет важную роль в обеспечении безопасности эргатических систем: голосовые сообщения дают возможность адекватно воспринимать информацию и принимать правильные решения в критических ситуациях.

Агентные технологии открывают новые перспективы решения проблемы конструирования интеллектуального интерфейса. Интерфейсные агенты создают единую интегрированную коммуникационную среду, обеспечивают гибкость интерфейсу - адаптируемость структуры и сценария развития диалога применительно к выполняемому заданию и текущим процессам ввода-вывода информации. Формирование грамматики диалога должно основываться на онтологической модели рабочей среды ЭСУ. Интерфейсные агенты способны одновременно (параллельно) выполнить ряд задач, включая корректировку или блокировку ошибочных действий оператора, ведущих к фатальным ошибкам.

Список литературы

- [1] Венда В.Ф. Системы гибридного интеллекта: эволюция, психология, информатика. М.: Машиностроение, 1990. - 448 с.
- [2] Городецкий В.И., Грушинский М.С., Хабалов А.В. Многоагентные системы (обзор) // Новости искусственного интеллекта. 1998. № 2. - С. 64-116.
- [3] Гибсон Дж. Экологический подход к зрительному восприятию. М.: Прогресс, 1988. - 464 с.
- [4] Искусственный интеллект и интеллектуальные системы управления / И.М.Макаров, В.М.Лохин, С.В. Манько, М.П.Романов. М.: Наука, 2006. - 333 с.
- [5] Леонтьев А.Н. Деятельность, сознание, личность. М.: Политиздат, 1975. - 304 с.;
- [6] Ломов Б.Ф. Методологические и теоретические проблемы психологии. М.: Наука, 1984. - 444 с.
- [7] Люгер Дж.Ф. Искусственный интеллект: стратегии и методы решения сложных проблем. М.: ИД «Вильямс», 2005. - 864 с.
- [8] Поспелов Д.А., Пушкин В.Н. Мышление и автоматы. М.: Сов. радио, 1972. - 222 с.
- [9] Поспелов Д.А. Ситуационное управление: теория и практика. М.: Наука, 1986. - 288 с.
- [10] Рассел С., Норвиг П. Искусственный интеллект: современный подход. М.: ИД «Вильямс», 2007. - 1408 с.
- [11] Рингланд Дж. Сценарное планирование для разработки стратегии. М.: ИД «Вильямс», 2008. - 560с.
- [12] Ронжин А.Л., Карпов А.А., Ли И.В. Речевой и модальный интерфейсы. М.: Наука, 2006. - 171 с.
- [13] Солсо Р.Л. Когнитивная психология. М.: Тривола, 1996. - 600 с.
- [14] Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. М.: Эдиториал УРСС, 2002. - 352 с.
- [15] Чекинов Г.П., Чекинов С.Г. Ситуационное управление: состояние и перспективы // Приложение к журналу «Информационные технологии». 2004. № 2. - 32 с.
- [16] Человеческий фактор: В 6 т. / Под ред. Г.Салвенди. М.: Мир, 1991.
- [17] Швецов А.Н. Агентно-ориентированные системы: от формальных моделей к промышленным приложениям / Всерос. конкурсный отбор обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы», 2008. - 101 с.
- [18] Cohen M.H., Giangola J.P, Balogh J. Voice User Interface Design, Publisher: Addison-Wesley Professional, 2004. - 368 p.
- [19] McCarthy J. Situations, Actions and Causal Laws // Stanford University Artificial Intelligence Project: Memo 2, Stanford. CA. 1963.
- [20] Schank R.C., Abelson R.P. Scripts, Plans, Coals and Understanding: An Inquiry into Human Knowledge Structures. Hillsdale, N.J.: Lawrence Erlbaum Ass., 1977. - 248 p.
- [21] Schank R.C. Dynamic Memory. Cambridge: Cambridge Univ. Press., 1982. - 234 p.

ENVIRONMENT CARE IN THE RUSSIAN FEDERATION: AN ANALYSIS FROM THE DUTCH PERSPECTIVE

A.C. Bijl

Project Officer – NL Agency
NL EVD International
The Hague, The Netherlands
arie.bijl@agentschapnl.nl
tel: + 31 88 6028 508

Key words: environment care, pollution, bi-lateral assistance

Abstract

Двусторонняя программа сотрудничества об охране окружающей среды осуществляется между Нидерландами (финансирующая сторона) и Российской Федерацией (страна-получатель). Инвестиционные проекты этой программы направлены на решения различных проблем охраны окружающей среды. Отличия между странами-участниками по организационным и экономическим условиям для охраны окружающей среды увеличивают риски для инвестиций, начатых в соответствии с данной программой, но ожидаемые структурные преобразования в России оправдывают эти риски.

Introduction

Environment is defined as the total of surroundings in which men, animals and plants live. By interaction with the surroundings, every creature has some influence on the composition of those surroundings. As an individual this is hardly noticeable but the development of society creates types of communal behaviour and the accumulated effects of that behaviour has significant influence on the composition of the surroundings. When this reaches a level where self regulating biological and chemical processes cannot cope with the seize, we speak of pollution. So by diminishing the effects of behaviour, men can control the composition of the environment it lives in.

There are three main spheres on which man can influence the negative effects of its behaviour on environment:

- the neutralisation of pollution in the past and the prevention of further pollution;
- the prevention of excessive use of natural resources;
- the prevention of negative effects on nature and landscape.

Care for the environment is an activity of every society. As survival and reproduction are traditionally items that the members of a society aim for, it is as well a characteristic that members of a society care for the environment as they want to enjoy living and want their offspring to enjoy a habitat that is healthy and clean.

When people live together in high density, the effects of pollution become more rapidly noticeable. That is why densely populated area's of the world demonstrate high activity on environmental care. They find solutions for problems as the legislation forces the members of society to diminish the effects of their behaviour.

Pollution is often not a local or regional problem. Many types of pollution have an international (water pollution) or global (warming) character. So pollution is also an important theme in international and global co-operation. It is evident that environment care must have a high priority. During the past decades man has significantly faster started to influence the environment and it seems this has irreversible effects. We can only try to slow down the process of changes.

The UN has taken up this challenge and aims for fruitful international co-operation. One of the aspects of this is treaties that countries that have a higher status in environmental care give assistance to countries where care for the environment is in an earlier stage of development. The Netherlands has taken up this challenge in programs where we provide assistance to countries in Eastern Europe to develop the care for the environment. We do this on 2 levels: Government to Government assistance on development of legislation and other forms government activity that influences the behaviour of society in this matter; the second program is an grant program for investments in processes that contribute to environmental care: the PSO Environment program. The Netherlands has chosen the Russian Federation as one of the countries for this co-operation. Other countries that are chosen for co-operation on this subject are: Ukraine, Turkey, Serbia and Georgia.

In consultations with national and/ or regional governments environment themes are chosen that need assistance in the recipient countries and projects are identified and developed. These themes are based on the principles of environmental care:

- measures to reduce emissions of polluting substances;
- measures to reduce the volume of polluting material and substances;
- structural measures to change the character of production- and consumption processes.

1 The PSO Environment program

The Netherlands Ministry of Housing, Spatial Planning and Environment has established a number of themes on which investments can be supported and knowledge can be transferred under the PSO Environment program:

- Waste treatment, including medical or industrial waste incineration;
- Agricultural or industrial wastewater treatment, including marine pollution treatment;
- Reduction of water, air or noise pollution by motor vehicles and/or equipment outdoors;
- Chemical or biological rehabilitation of contaminated soil;
- Soil and groundwater protection, landfills for hazardous and/or inert waste;
- Energy saving at the end-user, preferably a group of consumers, e.g. energy efficient buildings;
- Development of markets for sustainable goods and services;
- Environmental friendly buildings; and/or
- Renewable energy.

On any of the above themes Dutch institutions and companies have advanced knowledge and solutions to counter related problems. When it comes to technology, often manufacturers of other countries are involved as well in those solutions. Anyway, Dutch companies can offer advanced solutions to Eastern European governments.

1.1 The Russian situation

On the above themes different government bodies are involved in Russia. So suggestions for projects come from different sources, both on federal and on regional level. The paramount priority of the Russian federal government is at the moment energy efficiency. The ministries involved in this have a clear priority. However the ministry that is responsible for environment issues: the Ministry of Natural Resources and Environment is emphasizing the importance of waste treatment and wastewater treatment.

Different government levels and recent development of new legislation create a situation where institutions and companies involved in the day to day care for the environment have to deal with complicated legislations and permit issues. Practical solution for that situation is a system of tolerating which of course lead to very non transparent situations in practice.

When dealing with issues of energy efficiency and energy saving one has to deal with a complicated pricing system for energy. Measures for energy efficiency very often have little or no prospects of financial savings. Infrastructure for the distribution of heat are often inflexible and

systems where more responsibility for energy efficiency is assigned to the individual citizen, are difficult to implement.

The Russian Federation is a large country with a complicated history. There are many places where lack of care for the environment in the past have led to contaminated soil and water. On the other hand, the size of the country hide the effect of these contaminations in many situations.

1.2 Assessing the options

For Dutch government institutions that have to decide on the application of limited funding there is a difficult choice to be made. Criteria for choosing a project are many:

- the project size must match the funds that are available;
- the project must be an investment;
- the effects of the problem that is going to be subject of the project must preferably have an international character;
- the project preferably should have a demonstration character so it can be replicated in Russia in similar situations;
- the practical implementation of a project should not be frustrated by complicated legislation issues;
- personal views of the persons involved in the assessment on the Russian situation.

The selection of projects is made in 2 stages. First a rough assessment is made from different project idea's. This is made on the mentioned criteria, but also on the expected feasibility of the projects. After that an identification visit takes place. From this a decision document is drawn up.

1.3 Procedures of the program

After a project is identified, a more precise project design must be developed. For that purpose Dutch experts are assigned to the task of developing a Terms of Reference for each project. For identifying the consultants that can qualify for such a task NL EVD International has a continuous pre qualification program for interested consultants. As soon as a project has been identified, the formulation assignment is tendered for a limited number of consultants. They have to submit their quotation for the specific assignment. The selected consultant will design the project in the Terms of Reference (rough design) for the actual project, make a rough estimate of the costs and will get an approval for the design from the future recipient of the project and relevant authorities.

The project will be tendered with Dutch companies after a prequalification process. In the prequalification process the main tenderer will present his consortium and in that way he will demonstrate the fitness of the consortium for implementing the project. A maximum of 5 prequalifying companies will be invited to submit a quotation for implementing the project. After evaluating the quotations, the tenderer with the best offer on quality and price will be contracted.

For implementation of the projects a result based contract will be signed with the successful tenderer. The contract will be divided in different sub-results. For each sub result a certain amount will be awarded; last (significant) part will only be paid when the project is operational.

2 Implementation of the program

When choosing a project for further development from different themes, the options have to be evaluated. The different criteria are taken into consideration but it is obvious that evaluators from different background judge differently. As the program finances are to be implemented in Eastern Europe and the evaluator live in The Netherlands, it will be clear that the outcome of the evaluation process needs explanation.

2.1 Considerations

Projects for funding by PSO Environment first of all have to comply with the policy of the Netherlands Ministry of Environment. In general they support projects which can create an

international effect. That would mean that projects that mitigate a trans border pollution get a high score. That gives projects for waste water treatment and lowering greenhouse gasses a high rating. As second theme waste treatment, in a way that the volume is reduced, also have a high score.

For the projects that are identified, in general, limited funds are available from PSO Environment. That often results in a need for additional funding. The additional funding may be obtained from local budget or commercial funding. As providers of the PSO Environment we do not favour co-financing from other bilateral or international fund aimed at improvement of the environment or energy efficiency. Actually, as especially international programmes from EBRD or Worldbank institutions are aimed at much bigger projects, PSO Environment aims to support pilot investments that can prove feasibility of a system that can be replicated or increased to a much bigger scale with finance from these big international institutions.

If during the identification of a project theme and object it becomes clear that the future project will face complex issues with legislation and permits, it is a reason to reject the idea. Also complicated commercial structures surrounding a proposed project may lead to rejection of the idea.

2.2 Practical situations

For the Russian Federation the PSO Environment program started in 2009. Through contacts with federal ministries and regional governments 3 possible projects were identified:

- energy efficiency measures in public buildings in Moscow region;
- improving the water quality in the Baltic sea through mitigating the effects of land activities in Leningradskaya oblast;
- diminishing the emission of greenhouse gasses from a waste deposit and using these gasses as a source of alternative energy in Samara region.

Energy efficiency is clearly a theme with high attention in the Russian society. It has been declared a priority by high politicians and new legislation has been developed on the subject. In contradiction with the intensive political attention, it is observed that the economical conditions for energy efficiency do not support enough reason for replicating measures for energy efficiency. The low prices for oil, gas and electricity do often not justify investments in specific measures. Calculated pay back times for investments are often long and do not outweigh the risks of the investments. Still support from the PSO Environment program is considered justified as:

- energy efficiency contributed to the objective of reducing emissions that have a global effect;
- it is expected that in the near future the prices of energy in the Russian Federation will rise and that the logical demonstration effect of the project will become evident after a few years. Replicating the project will be a logical effect.
- Russian companies are forced by law to implement energy efficiency measures so the projects can serve as examples.

Mitigation of the volume of pollution in the waters discharged directly or indirectly to the Baltic sea will improve the quality of water in the Baltic sea. Especially in the Nord West region of the Russian Federation the awareness for working on this phenomenon is evident. A large investment program has been designed for waste water treatment. Companies are obliged to make investments for diminishing the pollution of the water they use in the manufacturing processes as a result of federal legislation. Small settlements do often not have the resources to make the necessary investments. Support is required.

As the investments are often made from general budget one may question the sustainability of the investments. When applying the principle “the one who pollutes will pay” one may expect a more sustainable system of maintaining and investing in waste water treatment. Payment structures that are directly channelled from citizens to organisations managing the waste water treatment are advised to create sustainability of the entire system. Pilot regions are already trying this type of financing.

Reduction of the emission of greenhouse gasses is clearly a measure with international effect. From waste disposal the greenhouse gas methane is emitted. By catching this gas it can be controlled.

The subsequent processing of the collected gas can be done by burning the gas. The result of this process is the production of carbon dioxide. By burning the pollution as emitted gas is reduced 20 fold. However, when having methane gas available it may as well be used as an energy source. The collected gas can be burned in a gas engine that is connected to an electricity generator. The generated electricity can substitute electricity generated from fossil fuel and in that way contribute to the goal of reducing the use of natural resources.

Not every idea is worth considering to become a project. A proposal to start a project for biogas production from poultry manure was rejected as it was considered not appropriate under the present circumstances. The reasons were:

- poultry manure cannot be a single source for biogas production as it is too rich in protein;
- combining poultry manure with more energy rich products could not be considered as a solution because availability of energy rich products was not demonstrated and public health legislation would probably not allow the combination;
- both dry poultry manure and the rest product of the biogas production are used as fertiliser for arable crops; as the field are remote transport of processed manure would be more costly as water has to be added before the biogas production process;
- the alternative: using unprocessed poultry manure as a fertilizer would be more appropriate but adjustments to (illogical) legislation would have to be made.

Conclusions

Investments on environmental care in the Russian Federation require a lot of attention. Bi-lateral assistance on the themes that are components of environment care helps to introduce new views and investments do bear risks as these investments should have a demonstrative effect. The economic circumstances of the project in different meanings of that term should get attention. In that way the development of a sustainable society in the Russian Federation will progress more quickly.

Literature

- [1] PSO Environment program information, NL EVD International.
- [2] Market survey Russia waste and waste water, NL EVD International, November 2009

ДЕМОКРАТИЗАЦИЯ ПРОЦЕССОВ УПРАВЛЕНИЯ ВУЗОМ

В.А. Виттих¹, А.Н. Гинзбург², О.Р. Козырев², С.В. Смирнов¹

¹ Учреждение Российской академии наук Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
vittikh@iccs.ru

тел: +7 (846) - 332-39-27, факс: +7 (846) - 333-27-70

² Нижегородский филиал Государственного университета – Высшей школы экономики
603155, Нижний Новгород, ул. Б. Печёрская, 25/12, Россия
aginzburg@bk.ru
тел: +7 (831) 419-55-81

Ключевые слова: высшее учебное заведение, сложные искусственные системы, холон, актор, образовательная услуга, интегрированная нормативно-правовая база знаний, децентрализация, онтологическая модель ситуации

Abstract

The article defines problems need to be solved to democratize the college management processes, while college is considered to be a complex “holonic” system, aimed at providing consolidated and basic educational services. All the services should be brought into effect in course of heterogeneous actors (lecturers, students and college officials) interaction, whose decision-making process is guided by integrated regulatory and legal knowledge frameworks. Theoretical foundations for decentralized college economics ensuring the passage to the new college economic relations need to be developed. There are considered functions and components of the decision-making support system, including empirical information structuring and visualization facilities, real-time resource management and etc. along with educational services portal as well as integrated regulatory and legal knowledge frameworks.

Введение

Несмотря на перманентное реформирование российской высшей школы, принципы управления вузами, как правило, остаются прежними – бюрократическими, предполагающими жесткую иерархию власти, разделение труда на основе функциональной специализации, регламентацию процедур, определяющих порядок действий во всех ситуациях, встречающихся в процессе функционирования вуза и т.п. Студент предстает здесь как «унифицированный», лишенный каких-либо особенностей, субъект, который под воздействием вузовской «обучающей машины» должен приобрести определенный набор знаний и навыков. Все решения, касающиеся того, чему учить и как учить, принимаются вузовскими руководителями различных рангов и, в какой-то степени, преподавателями. «Степень демократизации» здесь невысока, поскольку кафедры, как правило, рассматриваются как исполнители воли вышестоящих инстанций, а не как самостоятельные «бизнес-единицы».

Тем не менее, из вузовской практики можно привести примеры, когда *студент* рассматривается не как «пассивный объект управления», а как *партнер преподавателя* со своими собственными взглядами на мир и интересами, ведущий с ним диалоговое взаимодействие, что позволяет *формировать студента как личность*, а не превращать его в пополняемое «хранилище» знаний. Не секрет, что *персонализация обучения и демократизация процессов управления* вузом, предполагающая передачу кафедрам части полномочий принимать решения, создают основу для повышения качества учебного процесса. Вопрос состоит в другом: каким образом перейти от фрагментарного обучения отдельных, наиболее талантливых студентов, пре-

доставляя исключительно для них ресурсы, необходимые для персонального обучения, к новой, модернизированной системе обучения всех студентов вуза, учитывающей специфику, *ценностные ориентиры и интересы каждого студента* и обеспечивающей *эффективное распределение имеющихся в вузе ресурсов* для реализации индивидуальных программ обучения всех без исключения студентов, повысив при этом самостоятельность кафедр в процессах принятия решений (в том числе, экономических).

Для решения этой задачи в данной статье предлагается использовать парадигму ограниченной рациональности принятия решений [1].

1 Вуз как сложная искусственная система

Будем рассматривать вуз как сложную искусственную систему, создаваемую для предоставления образовательных услуг и объединяющую в себе:

- людей (*живые сущности*), в том числе *акторов* – лиц, участвующих в принятии и реализации решений;
- *неживые сущности* – различные материальные объекты (учебные аудитории, научные лаборатории, служебные помещения, приборы, оборудование и т.п.);
- *абстрактные сущности*, представляющие собой, в широком смысле, знания в форме научной и учебно-методической литературы, а также *нормативно-правовых актов*, регулирующих отношения между различными видами сущностей.

Актеры (преподаватели, студенты и руководители) являются *неоднородными*, т.е. каждый из них имеет свою точку зрения на мир и сложившуюся ситуацию, собственные интересы и шкалу ценностей. Актеры объединяются в *холоны*, организующиеся в качестве динамически упорядоченных целостностей, т.е. отношения между элементами сложной системы строятся на основе принципа *«часть – целое»*.

Иными словами, отношения между актерами базируются не на жесткой «вертикальной» схеме «начальник – подчиненный» (или «учитель – ученик»), а предполагают организацию «горизонтального» *взаимодействия* между ними. Появляющаяся основа для установления «равенства» отношений, тем не менее, не означает вседозволенности, поскольку речь идет о «равенстве перед законом», когда либерально-демократические *отношения* между преподавателями, студентами и вузовским руководством *регулируются соответствующими нормативно-правовыми документами*.

Демократизация управления вузом достигается, таким образом, за счет того, что и простые преподаватели, и студенты, и руководящий состав вуза в процессах принятия решений должны постоянно опираться на единую вузовскую нормативно-правовую базу. В настоящее время все они, конечно же, руководствуются учебными планами и программами, расписаниями занятий, зачетов и экзаменов. И этого достаточно, если исходить из традиционной «стабильности» учебного процесса, к которой стремится большинство вузов.

Если же ориентироваться на *динамику обучения*, которая неизбежна в условиях, когда получатель образовательной услуги – студент – активно вмешивается в учебный процесс, непредвиденным образом *изменяя и дополняя свои потребности в приобретении тех или иных знаний*, то ограниченные ресурсы вуза (кадровые, финансовые и иные материальные) должны гибко перераспределяться, а при их недостатке – пополняться. Здесь уже нельзя говорить не только о неизменном учебном плане и расписании занятий, но и о стабильной нагрузке преподавателей (а, соответственно, и об их зарплате), и о нагрузке учебного оборудования и т.п. Постоянными, таким образом, становятся только перемены.

Ключевую роль в обеспечении гибкой организации учебного процесса в вузе должны играть холоны [1]. *Холон*, как известно, представляет собой *самодостаточную целостность*, являющуюся частью другой, большей целостности. Холон – это, например, *кафедра с расширенной автономией* и располагающая всеми видами ресурсов (в том числе, финансовых) для

осуществления своей деятельности. Такая *кафедра-холон*, не дожидаясь команды «сверху», *самостоятельно вступает в договорные отношения* с другими кафедрами и, в отдельных случаях, со студентами, если того требуют интересы дела. Холон-деканат при этом может выполнять координирующую функцию по отношению к кафедрам, а «штабной холон» – функции ректората.

Таким образом, «входами» в сложную искусственную систему, именуемую вузом, являются заявки на оказание образовательных услуг. Система, основными структурными элементами которой являются холоны и их людские компоненты – *акторы* [1], должна гибко приспособлять свои ресурсы к постоянно изменяющейся нагрузке, обеспечивая актерам возможность оперативно получать интересующую их (релевантную) нормативно-правовую информацию для регулирования отношений целостностей [2]. На смену жесткому централизованному управлению здесь приходят *демократические принципы социальной самоорганизации* [3,4].

2 Консолидированные и базовые образовательные услуги

Будем рассматривать вузы экономического, технического, социального, социотехнического и иного подобного профиля, главной особенностью которых является то, что они готовят инженеров по разработке и эксплуатации сложных искусственных систем, т.е. преподают «науки об искусственном» [5]. В отличие от естественных наук, имеющих дело с верифицируемыми знаниями, приобретаемыми на основе оппозиции субъекта и объекта, науки об искусственном направлены на приобретение *аксиологических* (ценностно ориентированных) *знаний* в условиях суперпозиции субъекта и объекта, что предполагает *повышение роли субъект – субъектных отношений* [6]. Иными словами, изучить и понять поведение сложной искусственной системы возможно только путем привлечения специалистов различных предметных областей, каждый из которых со своей, в общем случае – субъективной, точки зрения может изложить видение ситуации, сложившейся в системе; а общность опыта взаимодействующих субъектов и общезначимость его результатов достигается за счет *интерсубъективности, указывающей на внутреннюю социальность индивидуального сознания* [7,8].

На этом основании введем понятие «консолидированной образовательной услуги» как совокупности взаимосвязанных базовых образовательных услуг, относящихся к относительно узким предметным областям и нацеленных на общую для них задачу обучения (по аналогии с консолидированной услугой как системой государственных и муниципальных услуг населению [2]).

Например, консолидированная образовательная услуга может оказываться по теме «Управление многоквартирными домами» и включать в себя следующие базовые образовательные услуги:

- экономические основы управления многоквартирными домами, включая ценообразование и тарифное регулирование;
- источники финансирования и способы привлечения финансовых средств на проведение капитального ремонта;
- расходные обязательства органов местного самоуправления при управлении многоквартирными домами;
- нормативно-правовые акты, регулирующие предоставление коммунальных ресурсов и услуг;
- объекты жилищного фонда, подлежащие управлению;
- основы электро-, тепло-, газо- и водоснабжения многоквартирных домов.

Базовые услуги могут предоставляться отдельными *структурными подразделениями вуза (кафедрами-холонами)* экономического, юридического, технического и какого-либо иного профиля, каждое из которых *располагает соответствующими кадровыми и материальными*

ресурсами, а также имеет счета (или субсчета) для проведения финансовых операций. Оказание консолидированной услуги может координироваться холлом-деканатом, взаимодействующим со студентами по организационным вопросам, с использованием финансовых механизмов и вузовской нормативно-правовой базы. *Штабной холлом*, определяющий стратегические направления развития вуза и распоряжающийся необходимой для этого частью вузовского бюджета, прогнозирует консолидированные услуги, которые могут оказаться востребованными в будущем, а какие, наоборот, – потеряют актуальность.

Все консолидированные услуги сводятся в реестр консолидированных образовательных услуг вуза, на основании которого должен быть разработан портал образовательных услуг вуза, с использованием которого любой пользователь (студент, абитуриент, сотрудник вуза и др.) может оперативно получить информацию о содержании любой консолидированной услуги и составляющих ее базовых услугах, о том, где и кем они оказываются, о времени предоставления услуг и их стоимости.

3 Интегрированные нормативно-правовые базы знаний

Демократизация управления вузом, связанная с расширением прав структурных подразделений и студентов в процессах принятия решений, предполагает повышение их правовой грамотности. С этой целью должна быть значительно повышена доступность нормативно-правовых документов различным категориям пользователей – от студента до руководящего работника вуза.

Дело в том, что в основе предоставления консолидированных и базовых образовательных услуг лежат нормативно-правовые акты федерального, регионального, муниципального и вузовского уровней, которые обычно хранятся в соответствующих полнотекстовых базах данных. Однако быстро найти нужную (релевантную) информацию в таких базах, относящуюся к конкретной ситуации, в которой оказался студент или вузовский работник, просматривая десятки и сотни страниц текста, написанного юридическим языком, трудно даже умелому пользователю компьютера. А если для поисков ответа потребуются значительные усилия и временные затраты, человек вообще может отказаться от получения таких «электронных услуг». Иными словами, традиционный подход, связанный с использованием баз данных, не может обеспечить высокой степени доступности нормативно-правовой информации.

Качественно новые возможности повышения доступности информации открывают технологии компьютерного представления и обработки знаний, которые позволяют превратить тексты законов (в рамках выбранной предметной области) в семантическую сеть понятий и отношений между ними и построенные на их основе наборы правил, которые можно использовать для простой и удобной гипертекстовой навигации по законам, автоматического нахождения нужных нормативно-правовых документов и для ответов на вопросы пользователей. Таким способом обработанные и преобразованные тексты законов и нормативных актов формируют интегрированные нормативно-правовые базы знаний, объединяющие в рамках единой компьютерной семантической сети федеральные, региональные, муниципальные и вузовские нормативные акты, относящиеся к общей для них предметной области (например, к той или иной консолидированной образовательной услуге). Тогда появляется возможность избавить пользователя, имеющего доступ к такой базе знаний через Интернет, от необходимости выполнения сложной цепи действий, связанных с нахождением интересующей его информации.

Такой подход был применен при создании региональной системы предоставления государственных и муниципальных услуг населению в электронной форме с применением интегрированных баз знаний и мультиагентных технологий в социальной сфере Самарской области [9]. Представляется целесообразным использовать его при предоставлении консолидированных образовательных услуг.

4 Децентрализованная экономика вуза

Понятие *децентрализации*, с одной стороны, предполагает независимость экономических решений акторов, а с другой, – взаимозависимость результатов этих решений. Причем неотъемлемой характеристикой процесса децентрализации принятия решений является *неоднородность* акторов и неизбежная *неопределенность*, сопутствующая экономическим решениям [10].

Классическая централизованная модель экономики вуза «стягивает» все решения к верхним эшелонам власти и опирается на веру в стационарность мира, единство человеческих желаний (однородность акторов), теорию общего равновесия и возможность реализации централизованных механизмов планирования и экономических расчетов.

Своеобразным «водоразделом» между классической экономической теорией и новым децентрализованным подходом является отношение к проблеме неопределенности. «Суть проблемы заключается в том, является ли экономика совокупностью однотипных агентов, однородных потоков и связей, или, напротив, она характеризуется качественной и разнообразной неоднородностью. В первом случае равновесие понимается как фиксированная точка, во втором – как состояние локально устойчивое, временное и случайное» [10].

Поскольку современный мир становится все более открытым и динамичным, на смену классической модели принятия экономических решений приходит парадигма ограниченной рациональности принятия решений, согласно которой объективность однородных акторов заменяется солидарностью неоднородных акторов, проблема истинности знания – его ценностью, а само *принятие решения осуществляется в реальном масштабе времени*, т.е. в темпе развития сложившейся проблемной ситуации.

Всё это приводит к необходимости создания теоретических основ децентрализованной экономики вуза, которые должны дать ответы на следующие актуальные практические вопросы. Каким образом должны формироваться и трансформироваться реестры и тарифы консолидированных и базовых образовательных услуг? Какие экономические решения целесообразно делегировать кафедрам, получившим самостоятельность и необходимое ресурсное обеспечение? Каким образом должны строиться экономические отношения между кафедрами, деканатами и ректоратом? Как будет функционировать экономика вуза в условиях, когда студентам предоставляется возможность изменять «траекторию» своего обучения? На каких принципах должно осуществляться разделение образовательных услуг на платные и бесплатные (финансируемые государством)? Какова будет специфика маркетинговых исследований в условиях индивидуализации процесса обучения?

5 Система поддержки принятия решений вуза

Как уже отмечалось выше, демократизация управления вузом предполагает участие всех акторов (преподавателей, студентов и руководителей) в процессах принятия решений, что практически невозможно осуществить, если не создать соответствующий комплекс компьютерных инструментальных средств – систему поддержки принятия решений (СППР) вуза.

В состав СППР входит *портал образовательных услуг вуза*, включающий в себя сведения о содержании любой консолидированной услуги и составляющих ее базовых услугах, о том, где и кем они оказываются, о времени предоставления услуг и их стоимости.

Интегрированные нормативно-правовые базы знаний и средства доступа к ним позволяют любому пользователю оперативно получать релевантную информацию, извлеченную из нормативных документов федерального, регионального, муниципального и вузовского уровней, действующих в сфере высшего образования.

СППР должна включать средства структуризации эмпирической информации и ее визуализации, обеспечивающие в удобной форме представление сведений о структуре вуза, функ-

циях его подразделений и должностных лиц [11], а также программные средства поддержки процессов управления распределением ресурсов в реальном масштабе времени [12].

СППР, которые предлагается создавать с применением онтологических моделей ситуаций (ОМС) – описаний ситуаций в форме персональных, групповых онтологий и объектных моделей, должны обеспечивать [1]:

- компьютерное представление персональных, групповых онтологий и объектных моделей;
- возможность решения задач в открытой форме в реальном масштабе времени, т.е. в темпе, соответствующем скорости развития ситуации;
- сетевой обмен информацией между акторами по принципу «каждый с каждым»;
- «дружественный интерфейс» для интерактивного взаимодействия акторов в процессах построения групповых онтологий и трансформации объектных моделей;
- регистрацию в системной библиотеке ОМС, которые были использованы при принятии решений, с тем, чтобы была возможность их применения в качестве прецедентов в процессах управления в будущем.

Заключение

Демократизацию процессов управления вузом необходимо развивать в следующих направлениях:

- повышение степени самостоятельности структурных подразделений вуза и роли «горизонтальных» связей между ними;
- индивидуализация обучения, обеспечивающая возможность студентам выбирать и корректировать «траекторию» обучения и участвовать в процессах принятия решений, затрагивающих их интересы.

Для достижения этих целей требуется решение ряда научных и практических задач.

Вуз необходимо рассматривать как сложную искусственную «холоническую» систему, предназначенную для предоставления консолидированных и базовых образовательных услуг. Должны быть разработаны реестры этих услуг и портал образовательных услуг вуза.

Все услуги должны реализовываться в процессах взаимодействия неоднородных акторов (преподавателей, студентов и руководителей вуза), которые при принятии решений опираются на предметно-ориентированные нормативно-правовые базы знаний вуза. Высокая степень доступности правовой информации должна достигаться путем создания и использования в процессах принятия решений интегрированных нормативно-правовых баз знаний, объединяющих в рамках единой компьютерной семантической сети федеральные, региональные, муниципальные и вузовские нормативные акты, относящиеся к общей для них предметной области.

Необходима разработка теоретических основ децентрализованной экономики вуза, которые должны обеспечить переход к новым экономическим отношениям в вузе и изменение подходов к маркетинговым исследованиям в области образовательных услуг. Неотъемлемой характеристикой процесса децентрализации принятия решений является неоднородность акторов и неизбежная неопределенность, сопутствующая экономическим решениям.

Демократизацию процессов управления вузом практически невозможно осуществить, если не создать компьютерную систему поддержки принятия решений, включающую в себя уже упомянутые портал образовательных услуг и интегрированные нормативно-правовые базы знаний, а также средства структуризации эмпирической информации и ее визуализации, управления распределением ресурсов в реальном масштабе времени и т.п. Разработку системы поддержки принятия решений вуза целесообразно осуществлять с применением онтологических моделей ситуаций, представляющих собой описание ситуаций в форме персональных, групповых онтологий и объектных моделей.

Список литературы

- [1] Виттих В.А. Парадигма ограниченной рациональности принятия решений-1. // Вестник Самарского государственного технического университета, серия «Технические науки», №3 (25), 2009, с. 22–31.
- [2] Виттих В.А. Управление как регулирование отношений целостностей. // Проблемы управления и моделирования в сложных системах: Труды X Международной конференции. – Самара, Самарский научный центр РАН, 2008. – С. 34-40.
- [3] Küppers G. Self – Organization – The Emergence of Order. From local interactions to global structures. // <http://www.unibielefeld.de/iwt/sein/paperno2.pdf>, July 1999.
- [4] Виттих В.А. Целостность сложных систем. // Проблемы управления и моделирования в сложных системах: Труды IV Международной конференции. – Самара, Самарский научный центр РАН, 2002. – С. 48-58.
- [5] Саймон Г. Науки об искусственном. – М.: Мир, 1972.
- [6] Виттих В.А. Аксиологические и верифицируемые научные знания. // Проблемы управления и моделирования в сложных системах: Труды XI Международной конференции. – Самара, Самарский научный центр РАН, 2009. – С. 449-454.
- [7] Словарь философских терминов. – М.: ИНФРА-М, 2004.
- [8] Современный философский словарь. – Лондон, Франкфурт-на-Майне, Париж, Люксембург, Москва, Минск: «ПАНПРИНТ», 1998.
- [9] Виттих В.А., Волхонцев Д.В., Гриценко Е.А., Светкина Г.Д., Скобелев П.О., Сурнин О.Л. Региональная система предоставления государственных и муниципальных услуг населению в электронной форме с применением интегрированных баз знаний и мультиагентных технологий в социальной сфере Самарской области. // Проблемы управления и моделирования в сложных системах: Труды XI Международной конференции. – Самара, Самарский научный центр РАН, 2009. – С. 411-414.
- [10] Сапир Ж. К экономической теории неоднородных систем (опыт исследования децентрализованной экономики). М.: Государственный университет – Высшая школа экономики, 2001.
- [11] Виттих В.А., Ситников П.В., Смирнов С.В. Онтологический подход к построению информационно-логических моделей в процессах управления социальными системами // Вестник компьютерных и информационных технологий, № 5, 2009. – С. 45-53.
- [12] Виттих В.А., Скобелев П.О. Метод сопряженных взаимодействий для управления распределением ресурсов в реальном масштабе времени // Автометрия, № 2, 2009. – С. 78-87.

СОСТОЯНИЕ ЭКОЛОГИЧЕСКОЙ БЕЗОПАСНОСТИ В АТОМНОЙ ОТРАСЛИ

В.Ф. Меньшиков¹, А.Г. Назаров², О.Е. Головихина³

¹Общественный Совет Государственной корпорации Росатом
119017, Москва, ул. Большая Ордынка, 24/26, Россия
os-ra@yandex.ru

тел: +7 (499) 949-21-88, факс: +7 (499) 949-23-20

²Экологический центр Института истории естествознания и техники РАН
117861, Москва, ул. Обручева, 30А, Россия
postmaster@ihst.ru

тел: +7 (495) 938-60-08 факс: +7 (495) 938-60-22

³Секретариат Общественного Совета Государственной корпорации Росатом
119017, Москва, ул. Большая Ордынка, 24/26, Россия
os-ra@yandex.ru

тел: +7 (499) 949-21-88, факс: +7 (499) 949-23-20

Ключевые слова: атомная энергетика, экологическая безопасность, экологически значимые показатели, зоны экологической безопасности предприятий

Abstract

For the first time in all the 65-year history of nuclear industry some unclassified reports have been published regarding environmental safety for all the major enterprises under State Corporation Rosatom. In the aim of assigning related level of environmental safety the whole range of enterprises have been ranked according to their type of activity and evaluated by area of environmental stress. The environmental index was specified on the basis of appropriate expertise using all the 12 strictly normalized environmental factors. Some nonuniformity of the environmental safety field for those nuclear industry enterprises has been revealed.

В России разработана национальная программа перехода от сырьевой к инновационной модели экономического роста, которая начала давать первые результаты. Одним из приоритетных направлений работы по модернизации и технологическому развитию экономики стала атомная энергетика. Здесь принят ряд стратегических решений, важнейшим из которых является принятие федеральной целевой программы по новой технологической платформе и замкнутому топливному циклу «Ядерные энерготехнологии нового поколения на период 2010–2015 годов и на перспективу до 2020 года». Была создана государственная корпорация "Росатом" и за последнее время атомная энергетика становится в некотором смысле фаворитом энергосектора. Государство определило стратегическое направление на развитие высоких технологий в атомной отрасли — одной из немногих, где Россия может конкурировать с зарубежными компаниями "Росатому" принадлежит 40% мирового рынка услуг по обогащению урана и 17% рынка по поставке ядерного топлива для АЭС. Предполагается, что в обозримом будущем России вполне по силам занять не менее 25% мирового рынка услуг по строительству и эксплуатации атомных электростанций.

В настоящее время доля отечественной атомной энергии в общей генерации электроэнергии порядка 16 %. В планах - довести ядерную генерацию сначала до 25% к 2025 г, а затем и до 40%. Для достижения таких показателей в России необходима господдержка атомной энергетике в ближайшие годы, и не только финансовая, но и политическая. Согласно программе развития, планируется потратить до 2018 г. 1,47 трлн руб., из них 674 млрд составят бюджетные средства. Из федерального бюджета на строительство АЭС в 2010 году

выделено более 68 млрд рублей, а всего в рамках инвестпрограммы концерна "Росэнергоатом" в этом году планируется вложить в атомные энергетические объекты 175 млрд рублей. Предполагается, что после 2018 г. господдержка отрасли будет сведена к нулю. Но без этой огромной господдержки невозможно выполнить задачу, сформулированную еще в 2006 г., развить нашу атомную энергетику до мирового уровня в плане выработки электроэнергии. Причем делать это предполагается в ускоренном темпе. Если за 1990-е - начало 2000-х в стране было введено лишь три новых энергоблока (всего 31 энергоблок в настоящее время), то в 2010-2030 гг. будет построено 26 новых блоков. На данный момент в России строится 10 новых энергоблоков, еще 5 – за границей. Выделение средств на атомную энергетику означает инвестиции и в отечественное машиностроение. А часть из них пойдёт на фундаментальную науку, образовательные программы, подготовку специалистов. В 21 веке отечественной ядерной энергетике потребуется полное обновление реакторного парка и современные высококвалифицированные кадры. Представляется, что "момент истины" наступит в 2014 году, когда определится тройка лидеров в новой мировой ядерной гонке - к безопасным реакторам с рециклингом топлива и дожиганием отработанного топлива. Безопасность АЭС является абсолютным приоритетом, как в нашей стране, так и в развитии мировой атомной энергетики. Следующей важнейшей задачей является обеспечение радиационной и экологической безопасности, где остается много нерешенных проблем.

Для нашей страны есть еще одна особенность. Атомная отрасль с момента ее становления была одной из самых закрытых среди других крупнейших производств. Главной причиной секретности являлось нахождение гражданского и военного сектора в одном ведомстве, в Министерстве по атомной энергии (Минатом) и общей атмосфере закрытости государства. Один из наглядных примеров – сокрытие крупнейшей радиационной катастрофой в истории человечества на тот период, получившей название Кыштымской аварии. И лишь через 30 лет она уступила этот страшный пьедестал Чернобылю. Катастрофа произошла 29 сентября 1957 г. на химическом комбинате ПО «Маяк», находящемся в небольшом закрытом городе «Челябинск-65» (в 1994 году он был переименован в город Озерск). Кыштымской авария зовется потому, что до недавнего времени Озерск был засекречен и отсутствовал на картах, а самым близким к нему городом был Кыштым. На комбинате проходил процесс изготовления материалов для атомной бомбы. Тогда из-за поломки системы охлаждения, взорвалась емкость, в которой содержалось более 85 тонн чрезвычайно опасных высокоактивных ядерных отходов. Взрыв был такой силы, что бетонное перекрытие весом 165 тонн было откинута на 25 метров как пушинка, а во всех зданиях в радиусе километра выбило окна. В результате взрыва в атмосферу попало более 20 млн. кюри радионуклидов. Большая часть радиоактивных веществ осела на территории ЗАТО (закрытого административно территориального образования химкомбината «Маяк»), а остальные подняло взрывом на высоту 1.5 – 2 км, где образовалось радиоактивное облако, состоящее в основном из газов и аэрозолей. Облако по направлению ветра устремилось на северо-восток и через несколько часов радиоактивные осадки выпали на территорию свыше 300 км. В загрязненной радионуклидами зоне оказалась территория площадью более 23 тысяч квадратных километров, включающая в себя 217 населенных пунктов с более 280 тысячами жителей, которых пришлось в срочном порядке переселять на другие территории. Эта зона с тех пор называется Восточно-Уральский радиоактивный след.

Прервать завесу секретности и законодательно начать организовывать помощь и реабилитацию пострадавших в радиационных авариях стало возможным после Чернобыльской аварии на рубеже 90-х годов прошлого века. Однако еще много лет вопросы радиационной и экологической безопасности на предприятиях Минатома России были недоступны для обсуждения независимыми экспертами и общественными организациями. Ситуация кардинально изменилась в последние пять лет в период преобразования атомного ведомства в Государственную корпорацию «Росатом» и с приходом нового руководства

отраслю. В феврале 2006 года в ГК «Росатом» (тогда Федеральное агентство по атомной энергии) был создан Общественный Совет для взаимодействия с российскими и международными общественными организациями в рамках обсуждения вопросов, связанных с работой и развитием предприятий атомной отрасли, с решением вопросов ядерно-радиационной безопасности, с решением социальных проблем, на территориях расположения предприятий атомной отрасли. В состав Совета входят ученые и специалисты, занимающиеся проблемами атомной энергии, охраны окружающей среды, ядерной и радиационной безопасности, радиобиологии и радиационной экологии, представители общественных и экологических организаций и объединений. Общественный Совет создавался как открытая площадка для диалога между, экспертным сообществом и общественными организациями. Целью деятельности Совета является коллективная выработка рекомендаций для принятия решений в области использования атомной энергии, охраны окружающей среды, ядерной и радиационной безопасности, охраны здоровья персонала атомных радиационных объектов, ликвидаторов и лиц, пострадавших от радиационных аварий и катастроф, а также населения, проживающего на радиационно-загрязненных территориях и территориях санитарно-защитных зон. К задачам Общественного Совета относятся:

- привлечение институтов гражданского общества к формированию политики в области использования атомной энергии;
- организация обмена мнениями между учеными, специалистами, представителями органов государственной власти, представителями общественных и экологических организаций и объединений, представителями общественности по вопросам, связанным с использованием атомной энергии, реализацией федеральных целевых программ в области ядерной и радиационной безопасности, здоровья персонала атомных радиационных объектов, ликвидаторов и лиц, пострадавших от радиационных аварий и катастроф, а также населения, проживающего на радиационно-загрязненных территориях и санитарно-защитных зон;
- повышение информированности общественности по основным направлениям деятельности Госкорпорации в сфере использования атомной энергии, реализации федеральных целевых программ в области ядерной и радиационной безопасности, ликвидации последствий оборонной деятельности;
- совершенствование взаимодействия Госкорпорации с общественными и экологическими объединениями и гражданами.

К концу 2008 года были завершены разработка, широкое обсуждение на Общественном Совете и принятие Экологической политики Госкорпорации "Росатом". В 2009 году все предприятия "Росатома" представили собственные документы по экологической политике, соответствующие Экологической политике Госкорпорации "Росатом". Проведенная работа явилась первым и методологически обоснованным шагом в оценке реального состояния в сфере экологической безопасности многочисленных атомных предприятий.

К сожалению, значительное число предприятий не смогло добиться доведения этого важнейшего документа до требований, которые определены в принятой Экологической политике Госкорпорации «Росатом», получившей высокую оценку внешних аудиторских проверок. За очень редким исключением (НПО «ЛУЧ», Ангарский электролизный химический комбинат (АЭХК), Приаргунское производственное горно-химическое объединение (ПГХО), Корпорация ТВЭЛ), в большинстве представленных документов не разработана и не декларирована научно обоснованная перспектива совершенствования экологической политики и экологической культуры в планах развития предприятий на ближайшее пятилетие. Частично это отражено в планах реализации экологической политики предприятий, но еще требует дальнейших больших усилий. Общественным Советом «Росатома» была проведена научно-экспертная оценка отчетов по экологической безопасности предприятий Госкорпорации. Для оценки степени экологической безопасности предприятия или группы (категории)

предприятий по видам деятельности. Соответственно, по его величинам были выделены «зоны экологической напряженности» предприятий и их категорий.:

10 АЭС (Балаковская, Белоярская, Билибинская, Волгодонская. Калининская, Кольская, Курская, Ленинградская, Нововоронежская, Смоленская);

10 предприятий, расположенных в ЗАТО и АЭХК;

10 производственных комбинатов;

17 научно-исследовательских институтов и конструкторских бюро;

16 федеральных предприятий по сбору и утилизации радиоактивных отходов (РосРАО).

Методика проведения анализа:

- Применение комплексной системы экологических критериев;
- Определение показателя экологической безопасности предприятия;
- Нормирование показателя экологической безопасности предприятия;
- Определение поля экологической безопасности предприятий;
- Определение показателя экологической безопасности для АЭС;
- Определение показателя экологической безопасности градообразующих предприятий ЗАТО+АЭХК;
- Определение показателя экологической безопасности производственных комбинатов;
- Определение показателя экологической безопасности НИИ
- Определение показателя экологической безопасности ФГУП«ДальРАО», ФГУП«СевРАО», филиалов ФГУП«РосРАО».

Экологический показатель определялся на основе экспертной оценки каждого из 12-ти строго нормируемых экологически значимых показателей (ЭЗП), по которым оценивается экологическая деятельность предприятия или их категорий по видам деятельности. К ним относятся: сбросы вредных химических веществ и радионуклидов в открытую гидрографическую сеть, наличие (или отсутствие) систем экологического менеджмента и менеджмента качества, производственный экологический контроль, забор воды из водных источников, выбросы в атмосферный воздух, обращение с отходами производства и потребления и с РАО, удельных выбросов, сбросов и отходов атомного объекта в общем объеме, рекультивация загрязненных территорий. Следовательно, вне зависимости от его названия, суммарный экологический показатель учитывает основные виды воздействия предприятий или их группы на окружающую среду и, опосредованно, на здоровье персонала и окружающего населения.

Комплексная оценка экологических критериев (ЭЗП подчеркнуты):

Общая характеристика объекта

Основная деятельность объекта

Экологическая политика объекта

Основные документы, регулирующие природоохранную деятельность объекта

Системы экологического менеджмента качества

Производственный экологический контроль

Воздействие на окружающую среду:

Забор воды из водных источников

Сбросы в открытую гидрографическую сеть

Сбросы ВХВ

Сбросы радионуклидов

Выбросы в атмосферный воздух

Выбросы ВХВ

Выбросы радионуклидов

Отходы

Обращение с отходами производства и потребления

Обращение с радиоактивными отходами

Удельный вес выбросов, сбросов отходов объекта в общем объеме по территории расположения объекта

Загрязненные территории и их рекультивация

Реализация экологической политики в отчетном году

Экологическая и информационно-просветительская деятельность

Обобщенный показатель, учитывающий суммарную оценку 12-ти экологически значимых показателей каждого предприятия, предложено называть «показатель экологической безопасности» (предприятия, категории предприятий по видам деятельности, атомной отрасли в целом). Этот показатель отражает тот уровень экологической безопасности, который сложился к состоянию первого обязательного представления экологических отчетов. Это условно начальный уровень, с которого начинается последующая динамика экологической деятельности предприятия. Начальный же, условно «нулевой» уровень экологической безопасности фиксирует все годы становления и развития предприятия, включая и те, когда понятие «экологической безопасности» не входило в практику деятельности. В большинстве случаев «возраст» предприятий составляет от 20-30 лет до 50-60 лет и более для предприятий, созданных с момента организации атомной отрасли (Радиохимический институт им. В.Г. Хлопина, ПО «Маяк» и др.).

Всего выделено 4 зоны: от «зоны допустимой безопасности» нормально функционирующего предприятия, соответствующего всем нормативным требованиям и экологически значимым показателям безопасности (ЭЗП) до четвертой «зоны высокого риска» (R1-R4). В зону R5, требующей оперативного реагирования, ни одно из предприятий атомной отрасли в нее не попадает (таблица 1).

Таблица 1 – Определение показателя экологической безопасности предприятия

Ранг 1 (R1)	Ранг 2 (R2)	Ранг 3 (R3)	Ранг 4 (R4)	Ранг 5 (R5)
Соответствие всем ЭЗП (без замечаний)	Незначительные отступления от ЭЗП (1 или 2 ЭЗП)	Существенные отступления от ЭЗП (1-3 ЭЗП)	Существенные отступления от ЭЗП (более 3 ЭЗП)	Существенные отступления от большинства ЭЗП

$P = \Sigma R_p / \Sigma R_{ip}$ – показатель экологической безопасности предприятия,

где ΣR_p – сумма рангов 12 ЭЗП,

ΣR_{ip} – сумма рангов 12 ЭЗП идеального предприятия, ($\Sigma R_{ip} = 12$)

Проведенное нормирование показателя экологической безопасности дает следующие данные (таблица 2):

Таблица 2 – Зоны экологической безопасности

IV зона	> 4,0	Зона реагирования	0 предприятий
III зона	> 2,5 – 4,0	Зона высокого риска	19 предприятий
II зона	> 1,5 – 2,5	Зона допустимого риска	41 предприятий
I зона	> 1,0 – 1,5	Зона допустимой безопасности	6 предприятий

Поле экологической безопасности предприятий, институтов и учреждений Госкорпорации Росатома показано на рисунке 1.

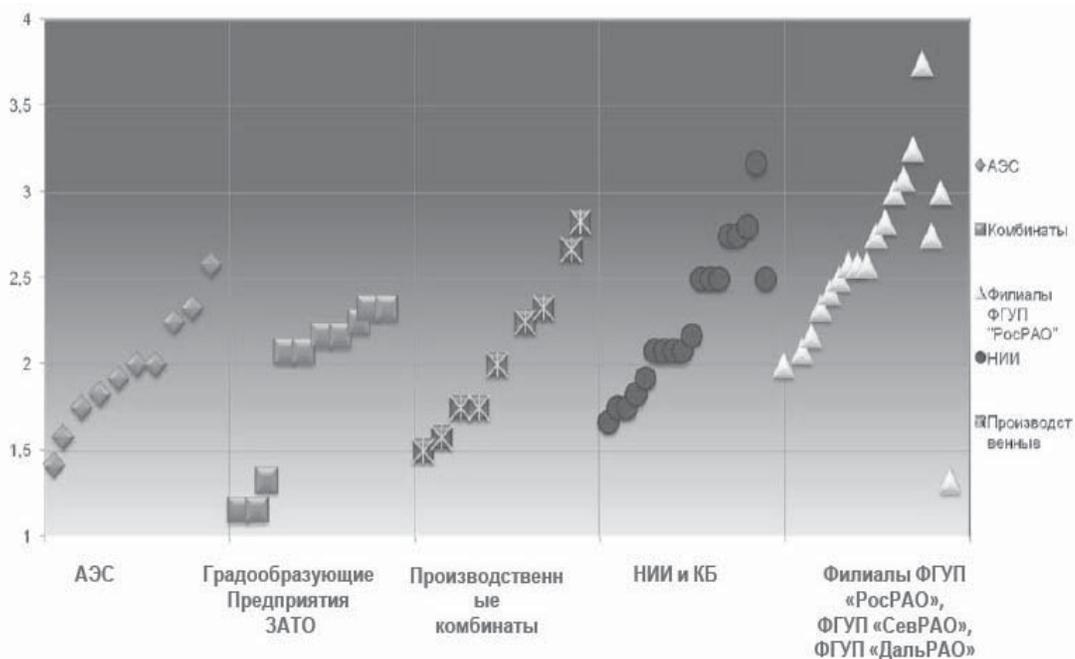


Рисунок 1 – Поле экологической безопасности предприятий

Как видно из приведенного рисунка, есть большие различия в уровнях экологической безопасности у всех категорий предприятий, НИИ и КБ. В первую очередь, мы должны сосредоточить внимание и анализ экологической обстановки к атомным электростанциям, как к объектам с близким расположением от места проживания населения. Среди научного и экспертного сообщества продолжается дискуссия о влиянии разрешенных выбросов и сбросов радионуклидов при работе АЭС на здоровье населения и окружающую среду. Среди огромного набора радиоизотопов, образующихся при штатной работе АЭС, есть и не учитываемые по разным причинам опасные радионуклиды. Существует достаточно большое количество научных исследований о последствиях дополнительного облучения в малых дозах на окружающее население и живую природу. Все эти дискуссионные моменты не позволили учитывать в данном исследовании эти возможные опасности и включать их в набор экологически и радиационно значимых факторов безопасности.

Для будущей объективной оценки экологической безопасности АЭС необходимо требовать в отчетах этих атомных предприятий информацию с обращением отработавшего ядерного топлива (ОЯТ). К сожалению, во всех странах, где работают атомные станции, нет удовлетворительного решения этой важнейшей для атомной энергетики проблемы. В России на пристанционных объектах хранения и в хранилище на Горно-химическом комбинате (г. Железногорск, Красноярского края) накоплено около 20 тысяч тонн ОЯТ. Заполнение некоторых пристанционных хранилищ подходит к рассчитанному пределу. Планы по регенерации ОЯТ и использования этого чрезвычайно радиоактивного материала для атомной энергетики требуют независимой научно-технологической и экономической экспертизы. В настоящее время в нашей стране нет даже регулирования на законодательном уровне чрезвычайно важной проблемы обращения с ОЯТ.

Учитывая эти, выше приведенные, замечания показатели экологической напряженности АЭС с учетом года ввода первого энергоблока можно представить на следующем рисунке (рисунок 2):

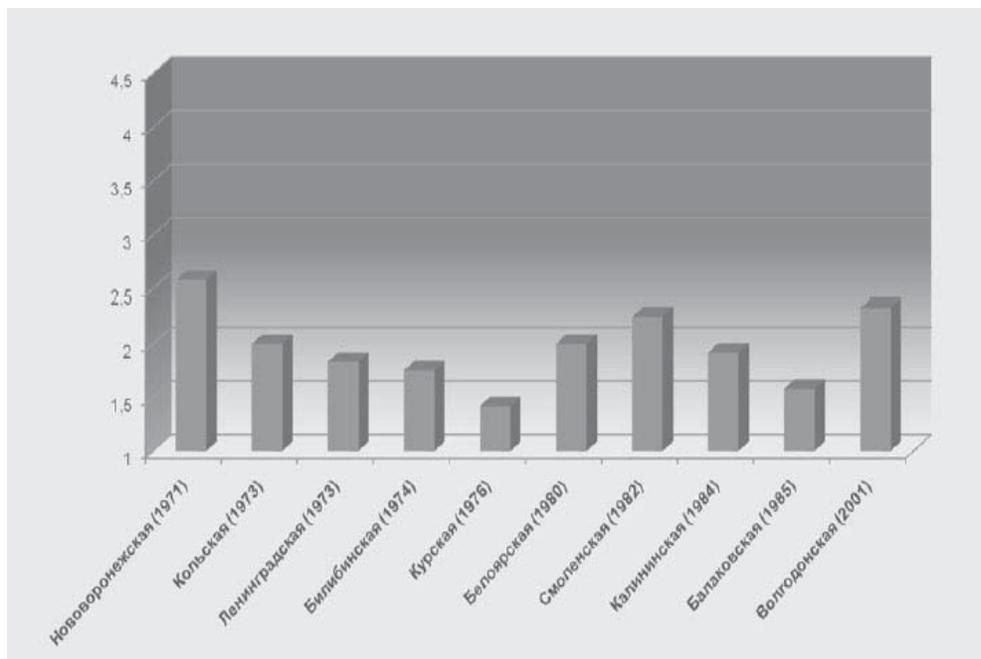


Рисунок 2 – Показатель экологической напряженности АЭС с учетом ввода первого энергоблока

В системе обращения с радиоактивными отходами (РАО) за последнее время произошли существенные позитивные изменения. В 2008 году Указом Президента Российской Федерации 15 специализированных предприятий «Радон» были переданы под управление ГК «Росатом» и было образовано ФГУП «РосРАО». Спецкомбинаты «Радон» были созданы в конце 50-х годов прошлого века для удаления и обезвреживания низкоактивных РАО и отработавших источников ионизирующего излучения из крупных городов и промышленных центров: Эти предприятия в настоящее время вошли в состав семи территориальных округов, сформированных согласно административному территориальному делению России и стали единственными профессионально . Была сформирована вертикально интегрированная структура, которая позволила расширить зоны ответственности филиалов, более эффективно осуществлять управление филиалами, принять активное участие в формировании Единой системы обращения с РАО в России, включая радиационное обследование и реабилитацию загрязненных территорий. Таких территорий за десятилетия развития военной и гражданской составляющей атомной отрасли в России оказалось недопустимо много. Перед «РосРАО» поставлена задача инвентаризации существующих объектов долговременного хранения РАО и существующих технологий обращения с РАО, а также формирование генеральной схемы размещения объектов долговременного хранения и окончательной изоляции РАО на базе действующих площадок «РосРАО». В рамках Федеральной целевой программы «Обеспечение ядерной и радиационной безопасности на 2008 год и на период до 2015 года» решаются неотложные вопросы по реабилитации загрязненных территорий. Необходимо также решить задачу систематизировать типы образующихся РАО по федеральным округам для приповерхностной или подземной окончательной изоляции в зависимости от типа отходов. Региональная структура мест дислокации пунктов долговременного хранения и окончательной

изоляции, которая будет создана в каждом федеральном округе, сейчас прорабатывается исходя из нескольких критериев. К примеру, в Приволжском федеральном округе для этих целей выбрано Саратовское отделение «РосРАО». Распределение показателей экологической безопасности ФГУП «РосРАО» и двух организаций, занимающихся в основном сбором и утилизацией РАО от военных объектов («ДальРАО» и «СевРАО»), показаны на следующем рисунке (рисунок 3):

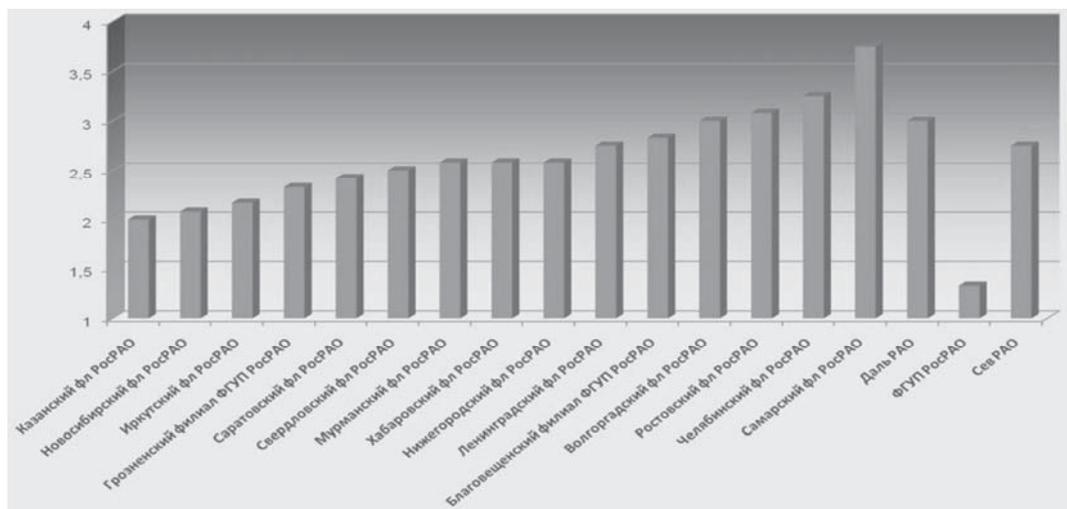


Рисунок 3

Выводы

В 2009 году впервые за 65-летнюю историю атомной отрасли отчеты по экологической безопасности всех экологически значимых предприятий Госкорпорации «Росатом» были опубликованы в открытой печати.

В процессе переработки и обобщения фактического материала отчетов по экологической безопасности полностью подтверждается первоначальный вывод о стабильности и высокой информативности интегрального показателя экологической безопасности. Для существенного его изменения в пределах или между границами зон экологической безопасности требуются экстремальные воздействия предприятия на окружающую среду при радиационных или техногенно-экологических авариях и катастрофах.

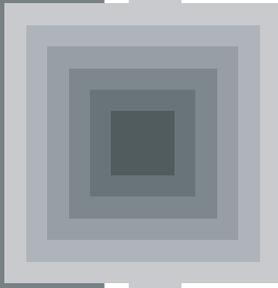
Таким образом, выделенные вертикальные зоны экологической безопасности, от «нормальной» к зоне «высокого риска», и горизонтальный ряд экологической безопасности от АЭС до предприятий РосРАО, отражают объективную начальную («стартовую») картину экологической безопасности атомной отрасли. Картина эта достаточно неоднородная, с превышениями показателя экологической безопасности над нормально функционирующей экологической системой безопасности ($P=1-1,5$) в 3-3,7 раза. Можно обобщить результаты проведенного анализа:

- представлена относительно объективная картина экологического состояния предприятий ГК «Росатом»
- выявлена неоднородность поля экологической безопасности, существующая на предприятиях атомной отрасли;
- группы предприятий по средней величине показателей экологической безопасности располагаются в следующий ряд: АЭС – ЗАТО – Производственные комбинаты – НИИ – Филиалы ФГУП «РосРАО»;

- представляется необходимым укрепление научно-экспертного звена, введения объектного мониторинга состояния недр, завершающего систему радиационно-экологического мониторинга, повышения квалификации работников экологических служб предприятий и организаций ГК«Росатом»;
- открытая экологическая отчетность на базе системы прозрачных экологических критериев необходима для экологически безопасной атомной энергетики.

Список литературы

- [1] Меньшиков В.Ф. Энергетика и экология в послекризисный период // Труды XI Международной конференции «Проблемы управления и моделирования в сложных системах». – Самара: СамНЦ РАН, 2009. – С. 44-56.



System Analysis and Control Theory
Системный анализ и теория управления

OPTIMIZATION TECHNIQUES FOR DESIGN AND CONTROL OF INDUCTION HEATING SYSTEMS

B. Nacke

Leibniz University of Hannover,
Institute of Electrotechnology
30167, Hannover, Wilhelm-Busch-Str. 4, Germany
nacke@etp.uni-hannover.de
phone: +49 (0) 511-762-5533, fax: +49 (0) 511-762-3275

***Key words:** induction heating, numerical simulation, coupled electromagnetic and thermal field, mathematical optimization*

Abstract

Индукционные нагревательные системы используются, в основном, для предварительного нагрева металлических изделий перед обработкой давлением, термообработки, сварки и плавки. Процесс нагрева является достаточно сложным, и часто требуется высокая точность получения пространственно-временного температурного распределения. В большинстве случаев только моделирование процессов нагрева помогает провести анализ процесса, правильно спроектировать индукционную нагревательную систему или найти оптимальные параметры процесса. В статье представлены три примера использования современных оптимизационных методов проектирования и управления индукционными системами. Рассмотрен пример численного моделирования процесса индукционного поверхностного упрочнения деталей сложной геометрической формы. Представлены результаты численного моделирования комбинированного процесса индукционного и лазерного нагрева, используемого для сварки. Описан пример оптимизации процесса индукционного нагрева для планетарного реактора.

Introduction

Induction heating systems are widely used in industry for heating and melting of metals and non-metal materials. In induction heating the main classical fields are heating for deforming, like for forging or rolling applications, heat treatment like the surface hardening of gears and heating for welding and soldering. Today also many modern applications of induction heating are on the way to be implemented in industrial production, like hybrid induction and laser welding, heating applications for non-metal materials like for semiconductor materials or for oxides and glasses.

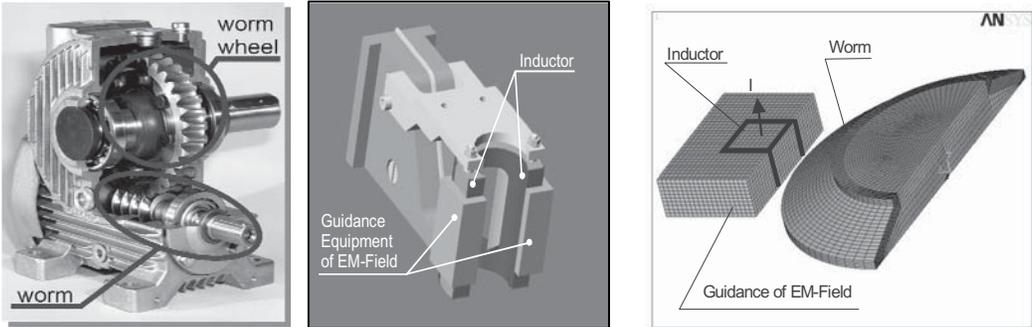
The requirements of all induction heating processes are an accurate time and space dependent temperature distribution in the work-piece depending on the application. In the case of heating for deforming and other applications a homogeneous temperature distribution is necessary while, for example, for surface hardening or welding a highly inhomogeneous temperature distribution is needed. In order to find the proper temperature distribution today modern simulation tools are available, in order to find an optimal solution in short time without long term and costly experimental investigations.

In the following three examples for the use of modern simulation and optimization techniques are presented, starting with two examples for the use of only numerical simulation of the heating process and using the results to find an optimal solution by expert knowledge and finishing with one example for the use of mathematical optimization in combination with numerical simulation which allows to find the optimal solution without special expert knowledge.

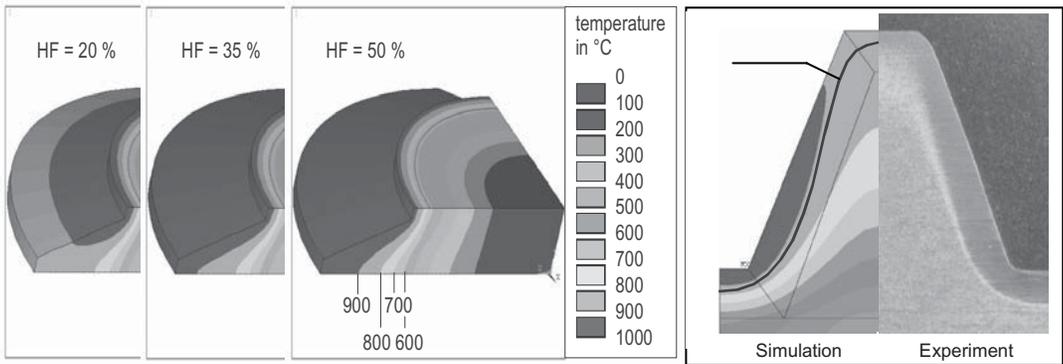
1 Design of complex induction hardening process using 3D numerical simulation

Induction surface hardening is an innovative electrothermal process for improving the quality of mechanical components and for increasing the productivity of production lines. The main advantages are the very short hardening time, high power densities and low space requirements. Therefore the integration into productions line is very easy and effective.

The design of induction hardening devices (Picture 1), especially of the inductor geometry and the choice of the hardening parameters like the inductor current and the necessary power are actually carried out by experience and experiments. But here are limits if complex work-piece geometries have to be handled.



Picture 1 – Complex geometry of a worm gear, the used line inductor and the numerical model



Picture 2 – Optimal heating profile for contour hardening using the simultaneous dual frequency method (SDF) and comparison with experiment

The 3D numerical simulation of the induction heating process (Picture 2 left) offers new chances for the design of necessary induction hardening devices for complex geometries. At the Institute of Electrotechnology new simulation models were developed which can successfully simulate the full 3D instationary heating process [1]. The model allows the solution of the coupled non-linear electromagnetic and thermal fields taking into account all temperature dependences of the material properties. Therefore the simulation program is able to investigate the influence of different hardening parameters on the heat profile in the work-piece and to optimize the hardening process (Picture 2 right). The numerical model additionally is able to simulate induction hardening problems for which several different frequencies are used, like for the simultaneous dual frequency hardening process (SDF) which was investigated for the configuration of Picture 1. Results for different power shares of the high and medium frequency are shown in Picture 2 left.

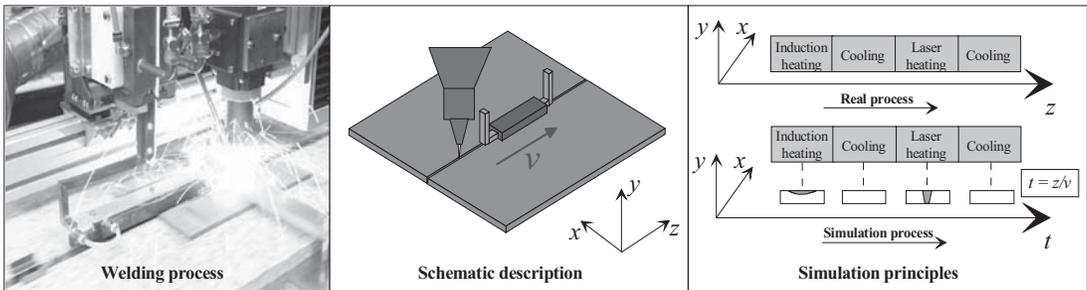
The project was supported by the German Research Foundation AiF.

2 Simulation of induction assisted laser welding using combined numerical simulations

High strength steel grades are coming into an increasing use in many industrial sectors because of their significant advantages as a high wear resistance combined with high toughness, weight saving aspects etc.. Unfortunately, fundamental metallurgical facts cause these steels to be relatively difficult to weld and thereby make their utilization complicated. This limitation can be overcome by the use of appropriate welding methods combined with induction heating. This part presents an overview of the progress in design and optimization of such induction heating systems with regard to their interaction with different welding sources.

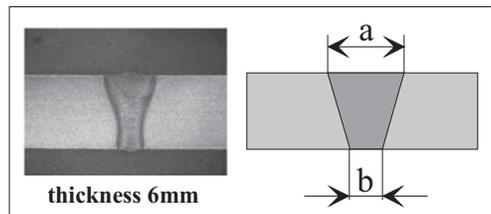
The investigations were performed in the frame of the European project INDUCWELD [2]. Its aim was to increase the weldability of high strength steels by induction assisted welding and it is oriented on higher steel thicknesses (from 3 mm up to 30 mm). In the frame of this project, the Institute of Electrotechnology investigated and designed induction heating systems for different welding techniques and applications. For this purpose, a special numerical model has been developed, which respects the mutual influence of welding and induction heating process and allows the modelling of the thermal behaviour within the complete process.

The model (developed using the commercial software package ANSYS) uses a 2D approach to describe the temperature behaviour of the welding process and respects the movement of the welding sources by their change in time (see Picture 3). It carries out the thermal analysis of different welding procedures as well as the electromagnetic and thermal analysis of induction heating process. As a



Picture 3 – Basic approach used for numerical modeling

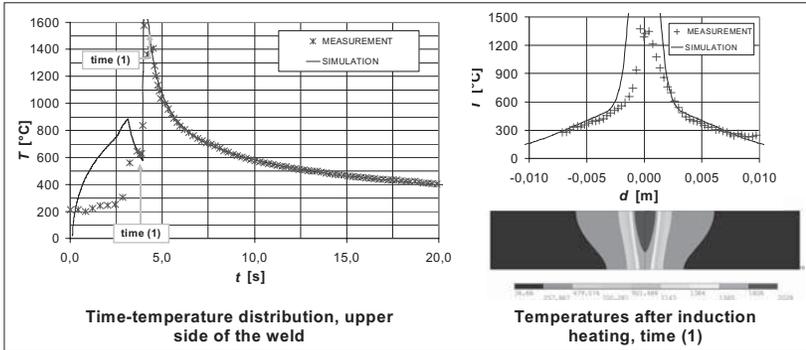
result the model provides the knowledge of the time and space temperature distribution within the welded steel plates and allows herewith the optimization of the induction heating system under actual process condition. Material properties used in the model are temperature dependent and ensure the correct behaviour of the model in the whole temperature range (from ambient temperature to maximum temperature during the welding process). Laser, plasma or GMA impacts are included as 2-dimensional fields of heat sources varying in time. Methods of their implementation, dimensions of the active weld area for different materials and thicknesses etc. have been determined using experimental results (Picture 4 shows an example of the actual shape of the weld and its simplified implementation in the model). The modular structure allows the simulation of all possible technological combinations of the welding process and possibilities of their induction heating support (pure welding process, postheating, preheating etc.).



Picture 4 – Laser and GMA welding energy implementation

Several existing welding installations have been successfully modelled in order to test and verify the capabilities of the model. Picture 5 shows the selected results for a laser welding process

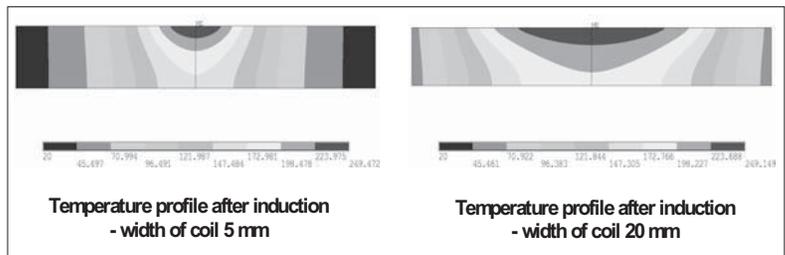
with integrated induction preheating. Both the time behaviour of temperature on the upper weld side (left) and the temperature profile on the upper side (right) show a good correspondence with experimental results. The 2D temperature distribution inside the welded steel parts illustrates one of the main advantages of numerical modelling – knowledge of process thermal behaviour at arbitrary time and location.



Picture 5 – Comparison between numerical results and experiments

welding methods. Various parameters (cross-section and the length of inductor, frequency, distance between laser and inductor etc.) can be investigated under actual process conditions. As an example, the influence of the width of induction coil on temperature profile in GMA welding process with induction preheating is shown in Picture 6.

The introduced simulation model represents a powerful tool for description of thermal behaviour of various welding processes with integrated induction heating support. The functionality has been verified by comparison with experimental data and the results confirm its usability for design and optimization of such induction heating systems.



Picture 6 – Temperature distribution after induction heating for two different inductors

Selected results of inductor optimization performed by means of this model illustrate the potential of the suggested strategy.

3 Design of an induction heater for a planetary reactor using mathematical optimization

Compound semiconductors provide the basis for a large number of electronic and optoelectronic applications in daily life. They are used by diverse range of customers worldwide to build advanced components such as photodiodes and lasers devices. Mass production deposition equipment in the semiconductor industry is based on the planetary concept (planetary reactor). The method of Metalorganic Vapour Phase Epitaxy - the deposition of thin layers in the form of crystal growth from the gas phase - is used in such deposition equipment.

Planetary reactor consists of water-cooled stainless steel housing with quartz and easily replaceable graphite components including susceptor and satellites (see Picture 7). The disc shape susceptor is used as holder for the wafers. The wafers are very thin crystal discs, on which the gas mixtures are deposited. The wafer holders, satellites, turn friction-free on gas cushions.

During the process for the production of compound semiconductor layers, the raw materials, gases, are dispersed inside the reactor across the wafers. To obtain the crystal growth deposits, the wafers must be heated up to the process temperature of above 1100°C. The heating is provided by a high-frequency induction system. An induction coil of “pancake” shape is mounted below the susceptor, where the Joule heat is generated.

The graphite susceptor is covered with thin layer of SiC to protect the graphite against chemically aggressive gas. The production cycles with heating up and cooling down lead to significant thermal stresses in the susceptor and the SiC layer. These stresses can result to cracking the SiC layer and its lifetime is reduced. The stress distribution in the susceptor and SiC layer depends on their temperature distributions. To provide as uniform as possible temperature field in the susceptor means to reduce the level of stresses in the SiC layer.



Picture 7 – View of the planetary reactor

The steady-state temperature field in the susceptor depends on the shape of the induction coil while the transient temperature distribution can be controlled by electric power. Optimization of the induction coil shape and investigation of temperature transients can be done by numerical simulation.

3.1 Numerical models

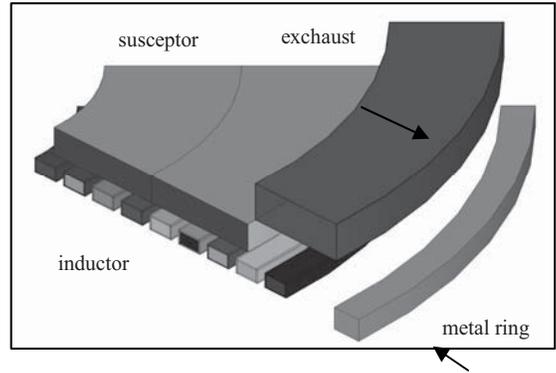
To solve the problem, a numerical simulation strategy was developed [3]. Generally, the problem requires a three-dimensional solution because the “pancake” shape induction coil is 3-dimensional. In spite of the susceptor is of axisymmetric disk shape, it has several technological drillings. So, the susceptor geometry is 3-dimensional as well. Finally, main cracks in the SiC layer depend on how big is the tangential component of the thermal stresses. This component of the stresses cannot be calculated in 2-dimensional axisymmetric approach. At the same time, the use of 3-dimensional modelling for mathematical optimization is extremely ineffective because of long runtime required.

To find a compromise, both 2- and 3-dimensional modelling have been applied using the commercial package ANSYS. 2-dimensional axisymmetric approach based on coupled electromagnetic and thermal analysis was used together with mathematical optimization to search for the optimal induction coil shape. 3-dimensional electromagnetic and thermal model has been created to investigate how the drillings inside the susceptor influence on the electromagnetic field distribution and, finally, on the temperature distribution in the susceptor. The structural analysis must be carried out in 3 dimensions all the time. Temperature distribution, received in the thermal analysis, is used as input data for the structural one. A special converter has been developed to input the temperature field from 2-dimensional thermal solution into the 3-dimensional structural model.

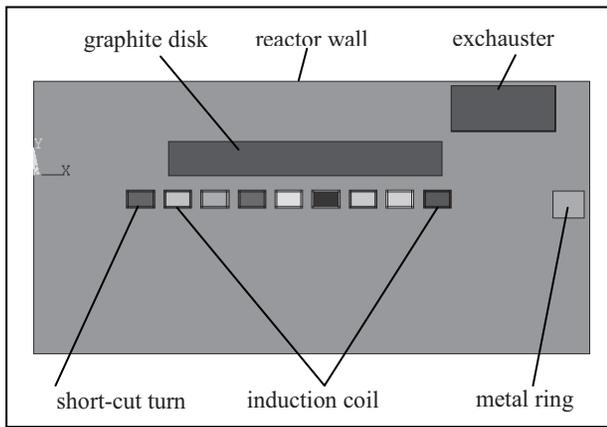
All created models include the electromagnetic and the thermal analysis as well as the structural calculation. They can simulate the steady-state or the transient heating and cooling processes in the planetary reactor. Temperature dependent material properties are taken into consideration by coupling between the electromagnetic and the thermal calculations. Heat exchange intensity on the susceptor surface is an extremely important parameter for the thermal analysis. The heat flux from the susceptor surface is a result of convection and radiation phenomena. The intensity of both convection and radi-

tion strongly depends on the gas temperature around the susceptor. Different effective values of the convective heat exchange coefficient, emissivity factor and reference temperature for all the susceptor surfaces were used.

The electromagnetic environment in the model includes all electromagnetically active components of the planetary reactor – induction coil, susceptor, exhauster and metal ring (see Picture 8) – as well as non-conductive space around them. The thermal environment consists of the graphite susceptor with SiC layer. The structural environment includes only the graphite susceptor and SiC layer as well. The temperature distribution from the thermal



Picture 8 – Electromagnetic components of the reactor (a sector of 45°)



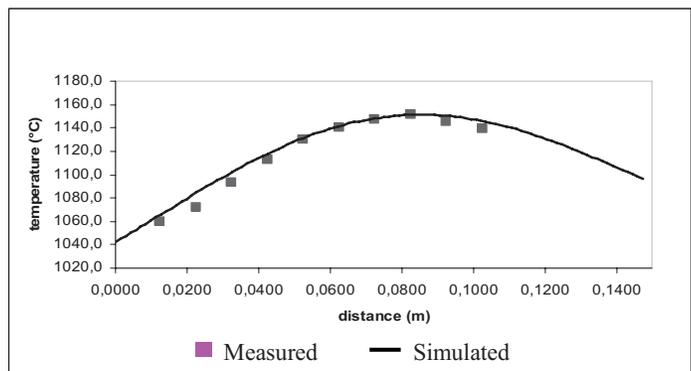
Picture 9 – 2-dimensional approach of modelling

analysis is used to calculate the thermal stresses and deformation as a result of the heat expansion phenomenon.

The 2-dimensional axisymmetric approach (see Picture 9) is based on the assumptions that the induction coil of spiral shape is presented as a set of concentric windings with the same current and the susceptor has no drillings. 3-dimensional models include the drillings inside the susceptor but no spiral shape of the coil is taken into account. This assumption can be used because of the susceptor rotation, which equalizes the temperature field in tangential direction.

3.2 Numerical simulation

The planetary reactor with the standard induction coil, shown in Picture 9 has been simulated in two dimensions and compared with experiments. The temperature profile on the top surface of the susceptor was measured. Picture 10 shows the measured and the simulated temperatures. Very good coincidence of numerical results with experimental data allows using the models for optimization of the induction coil. Additionally, to estimate the level of thermal stresses, the received temperature field has been used in 3-dimensional structural analysis of the susceptor without technological drillings. The calculations show that the



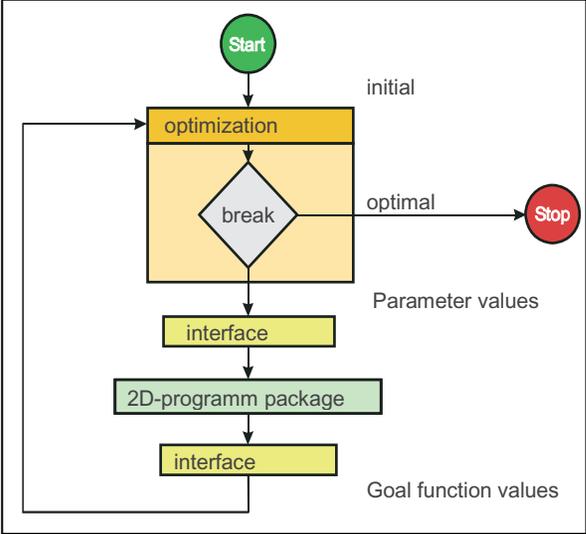
Picture 10 – Measured and simulated temperatures of the susceptor

maximum tension in the SiC layer was lower than the critical value for this material. However, the structural analysis, made for the susceptor with drillings, detected that the local tensions are concentrated around the ends of the drillings. Their level can exceed the level of tensions in the susceptor without drillings in two times. Thus, they can lead to cracking the SiC layer. To reduce this negative effect, the temperature distribution in the susceptor should be provided as homogeneous as possible by modifying the induction heating system.

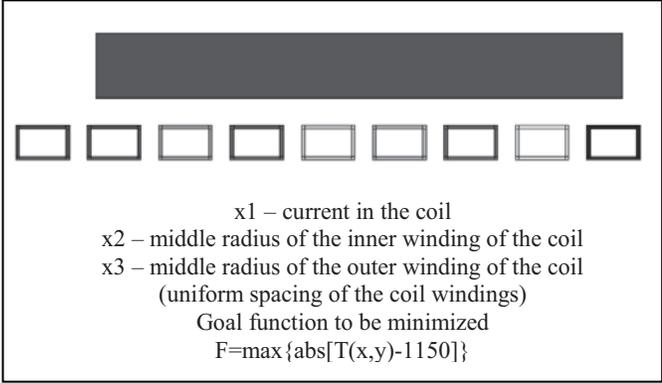
3.3 Optimization of the induction coil shape

The steady-state temperature distribution in the susceptor can be mainly homogenized by optimization of the induction coil shape. To solve the problem, the 2-dimensional electromagnetic and steady-state thermal model has been combined with mathematical optimization search code like it is shown in Picture 11. The optimization search is based on a Genetic Algorithm. Additionally, the data exchange between the model and the optimization code is organized via external data files. The set of independent optimization variables is transformed to the induction coil geometry in the INPUT interface before to run the model. The received temperature distribution in the susceptor is a basis to calculate the goal function value in the OUTPUT interface.

The next important step of the search preparation is the selection of the independent variables and formulation of the goal function. The optimization was started with 3 independent variables: two of them represent the coil geometry, while the coil current is used to reach the necessary temperature level (see Picture 12). The coil consists of nine windings of fixed profile. The required inner and outer radius of the coil is reached by spacing between the



Picture 11 – Mathematical optimization method

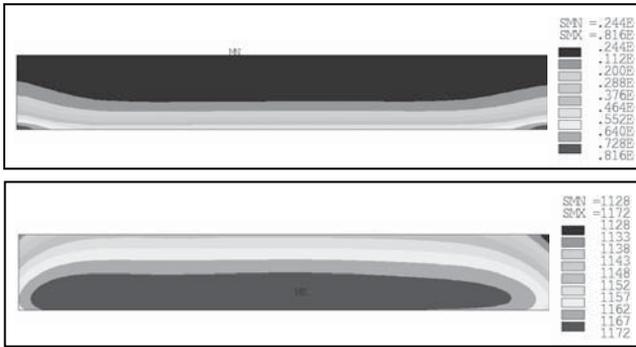


Picture 12 – Independent variables and goal function formulation

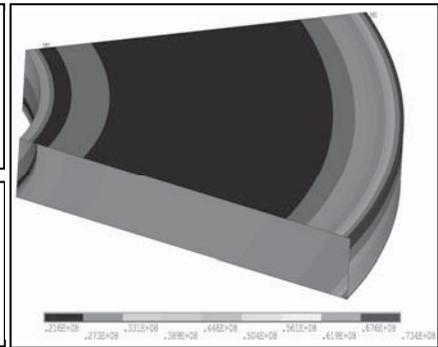
windings, which was chosen uniform. To find as uniform as possible temperature distribution in the susceptor the goal function was formulated as the maximum deviation of temperature from the prescribed value of 1150°C. This goal function should be minimized during the optimization procedure.

The found optimal coil shape provides the results shown in Picture 13. The Joule heat distribution has increased density near the inner and the outer borders of the susceptor to compensate additional thermal losses

in these regions. The temperature field is very uniform in the regular zone of the susceptor. The maximum temperature deviations are in the upper corners because of high heat losses. Nevertheless,



Picture 13 – Joule heat distribution and temperature field in the susceptor provided by the optimal induction coil



Picture 14 – Tangential component of tensions layer for the optimal coil

the reached uniformity of temperature field in the susceptor is of ± 22 K against ± 66 K in case of the standard induction coil.

The optimal temperature distribution was input to the structural model without technological drillings in the susceptor. The tangential component of tensions in the SiC layer is shown in Picture 14. The most dangerous tensions are concentrated at the outer boarder of the susceptor. However, the level of the tensions is much lower than in case of the standard coil.

The induction coil of the found optimal shape has been manufactured and tested in a laboratory set-up and afterwards in an industrial planetary reactor. The predicted temperature field of minimized deviation was confirmed by measurements. The industrial use of the optimal coil increased significantly the lifetime of the susceptors.

Literature

- [1] Wrona, E.; Nacke, B.; Schwenk, W.: Solving complex induction hardening tasks by numerical methods. Proceedings of the International Symposium on Heating by Electromagnetic Sources. Padua, June 22-25, 2004, ISBN 88-86281-92-7, pp. 593-598.
- [2] Mach, M.; Schuelbe, H.; Nacke, B.: Investigation and design of induction assisted welding processes. Proceedings of international UIE conference, Cracow, Poland, May 2008 and Przegląd Elektrotechniczny, R. 84, No. 11/2008, p. 128.
- [3] Nikanorov, A.; Zedler, T.; Nacke, B.: Optimized induction heating in planetary reactors for semiconductor industry. Proceedings of International Symposium on Heating by Electromagnetic Sources, Padua, Italy, June 20 – 22, 2007.

ANALYSIS OF METALLURGICAL MELTING PROCESSES USING NUMERICAL SIMULATION

E. Baake

Institute of Electrotechnology, Leibniz University of Hannover
30167 Hannover, Wilhelm Busch Str. 4, Germany
baake@etp.uni-hannover.de
phone: +49 (0) 511-762-3248, fax: +49 (0) 711-762-3275

Key words: *metallurgical processes, induction melting, numerical simulation, large eddy simulation, heat and mass transfer*

Abstract

Для эффективной реализации металлургических процессов важно уметь анализировать явления тепло- и массопередачи в металле, расплавленном в индукционных печах. Исследования, представленные в данной статье демонстрируют возможности использования известного LES (Large Eddy Simulation) подхода при создании переходных трехмерных моделей явлений тепло- и массопередачи в промышленных металлургических процессах. Результаты моделирования хорошо согласуются с экспериментальными данными.

Introduction

Industrial metallurgical processes like melting of alloys in induction furnaces has become a subject of numerical modelling since many years. A wide range of different modelling approaches for the simulation of the turbulent melt flow and the heat and mass transfer processes have been developed. But up to now the question about an universal and always reliable modelling approach, which can be used for the development and design of industrial metallurgical applications, remains open.

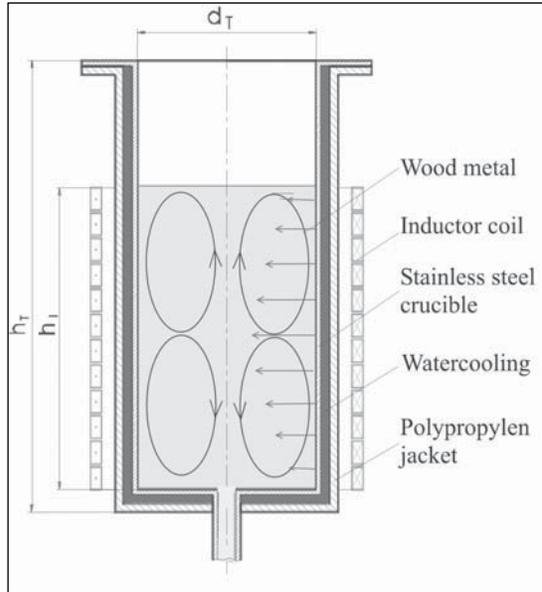
Melting of alloys in induction crucible furnaces can be mentioned as a wide spread example of numerical modelling, because this process can be approximated with two-dimensional (2D) axial-symmetric model. The flow pattern in these installations is formed by the influence of electromagnetic forces and usually comprises of two or more toroidal dominating recirculating vortices. Flow patterns obtained with two-dimensional solvers based on Reynolds Averaged Navier-Stokes (RANS) equations usually are in good agreement with estimated and measured time-averaged flow velocity values. The resulting spatial distribution of the temperature and alloys compound concentration depends strongly on the heat and mass exchange between vortices of mean flow. Numerical investigations show that two-dimensional turbulence models, e.g. $k-\epsilon$ and others, fail to describe correctly the heat and mass transfer processes between the main vortices.

At the present time, different modelling techniques are being used to achieve better agreement with the experiment [1-3]. Our own engineering approach developed for this problem is described in [4], however, it is necessary to investigate advanced simulation methods for more generic and therefore universal flexible solutions. Due to the permanent growth of accessible high powerful computational resources, nowadays, it is possible to run more complicated transient and three-dimensional (3D) numerical calculations of fluid dynamic problems using advanced turbulent models with higher time and volume resolution requirements and to get reliable results in reasonable time.

Concluding all these preconditions the calculations presented in this paper are devoted to the application of Large Eddy Simulation (LES) method for turbulent recirculating flows. The flows of this character often occur in various industrial processes where liquid metal is driven by electromagnetic forces.

1 Experimental set-up and measurement results

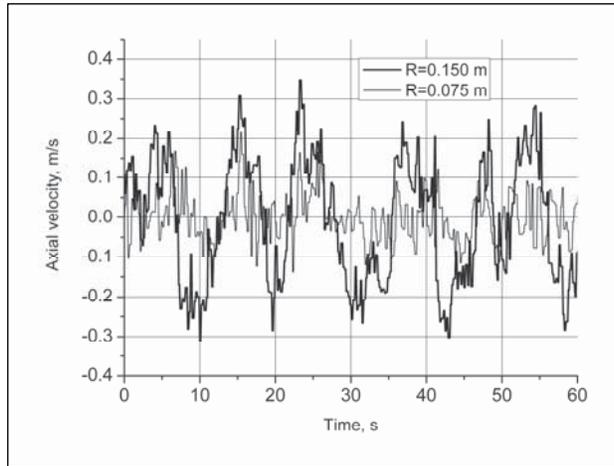
The experimental investigations of the melt flow velocities and temperature distributions in the melt are carried out in different laboratory induction furnaces like an induction crucible furnace (ICF), shown in Picture 1, where Wood's metal is used as the model melt.



Picture 1 – Laboratory induction crucible furnace with sketch of typical Lorentz force distribution and vortices of mean flow

The particular experimental ICF shown in the Picture 1 has a radius of 158 mm and a height of 756 mm, where the inductor height is 570 mm. Wood's metal, which has a melting point of 72°C , and a dynamic viscosity of $4.2 \cdot 10^{-3} \text{ kg/m}\cdot\text{s}$, a density of $9,700 \text{ kg/m}^3$ and a conductivity of $1 \cdot 10^6 \text{ S/m}$ serves as a model melt. Potential difference velocity probe with incorporated permanent magnet [5] is used to measure simultaneously two of three velocity's components with scanning rate from 20 up to 100 Hz. The water-cooled crucible wall keeps the melt temperature at the constant level slightly above 80°C , therefore thermal gradients are negligible and have no influence on the melt properties.

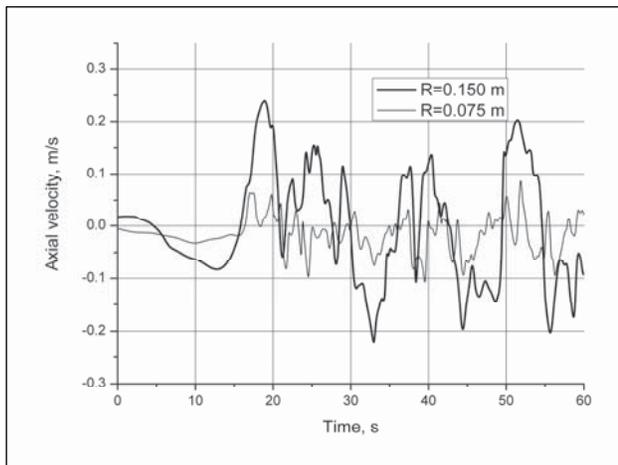
The flow pattern measured consists of two typical toroidal vortices of the mean flow. But, in the same time, velocity measurements revealed, that the melt flow is turbulent with $Re = v_{ch}R/\nu \sim 10^5$ and the amplitude of velocity oscillations is comparable with the time averaged flow velocity magnitude $|v'_{max}| \geq |v_{ch}|$. Also, the presence of low-frequency flow oscillations was exposed: most intensive of them are located close to the crucible wall between the main vortices and have the characteristic time period about 8-12 seconds depending on inductor current (Picture 2). Pulsations with shorter period (about 1-2 seconds) can be observed as well. The major oscillation frequency f increases in dependence on the time-averaged velocity: $f \sim v_{ch} \sim I_{ind}$.



Picture 2 – Axial component of the melt flow velocity oscillations measured in experimental ICF in the region between the dominating flow eddies near the crucible wall ($P \sim 55$ KW)

2 Numerical Modelling

For the numerical investigations of the turbulent melt flow as well as the heat and mass transfer two turbulence models were applied. The first was the well-known $k-\epsilon$ model, which has relatively low mesh requirements and is widely used and verified in various numerical engineering applications. This model usually produces fast good quantitative results for the time-averaged velocity distribution in case of stationary two dimensional calculations, but fails to describe correctly the heat and mass transfer quantities in the melt when the system contains at least two dominating recirculating flow eddies.



Picture 3 – Axial velocity oscillations in the ICF calculated with LES turbulence model

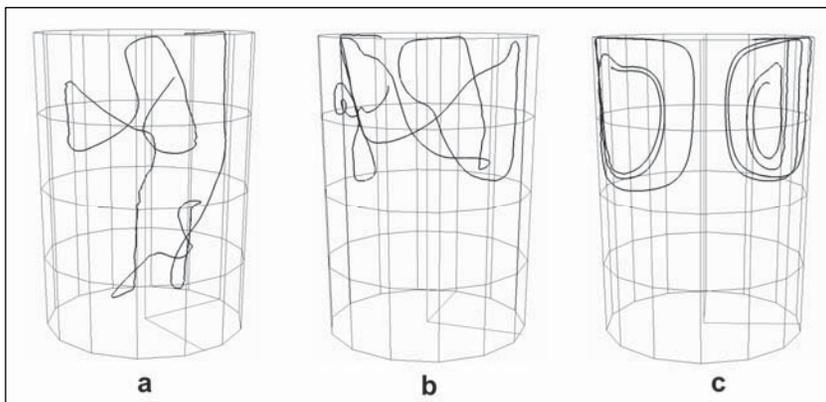
The main reason is that $k-\epsilon$ equations are unable to describe low-frequency pulsations, which arise due to the large scale flow dynamics. Furthermore, the discussed $k-\epsilon$ model is based on the hypothesis of isotropic turbulence, but experiments show, that significant anisotropy takes place not only close to the crucible walls, where axial oscillations are dominating, but also at the symmetry axis, where radial oscillations are prevalent. Relative isotropy can be observed only in the centres of large eddies. The Reynolds stress model can be applied in order to take into account local flow anisotropy at the cost of additional equations to be solved. The unsteady modification of this approach (URANS) allows to perform transient investigations of the flow [3].

Numerical investigations with 3D LES turbulence model revealed flow instabilities similar to those observed in the experimental induction crucible furnace (Picture 2). In order to investigate the reliability of these results, up to two minutes of the flow development were calculated after refining the mesh. Examining results, it was found out that axial flow velocity component oscillates with the amplitude of approximately 20 cm/s near the crucible wall in the region of vortexes interaction (Picture 2), but at the half-radius of the crucible these oscillations are approximately two times less intensive. The main oscillations, which have period about 10 seconds, are combined with less intensive high frequency ones, like in the experiment. The amplitude of these oscillations remains the same for different mesh resolution levels and coincides with experimental data.

Comparing the graphs with the Fourier analysis of the experimental and numerical data, one can see that the calculated results include not only the main frequency, but they also show several additional pulsations with higher and lower frequencies. The presented graph shows that the spectrum of the numerical results is shifted in the direction of the lower frequencies in comparison with the experimental data. This tendency remains when spatial grid resolution is decreased. Unfortunately, further experiments with mesh refinement are confronted with too high consumption of calculation time and computational resources.

The transient eddies dynamics can be examined on the animated sequence with vertical cross-section of the crucible. It becomes clear, that the melt flow loses any symmetry in phase of fully developed oscillations. Main vortexes of the mean flow break up into smaller ones, which move between different regions and therefore providing the intensive heat and mass exchange in the entire melt.

The discrete particle tracing approach has been carried out to investigate convective scalar transport mechanism in the considered flow. In the very beginning virtual particles are placed on the top of computational domain. These particles are assumed to have the same density as the fluid and this leads to the expectation that their path will coincide with the streamlines of the flow. When the



Picture 4 – Results of particle tracing a, b) in transient simulation and c) in long time period averaged flow

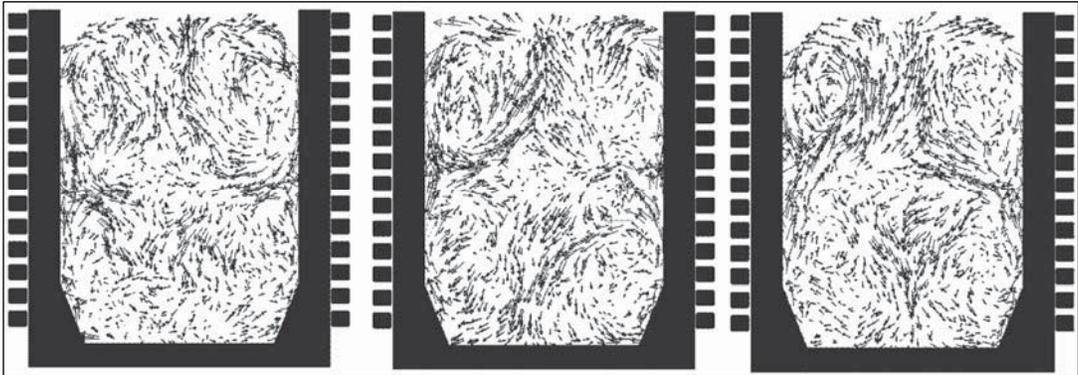
flow in closed domain without inlets and outlets is stationary, the streamlines are also closed and particle trajectories should be looped. In this case it is improbable that particle will penetrate into the neighbouring flow region if the turbulent transfer is neglected. Therefore, transport processes between the main flow eddies generally would have diffusive character in steady-state flow. In this case scalar exchange intensity will strongly depend on the semi-empiric turbulent parameters such as turbulent viscosity and turbulent Prandtl ($\delta_t = c_p v_t / \lambda_t$) or Schmidt ($Sc_t = \nu_t / \rho D_t$) number. Both latter parameters magnitude often depends on the type of fluid and flow properties and has to be determined experimentally. The flexibility of using such approach for different industrial installations is rather low.

The sub-grid viscosity models seem to have more universal character. Transient simulations with a small time step and appropriate meshing allow resolving the wide range of flow formations involved in scalar transport. Particle trajectories, traced in such instationary simulations, show that estimated convective mass exchange between the main flow eddies is quite intensive (Picture 5 a, b).

Four particles were launched simultaneously at $z = H$, $r = R/2$, $\delta\phi = \pi/2$ and traced for 20 seconds (about 5-6 eddy turnover times). The choice of starting position was caused by conformity with industrial alloying process, when additional components are added on the melt surface. The typical tracing result states, that most particles don't stay in one eddy longer than two or three turnover times. Hence, it is possible to come to the conclusion that convective transport mechanism plays significant role in the heat and mass exchange between the main flow eddies. The same tracing procedure was used with the time-averaged velocity field from transient LES calculations. As it was expected, all particles rotated in the initial eddy with relatively small azimuthally drift (Picture 5c). Probably, if longer time for averaging would be taken the trajectories would converge to those in steady-state flow pattern.

3 Industrial Applications

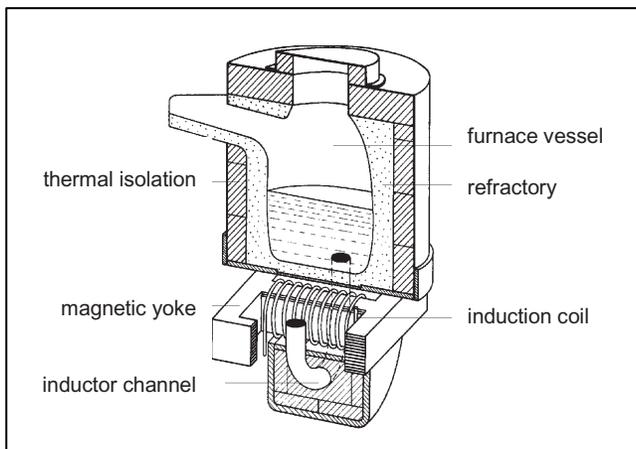
The simulation of the turbulent melt flow in an industrial crucible furnace is presented here as the first example of our LES numerical investigations. This furnace has a melt volume of about 0.9 m³ at 100% filling level. The radius of the crucible furnace is about 0.49 m and the height of the inductor is 1.34 m. The furnace is used for melting grey cast iron, which has a density of 6,800 kg/m³. The three-dimensional hydrodynamic model consisted of about 7·10⁵ elements and the time step in the transient calculations was 10⁻² s. Industrial crucible furnace differs from the experimental installa-



Picture 5 – Velocity distribution [m/s] in the industrial furnace calculated three dimensional instationary with LES after 20 (left), 21 (centre) and 22 seconds

tion with significantly larger linear dimensions, higher EM forces density and noticeable free surface deformation (meniscus).

The comparative analysis of LES and experimental data from model furnace allows applying this numerical method also for industrial scale installations, and qualitatively similar phenomena is achieved. The period of the low-frequency oscillations become smaller – 2 seconds – because of significant increase of the rotational velocity of the flow eddies. Initially axial symmetrical flow pattern becomes fully three-dimensional, but the symmetry remains in the time-averaging of the fluctuating flow.



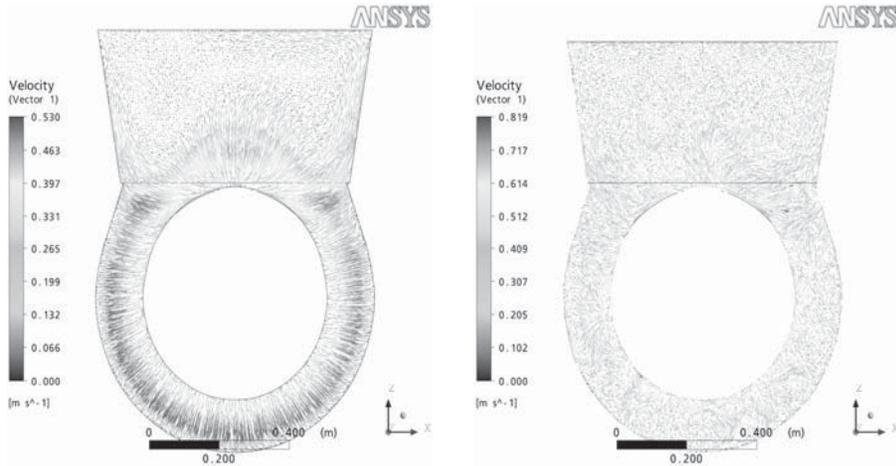
Picture 6 – Principle design of an one loop induction channel furnace [6]

The second example presented here is devoted to the induction channel furnace (ICF), which is used for holding and casting of ferrous and non-ferrous metals and due to its good efficiency for melting of non-ferrous metals. Picture 6 shows the principle design of a one loop ICF, which is typically used for holding and casting of grey cast iron. The ICF basically consist of a ceramic lined furnace vessel and one or several inductors. In principle, the inductor can be regarded as a transformer with iron yoke, where the induction coil is the primary circuit and the melt filled inductor channel represents the secondary short-circuited loop. For the safety and efficient operation of the ICF the heat transport from the channel, where the Joule heat is generated, to the melt bath in the furnace vessel is important in order to avoid a local overheating in the channel.

The melt flow in the channel itself and in the transition zone between the channel and the bath, the so-called inductor-throat, is very complex, highly turbulent and influenced mainly by electromagnetic forces but additionally by buoyancy forces. In order to investigate the operation behaviour of the ICF heat and mass transfer processes in the melt have been analysed applying the Large Eddy Simulation (LES) approach. The simulation results are verified by already existing data of melt flow velocity and temperature measurements [7].

The simulation results presented here are carried out for an industrial sized experimental furnace running with Wood metal as a model melt. The results are obtained using the two parameter or LES model show highly turbulent 3D dynamic flow vortex structures, with flow velocities up to 70 cm/s. The steady state velocity distribution in the symmetry plane calculated with the $k-\epsilon$ model (Picture 7 left) shows, that the flow is directed radial outwards in the channel and there is up-flow just above the channel forming two vortex loops in the bath. These loops have complex structure and are closed on front wall of the bath. Instantaneous velocity pattern calculated with the LES model (Picture 7 right) shows highly turbulent structure of the flow. However even here the regions of more or less symmetric structures and constant direction flow can be observed.

In comparison to these high local melt flow velocities, which are driven by EM-forces, the integral transition flow through the channel, caused mainly by buoyancy forces, is very small in the range of 5 cm/s and therefore only secondarily responsible for the heat transfer processes from the channel into the bath



Picture 7 – Melt flow velocity distribution on symmetry plane of the ICF. Left: steady-state distribution with k-ε model, right: intermediate situation at 9 sec with LES-model

The velocity distribution in the cross-section inside the channel is represented by two-vortex structure which also is in accordance with Lorentz force distribution and experimental measurements. Characteristic velocity magnitudes in the channel are approximately the same for both simulation models (0.5 – 0.7 m/s in maximum). The distance between the vortex centres is about channel radius.

Results of the simulation of the temperature distribution in the symmetry plane of the channel furnace are shown in Picture 8. The initial state with the channel maximal temperature in the bottom point is unstable and small velocity fluctuations can shift it to the left hand or right hand side on the channel, where it becomes stable corresponding to the transit convectional flow. Our still running investigations have shown, that the position of the magnetic yoke, which leads to unsymmetrical electromagnetic field and Joule heat distribution in the channel, influences the shift of the temperature maximum to the left hand or right hand side. Due to the dynamic interaction of the turbulent vortices the heat and mass exchange along the channel is well developed. The dynamic flow structure in the throat of the channel inductor leads to a smooth temperature gradient in the transition zone between the channel and the bath, which is shown as a result of the LES model (Picture 8 middle). But the results of the steady state temperature distribution, calculated with the k-ε model (Picture 8 right) show a precipitous temperature gradient at the exit of channel, because the interactions between the flow vortices in this region, which are characterized by low frequency oscillations, are not taken into account according to the real situation with the two parameter turbulence models.



Picture 8 – Temperature distribution in the channel furnace. Left: intermediate distribution after 45 sec (LES), middle: time averaged over 60 sec (LES), right: steady-state (k-ε model)

Conclusion

By Intensive numerical studies concentrated on applying the Large Eddy Simulation (LES) model for recirculating melt flows in induction furnaces were done together with experimental investigations in a model induction crucible furnace. The comparative analysis shows a good coincidence between the numerical and experimental results not only in terms of mean flow, but also for the turbulent fluctuations and their kinetic energy. The studies reveal that the low-frequency velocity oscillations play a main role in convective heat and mass transfer when flow structure contains two or more large vortexes of the mean flow. The modelling results show, that only the 3D transient LES is able to model correctly the heat and mass transfer processes in these recirculating flows. Therefore, the correct estimation of the characteristic parameters of these oscillations applying the LES method leads to creation of a universal and reliable numerical method, which can be used for solving fluid dynamics and thermal problems in practical metallurgical applications.

Literature

- [1] B.G. Thomas, Q. Yuan, S. Sivaramakrishnan, T. Shi, S.P. Vanka, and M.B. Assar: Iron Steel Inst. Jpn. Int., 2001, vol. 41, pp.1266-1276.
- [2] V. Bojarevics, G. Djambazov, R.A. Harding, K. Pericleous, M. Wickins: Magnetohydrodynamics, 2003, no. 4. pp. 395 - 402.
- [3] R. Schwarze, F. Obermeier: Modelling and simulation in materials science and engineering, 2004, vol. 12, pp. 985 - 993.
- [4] E. Baake, B. Nacke, A. Jakovics, A. Umbrashko: Magnetohydrodynamics, 2001, vol. 37, no. 1-2, pp. 13-22.
- [5] T. Weissenfluh: Int. J. Heat Mass Transf., 1985, vol. 28, no 8, pp. 1563-1574
- [6] Fasholz, J.; Orth. G.: Induktive Erwärmung. Verfahrensinformation, RWE-Energie, Essen 1991
- [7] Eggers, A.: Untersuchungen der Schmelzenströmung und des Wärmetransports im Induktions-Rinnenofen. Fortschritt-Berichte VDI, Reihe 19, Nr. 63, Düsseldorf 1993

ALUMINIUM BILLETS HEATED BY ROTATING MAGNETS: COMPARISON OF ANALYTICAL AND NUMERICAL RESULTS

F. Dughiero, M. Forzan, S. Lupi, F. Nicoletti

University of Padua, Department of Electrical Engineering,
35131 Padova, Via Gradenigo, 6/A., Italy

Key words: metal billet, heating, magnetic field, rotation, efficiency

Abstract

Металлические заготовки с низким электрическим сопротивлением могут быть нагреты с помощью токов, индуцируемых при вращении самой заготовки в поперечном магнитном поле постоянного тока. Из-за изменения магнитного потока, вращение заготовки влияет на распределение индуцируемых токов, которые генерируют источники тепла внутри заготовки.

Данный подход был подробно изучен в последние годы [1-3] поскольку он позволяет получить очень высокий КПД, практически соответствующий КПД электродвигателя, который обеспечивает механическое вращение заготовки. Такое техническое решение было успешно применено, и несколько установок работают в промышленности на алюминиевых литейных заводах. Главный недостаток описанной технологии связан с высокой стоимостью установки, поскольку индукционное поле постоянного тока создается катушкой из сверхпроводящих материалов.

Рассматриваемое в данной статье решение позволяет получить такой же высокий КПД при меньшей стоимости установки. Предлагается оставить заготовку неподвижной, а вращать постоянный магнит, расположенный внутри ферромагнитной рамки.

В статье результаты аналитического метода определения распределения мощности при предложенной конфигурации нагревательной системы сравниваются с результатами, полученными с использованием коммерческого пакета FEM с целью оценки аналитического подхода.

Description of the problem

The proposed system is presented in Figure 1: the first figure describes $\frac{1}{4}$ of the entire geometry for a 4 poles system while the second one shows the entire geometry of a 8 poles system.

FEM commercial codes are available since many years and they have proven to be reliable for the solution of eddy current problems. On the other hand, the electromagnetic problem can be solved only by means of a transient solution to take into account the motional effects. This solution reaches the steady state condition only when a high number of time steps have been solved and, as a consequence, the computational times are very long.

The analytical solution requires very short computational time, also because it gives directly the steady state condition of the system and, for this reason, it can be conveniently applied to an automatic design process.

On the contrary, the analytical code has been developed because several parameters of the system need to be optimized, in particular the permanent magnets represent the most expensive part and, consequently, the number of the magnetic poles, the type and the thickness of the permanent magnets have to be carefully chosen.

Figure 2 shows how the ratio between the induced power and the weight of the permanent magnets, computed from a parametrical analysis carried out by using FEM commercial code, varies with the magnet thickness.

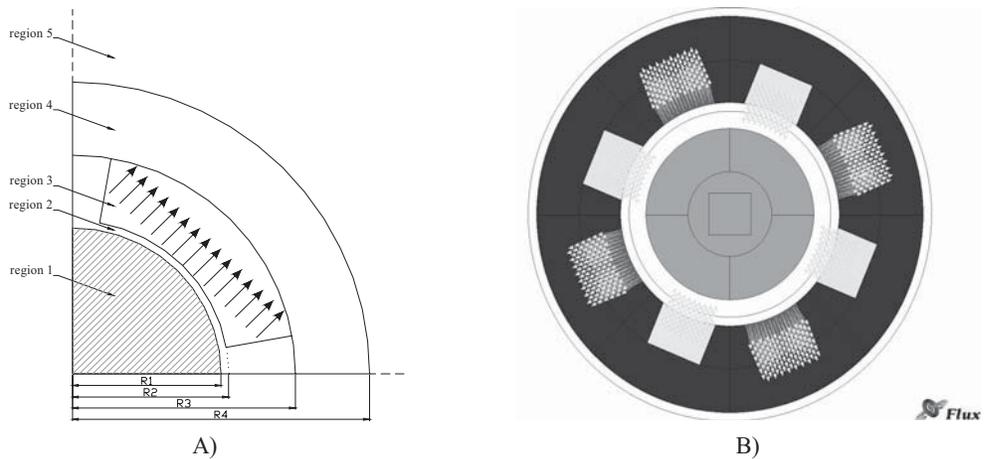


Figure 1 – System geometry. [A] 4 poles system. Region 1 – aluminum billet; region 2 – air-gap (which takes into account also the thermal insulating material to protect permanent magnets); region 3 – permanent magnets; region 4 – iron frame; region 5 – external air. B) 8 poles model]

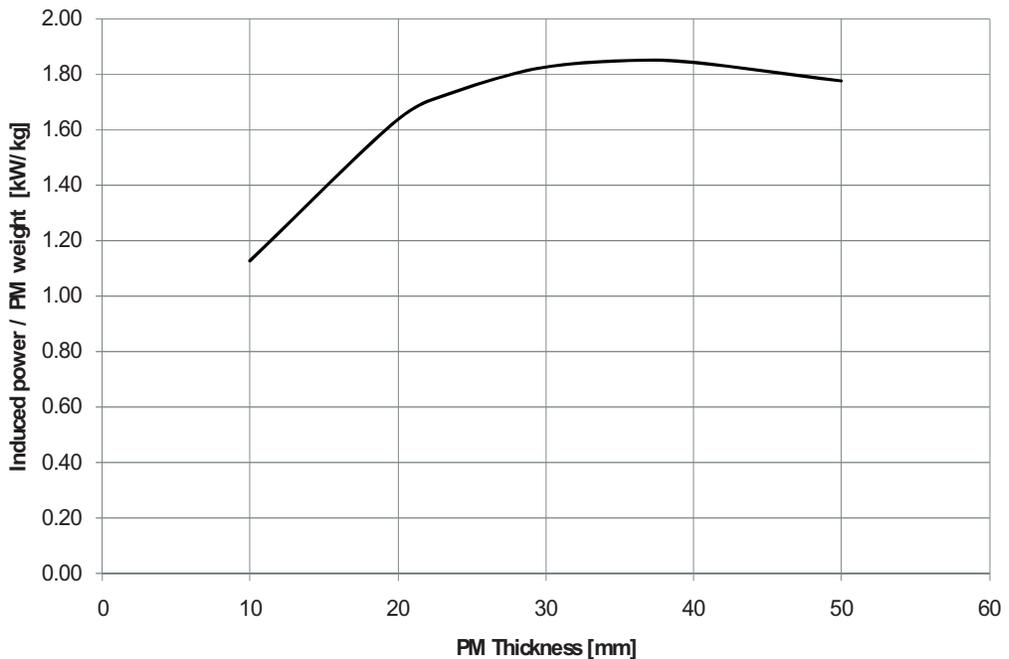


Figure 2 – Example of optimal design: ratio between induced power in the billet and weight of the permanent magnets, as a function of the permanent magnets thickness

Simulation strategy

FEM Solution

The FEM model has been solved taking into account the rotational movement of the permanent magnets by means of a transient magnetic solution.

The 2D electromagnetic formulation is based on the following PDE:

$$(1) \nabla \times \left(\frac{1}{\mu_r \mu_0} \nabla \times \mathbf{A} - H_c \right) + \sigma \left(\frac{\partial \mathbf{A}}{\partial t} + \nabla V \right) = 0$$

where μ_r – the relative permeability of the medium, \mathbf{A} – the vector potential [Wb/m], H_c – the coercive magnetic field for the permanent magnets, σ – the conductivity of the medium [S], V – the electric scalar potential.

The constitutive equations for magnetic materials are described by:

$$(2) \mathbf{B} = \mu \mathbf{H} \Rightarrow \mu_r \mu_0 \mathbf{H} \quad \text{for soft magnetic materials, e.g. region 4 of Figure 1}$$

$$(3) \mathbf{B} = \mu_r \mu_0 \mathbf{H} + \mathbf{B}_r = \mu_r \mu_0 \mathbf{H} + \mu_0 \mathbf{M} \quad \text{for hard magnetic materials, e.g. region 3,}$$

where B_r is the remanence and \mathbf{M} the magnetization vector.

The displacement of the rotating part modifies the geometry of the modelled device; consequently the computation domain must be re-meshed at each time step.

The state variable, the magnetic vector potential \mathbf{A} and the electric scalar potential V are initialized to zero for $t = 0$.

The state variables are time dependent and the differential equations of first order with respect to time are integrated by means of the Euler method in time domain. The time domain is sub-divided into small time steps; during each step the unknowns are supposed to vary linearly with respect to time. The numerical code applies then an implicit method where the derivative is computed by means of the current and the previous step values:

$$(4) \frac{dX}{dt} = \frac{X_{i+1} - X_i}{\Delta t} .$$

For the 8 poles case, some results are presented in Figure 3, where the colour maps describe the induced current density in the aluminium billet and the lines are the flux lines.

Analytical solution

The analytical solution of the problem has been developed using a cylindrical coordinates system (r, ϑ, z) with the origin in the center of the billet. In the following formulas, the indexes of the regions makes reference to the description of Figure 1.

The applied analytical method is completely described in [11].

The computation of the magnetization \mathbf{M} has been carried out in the permanent magnets (region 3) in the hypothesis of a 2D field that can be described in polar coordinates as:

$$(5) \vec{\mathbf{M}} = M_r \vec{\mathbf{u}}_r + M_\vartheta \vec{\mathbf{u}}_\vartheta$$

by means of Fourier series:

$$(6) M_r = \sum_{n=1,3,5,\dots}^{\infty} M_{rn} \cos(np\vartheta) \quad M_\vartheta = \sum_{n=1,3,5,\dots}^{\infty} M_{\vartheta n} \sin(np\vartheta) ,$$

where $2p$ is the number of poles.

The two components of the magnetization are illustrated in figures 4 and 5, for a 8 poles system.

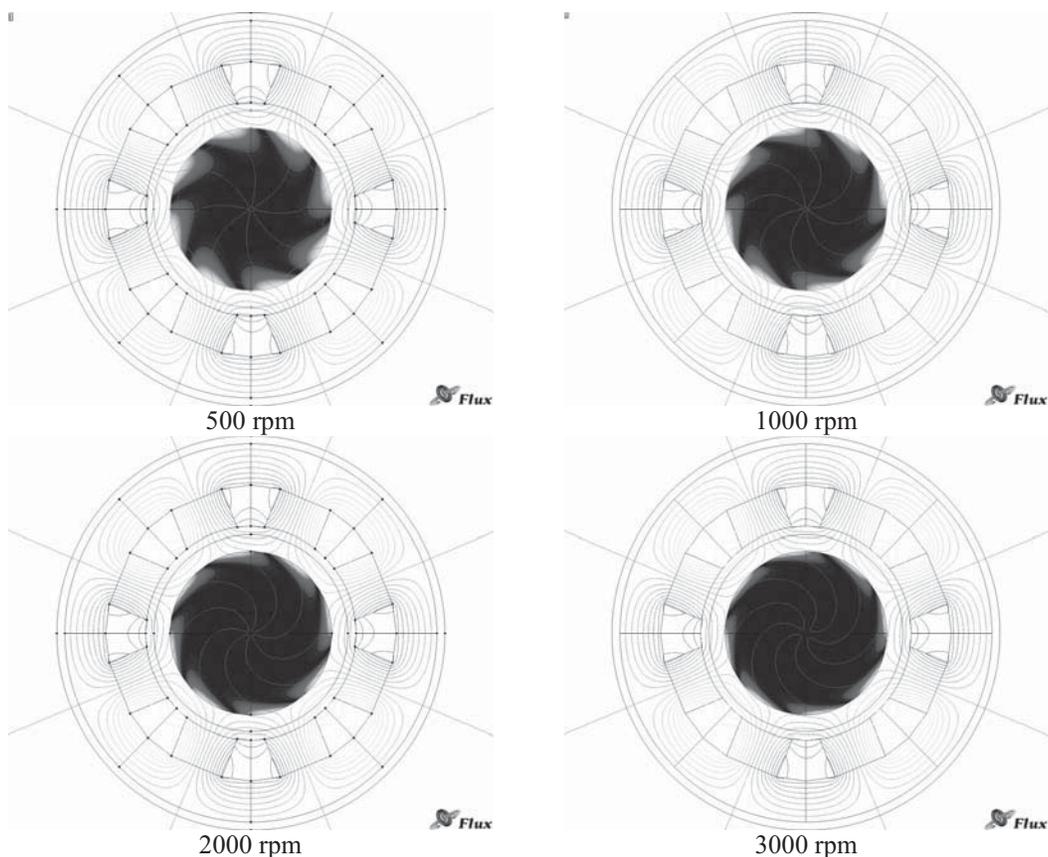


Figure 3 – Induced current density distribution (color shade plot) and equipflux lines resulting in the steady state condition of a transient FEM solution at different rotational speeds

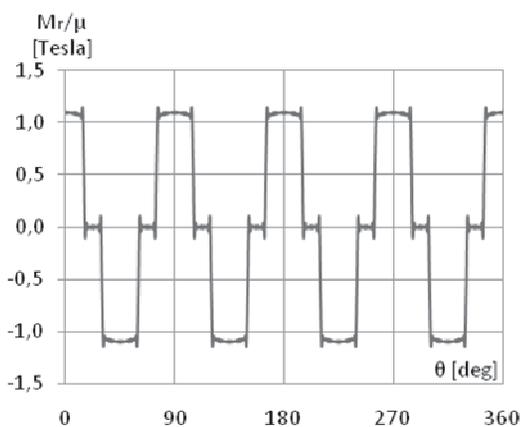


Figure 4 – Magnetization M_r

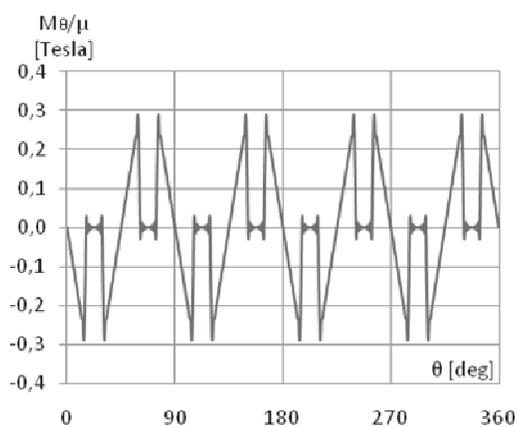


Figure 5 – Magnetization M_θ

The general solution of the problem can be obtained by the diffusion equation in terms of the magnetic vector potential that has only one component along z direction:

$$(7) \quad \frac{\partial^2 A}{\partial r^2} + \frac{1}{r} \frac{\partial A}{\partial r} + \frac{1}{r^2} \frac{\partial^2 A}{\partial \theta^2} = \sigma \mu \Omega \frac{\partial A}{\partial \theta} + \frac{1}{r} \frac{\partial M_r}{\partial \theta} - \frac{1}{r} M_\theta - \frac{\partial M_\theta}{\partial r}.$$

The induced current density in the complex domain can be computed from the values of A:

$$(8) \quad \underline{I}_1 = -j \frac{k^2}{\mu_1} A_1 .$$

The instantaneous value of the power losses in the billet, integrated over the volume of the cylinder, is computed by integrating the specific power density:

$$(9) \quad P = \int_0^{R_1} \int_0^{2\pi} \frac{J_1^2(r, \vartheta)}{\sigma} r d\vartheta dr .$$

RESULTS

Some comparisons between the results obtained by the analytical formulation and the ones resulting from the FEM solution are presented in the following for 3 different cases (4 poles, 8 poles, 16 poles) and different rotational velocities. The main geometric dimensions of the system, taken into account as reference for this comparison with reference to Figure 1, are: R1 = 100 mm; R2 = 130 mm; R3 = 180 mm ; R4 = 230 mm; α_p , the magnet pole-arc to pole-pitch ratio = 0.666; the axial length = 1 m.

The rare earth permanent magnets considered for the calculations have a remanence $B_r = 1.1$ T and a coercive field $H_c = 838$ kA/m.

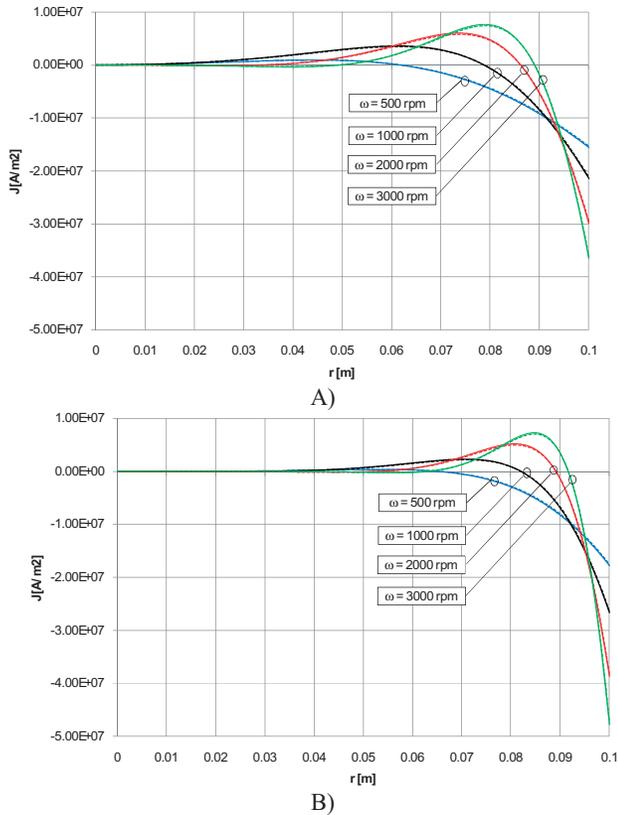


Figure 6 – Current density distribution evaluated along the radius for different rotational speeds
A) 4 poles system B) 8 poles system. FEM solution - dotted lines; analytical solution - continuous lines

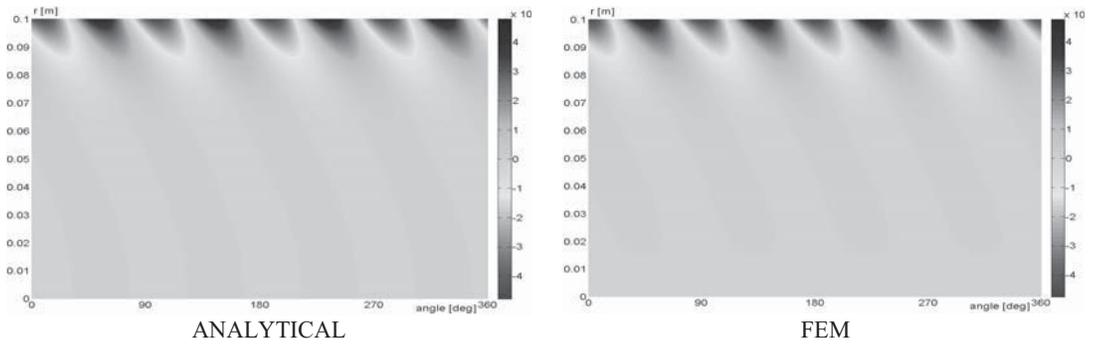


Figure 7 – Current density distribution as resulting from the analytical solution and the FEM one for 8 poles system: on x axis the azimuth [angle in degree], on y axis the radial position [m].

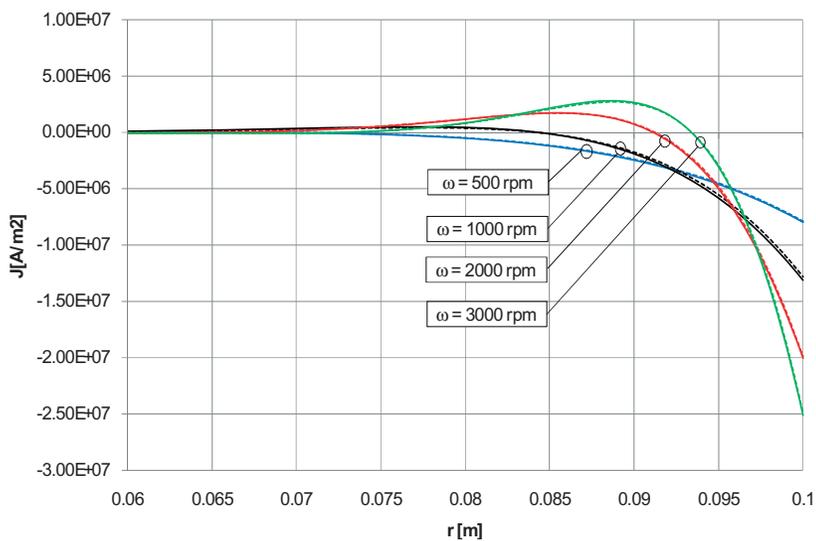


Figure 8 – Current density distribution evaluated along the radius for different rotational speeds for 16 poles system. FEM solution - dotted lines; analytical solution - continuous lines

Table 1 – Induced power in the billet computed by the FEM and the analytical (AN) method for the different configurations and rotational speeds

	p = 4		p = 8		p = 16	
	FEM P [W]	AN P [W]	FEM P [W]	AN P [W]	FEM P [W]	AN P [W]
500	9.95 E+04	1.01 E+05	5.82 E+04	5.93 E+04	6.20 E+03	6.31 E+03
1000	1.67 E+05	1.69 E+05	1.19 E+05	1.21 E+05	1.58 E+04	1.61 E+04
2000	2.61 E+05	2.65 E+05	2.08 E+05	2.12 E+05	3.21 E+04	3.28 E+04
3000	3.33 E+05	3.38 E+05	2.78 E+05	2.84 E+05	4.54 E+04	4.65 E+04

Conclusions

In the paper, a comparison between numerical and analytical methods for the prediction of current and power density induced in a conductive billet inside a rotating magnetic field is proposed. The two different methods can be used in order to design and optimize this innovative technique to heat billets.

The proposed method looks very attractive because it overcomes the main drawback of the previously proposed approach where the billet is forced to rotate inside a DC induction field produced by superconductive coils, that requires an adequate refrigerating system. The system reaches the same efficiency values of the previously proposed approach, efficiency that mainly depends upon the efficiency of the motor and its driving system, achieving strong induced power values using a set of rare earth permanent magnets instead of a more expensive superconductive system.

Literature

- [1] F. Dughiero, M. Forzan and S. Lupi, (2006). Induction heating of aluminium billets rotating in a DC magnetic field. Proc. of VIII Int. Conf. on Problems of Control and Modeling Complex Systems, Samara - Russia, June 24-29, pp. 171-176.
- [2] Araneo R., Dughiero F., Fabbri M., Forzan M., Geri A., Lupi S., Morandi A., Ribani P., Veca G. (2008). Electromagnetic and thermal analysis of the induction heating of aluminum billets rotating in dc magnetic field. HES-07 – Heating by Electromagnetic Sources. June 19-22, 2007, vol. 1, pp. 487-496. ISBN 88-89884-07-x, Padua: Sgeditoriali (Italy); printed in COMPEL, vol.27, No.2, 467-479, ISSN 0332-1649.
- [3] Fabbri M., Forzan M., Lupi S., Morandi A., Ribani P.L. (2009). Experimental and Numerical Analysis of DC Induction Heating of Aluminium Billets. IEEE Trans. on MAGNETICS. Vol.45, n.1, January 2009, 192-200.
- [4] Watanabe T., Todaka T., and Enokizono M. (2005). Analysis of a New Induction Heating Device by Using Permanent Magnets. IEEE Trans on MAGNETICS. Vol. 41, NO. 5, May 2005, 1884-1887.
- [5] Zhu Z. Q., Howe D., and Chan C. C. (2002). Improved analytical model for predicting the magnetic field distribution in brushless permanent-magnet machines, IEEE Trans. on Magnetics, vol. 38, no. 1, pp. 229-238.
- [6] Lubin T., Netter D., Leveque J., and Rezzoug A. (2009). Induction Heating of Aluminum Billets Subjected to a Strong Rotating Magnetic Field Produced by Superconducting Windings;. IEEE Trans. on MAGNETICS. Vol.45, n.5, May 2009, 2118-2127.
- [7] Boughrara K., Chikouche B. L., Ibtouen R., Zarko D., and Touhami O.(2009). Analytical Model of Slotted Air-Gap Surface Mounted Permanent-Magnet Synchronous Motor With Magnet Bars Magnetized in the Shifting Direction. IEEE Trans. on MAGNETICS. Vol.45, n.2, February 2009, 747-758.
- [8] Markovic M., and Perriard Y.(2007). An analytical determination of eddy-current losses in a configuration with a rotating permanent magnet. IEEE Trans. on MAGNETICS. Vol.43, n.8, August 2007, 3380–3386.
- [9] Markovic M., and Perriard Y.(2008). Analytical solution for rotor eddy-current losses in a slotless permanent magnet motor: the case of current sheet excitation. IEEE Trans. on MAGNETICS. Vol.44, n.3, March 2008, 386–393.
- [10] CEDRAT, <http://www.cedrat.com/>, FLUX users guide.
- [11] Dughiero F., Forzan M., Lupi S., Nicoletti F., Zerbetto M.(2010). A new high efficiency technology for the induction heating of non magnetic billets. HES-10 – Heating by Electromagnetic Sources. May 19-21, 2010.

THE REVIEW OF COMMON ACTIVITY OF ETP AND SAMSTU IN THE FIELD OF SIMULATION AND OPTIMISATION OF ELECTROTECHNOLOGICAL PROCESSES

B. Nacke¹, Yu.E. Pleshivtseva²

¹Leibniz University of Hannover, Institute of Electrotechnology
30167 Hannover, Wilhelm-Busch-Str. 4, Germany
nacke@etp.uni-hannover.de
phone: +49 (0) 511-762-5533, fax: +49 (0) 511-762-3275

²Samara State Technical University,
443100, Samara, Molodogvardeyskaya Str., 244
yulia_pl@mail.ru
phone: +7 (846) 332-42-34, fax: +7 (846) 278-44-40

Key words: *simulation, numerical models, optimal control, algorithm, optimisation procedure, hot forming, induction heating, design, forging, production line.*

Abstract

В статье представлены основные направления работ, проводимых совместно Институтом Электротехнологий Университета им. Лейбница (г. Ганновер, Германия) и Самарским государственным техническим университетом (Россия) в рамках многолетней программы сотрудничества в области моделирования и оптимизации электротехнологических процессов. Рассмотрены основные образовательные и научно-исследовательские проекты, реализуемые в течение последних 12 лет.

Introduction

The role of international co-operation in scientific researches and education increases constantly worldwide. The integration of scientists from German and Russian universities is seen today as necessary condition for the advancement of science and education at the national level in Germany and Russia.

The Institute of Electrotechnology (ETP) at the Leibniz University of Hannover (LUH) is the only university institution in Germany which covers the full range of all fields in Electrotechnology in teaching and research. The institute has a lot of relations to industrial partners and nearly all scientific projects carried out in the institute are in co-operation with industry. It is one of the worldwide leading research institutions for induction heating and melting and has long-term co-operations of more than 15 years with signed co-operation agreements with many western and Russian universities.

At Samara State Technical University (SamSTU) investigations have been conducted in the fields of mathematical modelling and optimal control of interconnected electromagnetic, thermal, and structural mechanics processes for more than 40 years. During this period, the theoretical basis has been developed for methods of optimisation of technical processes with respect to different technical and economic cost functions. The Alternance Method is developed at SamSTU as the basis for the solution of optimal control problems [1]. Special techniques for computations of optimal control algorithms have been created for thermal physics systems with distributed parameters.

The successful and fruitful co-operation between the ETP and SamSTU has started already more than 12 years ago. Many scientific projects have been and are carried out during this time. The content of this cooperation is fixed in the programs of the mutual work. In the frame of this cooperation some educational projects have been successfully carried out as well, for example, TEMPUS/TACIS European Compact Project CP_20021-98 "Dissemination of knowledge on modern energy saving electrotechnologies" in 1999-2001.

The goal of this paper is to describe the main ideas and directions of recent educational and research activities carried out in the frame of the close co-operation between ETP and SamSTU.

1 Background of co-operation

The use of electrical energy is the basis of modern industrial development of a country and plays an even increasing role in people welfare. Due to the amount of energy required and the costs involved, even small increases in efficiency and small percentual economic savings in the application of traditional and innovative energy saving electrotechnologies represents large amounts of money. These economic savings can be accomplished by the optimisation of electrotechnological processes both at planning and operational level. Only through a quantified, well balanced analysis, development, simulation, optimisation, and automatic control of technological processes, and a constant awareness of the economical factors that they involve, the industry can stay competitive.

The main goal of presented scientific co-operation consists of combination special optimization approaches with advanced mathematical models and application of the developed optimization strategy to the optimal design and control of different induction heating systems and manufacturing lines.

To create competitive technology it is necessary to integrate mathematical models with such optimization methods that would be able to give an alternating technical solutions, which could not be improved regarding to chosen optimization criteria. In this case it is necessary to develop special optimization method, which combines problem-oriented process model with appropriate optimal control tool. These purposes could be reached due to close cooperation of high competitive specialists in these fields from ETP and SamSTU.

Improvement of diploma and post graduate students' education and their involvement into R&D Projects should help to decide the problems and to prepare high level specialists for industry and universities. It is also planned to develop and implement different courses for students and PhD studies in the area of energy saving electrotechnologies in line to the Bologna principles that provides development of very close interconnection between scientific researches and educational programs in European and Russian universities, in particular, at LUH and SamSTU.

2 R&D Projects

The main goal of common multidisciplinary scientific research Projects is the development of new theoretical approaches, methods and engineering techniques for mathematical simulation, optimisation and automatic control of industrial high effective electrotechnologies. This theory is considered as a basis for development of problem-oriented mathematical models, optimisation procedures, automatic control systems providing optimal operational modes of energy saving electrotechnological processes increasing the efficiency of industrial technologies, improving the end product quality, and achieving technical process flexibility. Finally, this will lead to economic advantages regarding saving of energy, that is being of vital importance considering main aspects of industry development. Highly efficient technologies could be developed due to the combination of latest achievements of the Projects partners.

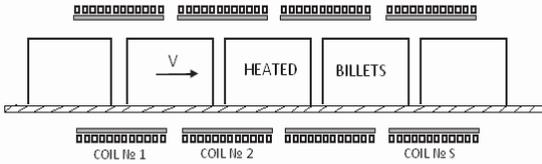
2.1 Optimal control of continuous induction heaters at steady-state and transient operation modes

The majority of induction billet heaters utilize heating mode where billets are moved through a heater. This Project provide a study of the wide range of optimal control problems for continuous induction heating modes.

With the continuous heating mode the workpiece is moved in a continuous motion through one or more in-line induction heating coils. This heating mode is commonly used when it is required to heat long components such as bars, slabs, strips, tubes, wires, blooms, and rods. Continuous heaters, being in common use, are considered at steady-state and transient operation modes. In the researches

the optimal control problems for continuous heating modes are based on optimal control method taking into consideration the most typical real-life operating requirements.

One of the main technological requirements to through heating is to provide an outlet temperature distribution along the billet's cross-section on the level of a required temperature before the consequent stages of the technological processes such as forging, rolling, forming, extrusion, etc.



Picture 1 – Induction through heater

That is why the first optimization goal is to minimize the deviation of final temperature along cross-section of the billet from the prescribed one during steady-state mode of heater operation.

In the most typical technological processes admissible deviation ε of temperature from the required outlet temperature is prescribed. It means that at the end

of the heater temperature in any point of the billet cross-section should deviate not more than the value ε from the required temperature θ . If the number of sections S and their sizes are defined by the given IHI design, the chosen voltages U_1, U_2, \dots, U_S (currents or appropriate heat powers) for all S coils of induction heater can be considered as control inputs. In this case the typical problem statement for induction through heating optimization can be formulated as a problem of determining optimal coil voltages (currents, powers) that provide the most uniform heating.

However in practice the steady-state modes can be very often violated by disturbances connected with changes of billets' sizes and properties; changes of heater throughput; technological breaks in heater operation; start and stop modes and so on. Under these conditions the main goal of control process is to minimize the deviation of outlet temperature distribution along the billet cross-section from the required one during transient modes.

So as steady-state or transient operation modes of induction heating installations should provide desired billets' temperature with minimal deviation during a fixed time. In such case the induction heating process can be considered as the control object and problem of desired temperature achievement can be formulated and solved as optimal control problem.

In the case of transient mode optimization for the chosen set of data it is necessary to find out the functional dependences of optimized parameters from time, which provide extreme value of one of the optimization criteria.

The optimal control problems have been solved for the number of industrial through heaters with 1, 3, 4 and 10 inductors [2, 3] operating in steady-state modes. Optimization results have been obtained for different variants of inductors distribution to the sections. It was proved that two sections under independent control are already enough to reach the temperature deviation ε on the level of 8-10°C. Further increasing the number of sections does not lead to a reduction of temperature deviation from required one at the heater outlet. The optimally controlled heaters have been investigated for different throughput of the steady-state process. Even when throughput decreases from 100% down to 30% the optimal control method allow to provide not significantly increasing of the maximum temperature deviation ε in all considered cases.

The optimal control problem have been also solved for the transient modes corresponding to initial start of the heater, restart after technological brakes of different durations, changing the work-piece specification or production rate of the heating installation [2, 3].

In this Project different 2D steady-state and transient numerical models of induction through heaters with continuous workpiece movement have been developed at the ETP and successfully used in optimisation procedures.

The investigations fulfilled lead to the following conclusions. A significant effect is achieved for single- and multi-sectional heaters both in steady state and transient modes. Comparing to direct search algorithms the optimal control techniques significantly reduce a number of necessary iterations. The new method could be successfully applied to improve heating processes with use of corresponding more complicated mathematical models. There are also no fundamental difficulties in use of this method for continuous induction heaters optimization regarding different optimization criteria, for example, energetic or economical ones.

2.2 Optimal design of induction heaters for forging

As applied to steady-state continuous heating there is a variety of design concepts of induction heater that provide required heat generation power distribution along a heater length.

Many variants of inductor designs are available but all of them have certain limitations. The list of appropriate design solutions often used in practice includes but not limited to:

- Multi-coil inductor design (number of in-line coils with different geometry and power);
- Proper choice of coil winding (taper winding);
- Applying special feed circuits;
- Different design approaches in regards to system “power supply-inductor-metal”;
- Proper choice of coil overhang.

One of the most typical implementation of space-time control is multi-coil design of continuous heater with independently controlled separate coils. In this case appropriate number of in-line coils is used. Each coil may have different diameter, length and number of turns. In addition, power of each coil varies within certain limits and can be different from coil-to-coil. Such approach provides desired change of heating power in time and along heater length. It is also one of the most widely used design solution to control transient heating modes of induction heater [1].

The Project deals with the optimal design problem for induction heaters for forging industry.

As a typical example the problem that minimizes heater length L_{\min} has been considered for the case when speed V of billet movement through an inductor heater is constant and a required heating accuracy should be obtained at the inductor exit. Then cost function represents total heater length L . Spatial distribution of heat generation power along the heater length has been chosen as a control input function. A distribution of internal heat sources can be varied by variation of currents (voltages) of power supplies applied to heater sections. That is why the power supply currents (voltages) can be chosen as an optimized parameters as well. It was obtained that it is possible to reduce the heater length by 5-12% depending on particular minimum length problem [3].

The optimal control method leads to good results in problems of optimal design of induction heating installations for forging industry. Obtained results can be extended onto other optimal design problems, they exhibit quantitative and qualitative properties of optimal heating processes and deal with estimation of a maximum of capability of induction heater. This can be very helpful in order to have a rough idea in regards to number of sections, power levels and required shop floor space.

2.3 Optimisation of Alu billet heating by rotation in DC magnetic field

For improving the energy efficiency of induction heating processes, an advanced induction heating technology has been developed last years (Fig. 2) [1-4]. It was suggested to use DC superconducting windings producing high intensive magnetic field. The technology is based on the ability of superconductors to carry DC currents loss-free. The heating is performed due to induced currents obtained by rotation of billets in DC magnetic field. The efficiency of such heating technology could theoretically approach to 90% as it fully utilizes the potential of the superconductor materials.

The main goal of the Project consists of combination special optimization approaches with advanced mathematical models and application of the developed optimization strategy to this innovative heating technology. The Project deals with the time-optimal control problem for induction heating of billets rotating in DC magnetic field. The temperature distribution within billet cross-section at

the end of heating could be treated as an output controlled function of the process. Variation of rotation frequency over time $f(\tau)$ is chosen from the set of admissible controls to influence temperature distribution and dynamic behavior of induction heating system.

In order to investigate heating process the numerical 2D ANSYS model has been developed at ETP that allows to simulate the DC magnetic field, the rotating billet inside the field and the generation of eddy currents inside the aluminum billets. The results of modeling represent electric current, heat density and temperature distribution within the billet cross-section. The simulation results are in good agreement with analytical estimations and the data based on experiments.

When a maximum productivity is required, a minimal total heating time τ^0 could be considered as a cost function and time-optimal control problem should be formulated. Metal heating processes impose the two most general requirements with regard to the temperature distribution during the heating process. The first requirement demands that the maximum value of the temperature within the heated billet should be below a certain admissible value θ_{adm} . If this value will be exceeded, then irreversible adverse changes in material structural properties (i.e., grain boundary liquation, burning, etc.) and even metal melting could take place. Secondly, the temperature differences within the whole volume of heated workpiece should be restricted during heating in such a way that the maximum value, σ_{max} , of tensile thermal stresses (due to thermal gradients) would not exceed prescribed admissible value, σ_{adm} , that corresponds to ultimate stress limit of the heated material.

In this case the optimal control problem should be formulated as follows. It is required to select such variation of rotation frequency over time that provides steering billets' initial temperature distribution to desired temperature with maximum heating accuracy and holds maximum temperature during heating at maximum admissible level. Additional technological constraints add complexity to optimal control problem solution and appropriate computational procedures [4].

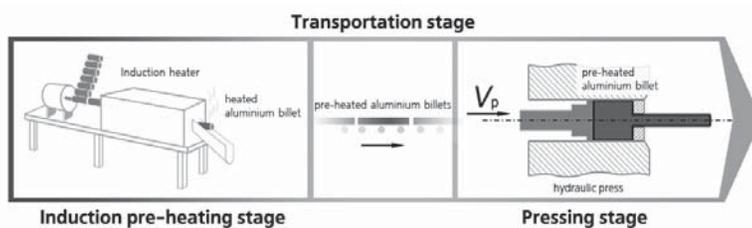
The time-optimal control algorithms with and without taking into account the main technological restrictions on maximum admissible temperature and maximum admissible thermal stresses, and problem-oriented optimization software have been developed in application to this innovative technology. Optimization procedures are based on alternance method of optimization of systems with distributed parameters.

From the investigations performed it is possible to conclude that only two-stage heating could provide required heating accuracy at minimum heating time. Even when technological constraints are taken into account, at the end of heating process it is possible to obtain the temperature non-uniformity along billet cross-section on the level of 5°C that is usually acceptable for the majority of hot forming applications.

2.4 Optimisation of manufacturing lines for metal hot forming with induction pre-heating

The main goal of the project is the development of problem-oriented methods, algorithms and mathematical models for optimal control of interrelated electromagnetic, temperature and structural deformation fields in technological lines for metal hot forming with induction pre-heating in order to achieve best major technical and economic quality criteria.

The system optimisation approach allows to take into consideration the sequence of the interrelated technological stages, including billet pre-heating by induction, its cooling during transportation, and subsequent plastic deformation during hot working, as a whole technological line being controlled. Combined optimisation criteria can be used for optimisation all these subsequent processes as a whole. At the same time the local cost functions for optimisation of induction heating and pressing stages could be proper selected based upon valid decomposition of a combined optimisation problem into separate optimisation problems for processes of heating and hot working [1].



Picture 2 – Structure of the technological line for metal pressing with induction pre-heating

In order to describe physical processes of production line, the system of interconnected numerical non-linear models are developed at ETP and SamSTU. Since the on-line optimisation procedures are planned to be used, some assumptions and simplifications are applied in the process of models development to reduce their time-consuming. The new approach to numerical simulation of pressing stage which is in fairly correspondence with known physical processes regulations was developed during Project running. This axis-symmetrical two-dimensional model allows to simulate numerically the temperature and stress fields evolving over pressing of aluminium alloy cylindrical billet. There are big potentials in application of this simulation method for optimization of hot forming stages in manufacturing line “induction heater – hot working equipment”. So the subsequent work should be aimed to verification and experimental model validation and adjustments of the model to be used in optimisation procedures. At the same time, the special optimisation algorithm that allows dramatically reduce the number of necessary iterations and models calls is developed to be applied in the optimisation procedures.

A significant effect is expected to be achieved both for local and combined optimisation of steady state operation modes of technological line for metal hot forming with induction pre-heating.

In future it is planned to develop an applied theory of mathematical simulation and optimal control of interrelated electromagnetic, thermal, and structural mechanics processes. This theory shall serve as a basis for increasing the efficiency of industrial technological lines for metal hot forming with pre-heating by induction. Problem-oriented mathematical models and innovative methods will be developed to define optimal control algorithms and to create control systems providing optimal operational modes for the whole technological line as well as for its separate installations with respect to the main joint technical and economic criteria. A system of problem-oriented mathematical models of different complexity levels that describe interrelated electromagnetic fields, temperature fields, and structural deformation will be developed to carry out the optimisation procedures [5].

It is necessary to underline that the described approach leads to determining capabilities of a control algorithm out of the framework of traditional “servicing of” technology and ensures active participation in technology formation.

3 Educational activities

During the time of co-operation between ETP and SamSTU many intensive exchanges of students, post graduate students, young scientists and Professors from both sides have been done.

In the frame of the “Mikhail Lomonosov” and “Leonhard Euler” Programmes the German Academic Exchange Service (DAAD) helps to support and to intensify internships of young Russian student, PhD-students as well experienced post-doc researchers. Up to now 10 students and PhD-students have received an important financial support in Russia and the unique possibility to stay and work in. The results of students and PhD-students internships have indicated that they contribute significantly to the better mutual understanding and to faster achievement of educational and scientific objectives in both countries.

At present time, the ETP and Sam STU plan to participate in new educational Project. The wide objective of this Project is the development and implementation of PhD modular curricula in the field

of Electrotechnology in RU. The Russian Federation has signed the Bologna declaration. Therefore within next years all Russian universities have to reform the PhD curricula according to the Bologna process. Several Russian universities offer a PhD education in Electrotechnology, but actually specific courses in Electrotechnology are not available for PhD students at RU universities. The planned courses, which will be taught in English, are based on the most recent theories on numerical simulation, control and optimisation in the field of electrical engineering, enabling PhD- students to acquire an in-depth understanding of electrotechnologies development, to develop analytical skills and methods in applied researches and to put theoretical results into industrial application in real life. The methodology used is interactive and task-based giving students the opportunity to use their knowledge in practical situations. A one- and two-month internships will allow participants to use their newly acquired skills and knowledge in partner universities activities.

4 Conclusions

It is necessary to underline, that most of the Projects mentioned above are not at the finishing projects status. Described optimization strategy is under development and could be improved, adjusted and applied for similar problems and for problems that remain beyond the bounds of the Projects carried out. So the subsequent work should be concentrated on expanding of this method applicability sphere.

Presented results of optimization illustrate the big potential of the optimisation strategy developed. The results obtained will allow for improving the end product quality, increasing the output rate, saving floor space, decreasing idle time of deforming equipment, and achieving technical process flexibility. Finally, this will lead to economic advantages regarding production costs and saving of energy, the latter being also of vital importance considering environmental aspects. Highly efficient technologies will be developed due to the combination of latest achievements of the project partners in the areas of mathematical simulation and optimal control theory.

Literature

- [1] E. Rapoport, Yu. Pleshivtseva. Optimal Control of Induction Heating Processes. DK6039, CRC Press/Taylor & Francis Group, 6000 Broken Sound Parkway, NW Suite, 300. Boca Raton, FL 33487 (USA), 2007, 349 p.
- [2] Pleshivtseva Yu., Nacke B. Optimal control of induction heaters for forging. Heat Processing (7) Issue 1, 2009, pp. 57-62, ISSN 1611-616X, Vulkan-Verlag GmbH, Essen (Germany).
- [3] Pleshivtseva Yu., Rapoport E., Efimov A., Nacke B., Nikanorov. Optimal design and control of induction heaters for forging industry. Proc.: International Seminar "Heating by Electromagnetic Sources" HES-07, Padua, Italy, June 19-22, 2007. – pp.251-258.
- [4] Pleshivtseva Yu., Zaikina N., Nacke B., Nikanorov. Time-optimal control of energy-efficient heating of aluminum billets rotating in DC magnetic field. Przegląd Elektrotechniczny (Electrical Review), ISSN 0033-2097, R. 84 NR 11/2008, pp. 120-123.
- [5] Pleshivtseva Yu., Rapoport E., Afinogentov A., Yu. Shemyakin, Nacke B., Nikanorov A. Application of optimal control theory for optimisation of metal hot forming lines with induction pre-heating, Proc.: International Seminar "Heating by Electromagnetic Sources" HES-10, Padua, Italy, May 19-21, 2010.

OPTIMISATION OF PRODUCTION COMPLEX FOR METAL PRESSING WITH INDUCTION PRE-HEATING

Yu.E. Pleshivtseva

Samara State Technical University
443100, Samara, Molodogvardeyskaya Str., 244, Russia
yulia_pl@mail.ru
phone: +7 (846) 332-42-34, fax: +7 (846) 278-44-40

Key words: *production complex, optimisation, control algorithm, hot forming, induction heating, numerical models, local and joint optimisation criteria*

Abstract

Статья посвящена актуальной задаче оптимизации электротехнологических комплексов для обработки металлов давлением, которая может быть сформулирована как задача достижения предельных технико-экономических показателей технологических линий в условиях, обеспечивающих гибкость производственного процесса. Задача рассматривается на примере технологического комплекса для предварительного нагрева цилиндрических слитков из алюминиевых сплавов и их последующего прессования.

Представлены некоторые результаты решения поставленной задачи, полученные путем оптимизации режимов работы и конструктивных характеристик, как отдельно технологической стадии прессования, так и производственного комплекса в целом.

Introduction

Modern industry development demands very strict technological requirements to all stages of technological processes and a number of such requirements increases constantly. That is why the optimization of the production complexes by certain quality criteria becomes vitally important. The main objectives of optimization could be reached on the basis of modern theory of optimal control.

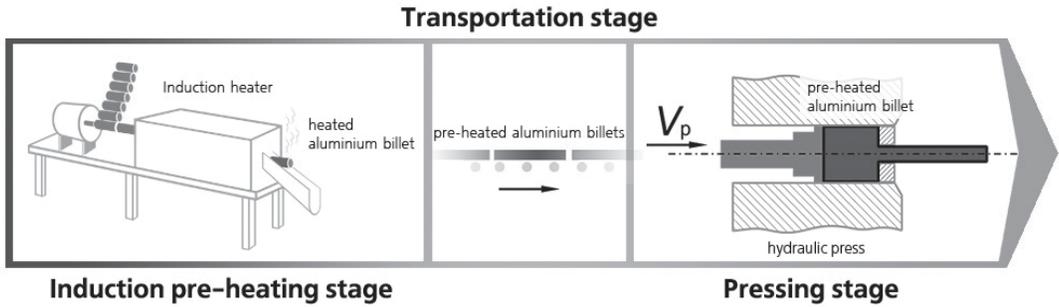
The major purpose of this paper is to describe how to use methods and tools of applied optimal control theory for optimization of technological lines “heater - pressing equipment”.

Representative example of the “heating – metal treatment by pressure” complex is a technological line for induction mass heating at the first stage and the pressing of aluminum alloy cylindrical billets at the final stage (Picture 1).

In the paper optimisation of a technological production complex for induction heating of cylindrical billets and subsequent hot forming operation by hydraulic presses will be used as an example of application of optimisation methods developed at Samara State Technical University.

Principal distinctive feature of the system optimisation approach is a necessity to take into consideration the sequence of the interrelated technological stages, including billet pre-heating by induction, its cooling during transportation, and subsequent plastic deformation during hot working, as a whole technological line being controlled. Combined optimisation criteria can be used for optimisation all these subsequent processes as a whole.

At the same time, for particular technological complex, it is important to reveal the cost function for each separate operation or stage of the technological production complex. A proper selection of local cost functions for optimisation of induction heating and pressing stages is based upon valid decomposition of a combined optimisation problem into separate optimisation problems for processes of heating and hot working [1].



Picture 1 – Structure of the technological line for metal pressing with induction pre-heating

In order to describe physical processes of production complex, it is necessary to use the system of interconnected numerical non-linear models which are extremely time-consuming. Since the optimisation procedures are planned to be used in real time, some assumptions and simplifications are applied in the process of models development.

The special optimisation algorithm that allows dramatically reduce the number of necessary iterations is used in the optimisation procedure.

A significant effect is achieved both for local and combined optimisation of steady state operation modes of technological line for cylindrical billet pressing with induction pre-heating.

1 Maximum productivity problem

Typically, a system productivity or production cost is considered as objective function I in optimisation problems. If maximum productivity should be obtained, then a minimal time, t_{end} required for the total cycle of a single billet processing can be considered as a cost function:

$$(1) \quad I_1 = t_{end} \rightarrow \mathbf{min}.$$

When minimum product cost is required, the following overall function can be used as a combined criterion of the optimal control problem:

$$(2) \quad I_2 = \sum_{i=1}^s C_i P_i + C_t t_{end} \rightarrow \mathbf{min}.$$

This criterion includes all costs components $P_i, i = \overline{1, s}$ with weighting coefficients C_i . Those coefficients make sense with respect to a relative costing representing an importance of each of the overall costs. Increase of the heating time leads to expense growth, according to criterion I_2 with correspondent weighting coefficient C_i . The value of P_i is primarily defined by energy expenses for heating and pressing, as well as by metal losses due to scale formation during the process of heating.

Let us consider the problem that provides maximum productivity of production complex “induction heater — pressing equipment” in the steady-state operation mode. It means that the goal of optimisation is to obtain minimum of criterion in formula (1) [1].

For this case the general statement of the optimal control problem could be formulated as follows. It is necessary to select such controls by heat power $u^*(t)$ and/or extrusion speed V_p^* restricted by pre-assigned set of constraints that provide a final press product of a desired length and hold overall cost criteria I at extreme value.

The formulated problem provides freedom in choosing the billet temperature distribution before the hot forming stage in a way that would satisfy requirements imposed by heating and pressing operations (with regard to transportation time). An appropriate temperature distribution that provides

minimal value of I_1 should be found as well as optimal control $u^*(t)$ and V_p^* . The selection of the billet temperature before hot forming out of set of admissible temperature distribution allows varying technological parameters within process flow sheets in order to provide optimal operational modes of induction heating installation (IHI) and pressing equipment.

1.1 Mathematical models of production complex stages

The ultimate goal of heating stage optimisation is to provide required thermal conditions of billet just before hot working operations [1]. From other side, pressing operation imposes certain requirements for initial temperature distribution. For example, there is the maximum range of admissible temperature variations within the heated billet prior to its hot forming by press. The temperature distribution within heated billet before the pressing stage affects essentially the overall performance index of technological complex. These are the reasons why the temperature distribution within the billet can be regarded as an output-controlled function. Therefore, as a subject of optimal control a technological line “induction heating – pressing” can be described by a system of heat transfer equations that represent heating, cooling (billet transfer) and pressing stages.

Required accuracy of temperature distribution within the billet is usually provided by acceptable level of three-dimensional temperature variations. For majority of practical applications the temperature distribution within the billet during the heating stage can be described by two- and three-dimensional mathematical models. Temperature distribution within cylindrical billet $\theta_1(l, y, t)$, evolving over heating process along coordinates (l, y) and in time t , can be described by two-dimensional heat transfer Fourier equation complemented by boundary condition of the third kind. Temperature distribution during billet transportation stage $\theta_2(l, y, t)$ can also be described by the same Fourier equation with only exception that there are no any heat sources [1]. Therefore, two-dimensional non-linear mathematical model is applied here for mathematical modeling of cylindrical heating system [1, 2]. This numerical two-dimensional model provides sufficient information regarding temperature profiles in axial and radial directions.

Computation of temperature distribution $\theta_3(l, y, t)$ evolving over extrusion process represents a problem that cannot be solved easily. One of difficulties deals with necessity to solve previously a highly complicated problem of mechanics of continua in order to define velocity field of metal flow and spatial distribution of plastic deformation energy that should be taken into account in heat transfer equation [1].

A simplified approach that is provided in [3] allows to solve the problem for stage of extrusion with constant speed $V_p = \text{const}$. In the considered case of axially symmetric deformation, simple analytical approximations for sought-for metal flow velocities $V_y(l, y, V_p)$, $V_l(l, y, V_p)$ and plastic deformation energy distribution $W^*(l, y, V_p)$ can be found according to well-known hypothesis of spherical sections. The numerical model that describes temperature distribution $\theta_3(l, y, t)$ during a pressing process can be applied to optimization of different modes of such a kind of hot forming operation. This numerical model has been obtained using finite-difference method [4, 5].

Unique feature of this production system deals with the fact that the separate operations are performed one after the other. The initial temperature distribution at the beginning of the billet transportation stage represents final temperature distribution at the end of the heating stage $t = t_1^0$. The temperature distribution before the hot forming stage represents the final temperature distribution at the end of the billet transportation stage $t = t_1^0 + \Delta_T$. These facts can be written in the form of the following equalities:

$$(3) \quad \theta_1(l, y, t_1^0) = \theta_2(l, y, t_1^0); \theta_3(l, y, t_1^0 + \Delta_T) = \theta_2(l, y, t_1^0 + \Delta_T),$$

where Δ_T is a duration of transportation stage; $t = t_1^0$ and $t = t_1^0 + \Delta_T$ are times of transitions from one stage of technological cycle to another. The equalities in formula (3) are imperative to bind billet temperature fields at times $t = t_1^0$ and $t = t_1^0 + \Delta_T$.

1.2 Decomposition of optimisation problem

For each fixed temperature distribution θ_2^* at the end of the heating stage the time-optimal modes of heater operation and extrusion press operation correspond to appropriate values of $t_{1\min}^0(\theta_2^*)$ and $t_{pr\min}(\theta_2^*)$. Here the value of $t_{1\min}^0(\theta_2^*)$ represents the minimum time required for billet heating up to the temperature $\theta_2^* \pm \varepsilon$. This also includes billet transportation stage. On the other hand, the value of $t_{pr\min}(\theta_2^*)$ is minimal possible time that is required for pressing the product of given length z_{end} under initial temperature $\theta_2^* \pm \varepsilon$, where $t_{pr\min} = (t_2^0 - t_1^0 - \Delta_T)_{\min}$; $t = t_2^0$ is end time of the pressing stage.

For steady-state mode, the minimum duration of production cycle $t_{c\min}(\theta_2^*)$ can be chosen as maximum between the values $t_{1\min}^0(\theta_2^*)$ and $t_{pr\min}(\theta_2^*)$ keeping in mind the type of IHI and specific way of combined operations of IHI and press. Obviously, the minimum value $t_{c\min}(\theta_2^*)$ is the optimal duration of the production cycle t_c^* that provides maximum possible productivity of the technological complex “induction heater — extrusion press”. Therefore, it is possible to write:

$$(4) \quad t_c^* = \min_{\theta_2^* \geq \theta_{2\min}^*} t_{c\min}(\theta_2^*) = \min_{\theta_2^* \geq \theta_{2\min}^*} \left[\max \left(\frac{1}{B_1} t_{1\min}^0(\theta_2^*), t_{pr\min}(\theta_2^*) + \psi \Delta_T \right) \right].$$

Here B_1 is the number of the billets that can be simultaneously heated in the IHI. The condition $B_1 = 1$ is met for static heating. The expression $B_1 > 1$ holds true for progressive heating mode with step-by-step in-line processing of billets through induction heater. In this case the time of extrusion t_{pr} is close to step duration that decreases in B_1 times compared to t_1^0 . The value of $\psi \in [0,1]$ in formula (4) is the factor that takes into account all possible variants of the correlation between the moment of unloading of the next billet and completion of the previous billet pressing [1].

If the functions $t_{1\min}^0(\theta_2^*)$ and $t_{pr\min}(\theta_2^*)$ are known the optimal temperature profile at the end of heating stage can be defined from formula (4) as follows:

$$(5) \quad \theta_{2e}^* = \arg \min_{\theta_2^* \geq \theta_{2\min}^*} \left[\max \left(\frac{1}{B_1} t_{1\min}^0(\theta_2^*), t_{pr\min}(\theta_2^*) + \psi \Delta_T \right) \right].$$

This value corresponds to the optimal duration t_c^* of production cycle and, therefore, the value of t_c^* can be defined as well. When the temperature distribution θ_{2e}^* is found then it is possible to perform the valid decomposition of the joint maximum productivity optimization problem of induction heater and extrusion press [1].

Decomposition allows to reduce the initial statement of the maximum productivity for the whole production line to the optimization problem for the cycle stage where duration is longer under condition of $\theta_2^* = \theta_{2e}^*$. This means that in order to decrease the value of t_c^* (see formula (4)) the technological stage with a maximum duration should be optimized with respect to time-optimal criterion. The technological stage that has the shortest duration should be optimized with respect to a cost func-

tion of the type of formula (2). The values of $t_{1\min}^0(\theta_2^*)$ and $t_{pr\min}(\theta_2^*)$ should be defined as a result of solutions of separate local time-optimal control problems for heating process and billet extrusion stages. The first problem can be solved by alternance method [1], but the second one represents independent problem and requires special consideration.

2 Results of optimisation

2.1 Optimisation of pressing stage

The problem of pressing stage optimisation has been solved for the most typical case when the extrusion operation is performed under constant speed V_p . It means that the pressing stage cannot be optimised with respect to a speed of extrusion. Under this condition, a maximum of possible productivity of the technological complex “induction heater — extrusion press” is restrained by the press equipment capabilities. In this case, the optimisation problem could be formulated as a problem of maximum approximation to the isothermal pressing mode that takes place when temperature in the die hole is held at maximum admissible level θ_{3ad}^* [1, 2]. In this problem the temperature distribution before the pressing stage could be considered as a sought-for output function to be optimised. As far as the value θ_{3ad}^* affects essentially the quality of press-product material structure, energetic characteristics of pressing equipment operation, and therefore, overall performance index, it is necessary to provide the minimum deviation of temperature in the die hole $\theta_3(l_{dh}, 1, t)$ from maximum admissible value θ_{3ad}^* at each moment of pressing time. This deviation could be estimated by uniform Chebyshev' metrics and the following cost function could be considered for the optimisation of pressing stage:

$$(6) \quad I = \max_{t \in [t_1^0 + \Delta_T; t_2^0]} \left| \theta_3(l_{dh}, 1, t) - \theta_{3ad} \right| \rightarrow \min_{\theta_3(y, t_1^0 + \Delta_T) \in \Omega} .$$

If a non-uniformity of temperature distribution along billet diameter could be neglected, one-dimensional temperature distribution along a billet length $\theta_3(y, t_1^0 + \Delta_T)$ could be chosen as an output function to be optimised. This function has been represented as a piece-wise constant function of the parameter vector $\Delta^{(S)}$ in the following form:

$$(7) \quad \theta_3(l, y, t_1^0 + \Delta_T) = \theta_3^0(y, \Delta^{(S)}) = \begin{cases} \theta_{02} + \frac{(\theta_{01} - \theta_{02}) \cdot y}{y_{01}}, & 0 < y \leq y_{01}, \\ \theta_{02} + \frac{(\theta_{00} - \theta_{01}) \cdot (y - y_{01})}{1 - y_{01}}, & y_{01} < y \leq 1, \end{cases}$$

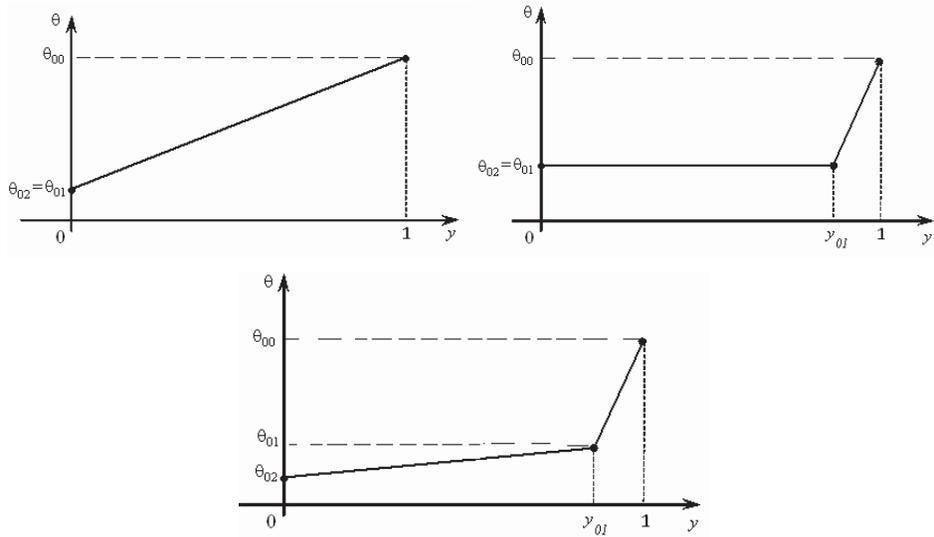
where $\Delta^{(S)} = (\Delta_i), i = \overline{1, S}, 2 \leq S \leq 4: \Delta_1 = \theta_{00}; \Delta_2 = \theta_{02}; \Delta_3 = y_{01}; \Delta_4 = \theta_{01}$. The function of formula (7) defines the set of possible temperature profiles along a billet length before the pressing stage (Picture 2) and allows to reduce the initial optimisation problem to the problem of mathematical programming that could be written as follows:

$$(8) \quad I(\Delta^{(S)}) = \max_{t \in [t_1^0 + \Delta_T; t_2^0]} \left| \theta_3(l_{dh}, 1, t, \Delta^{(S)}) - \theta_{3ad} \right| \rightarrow \min_{\Delta^{(S)}} .$$

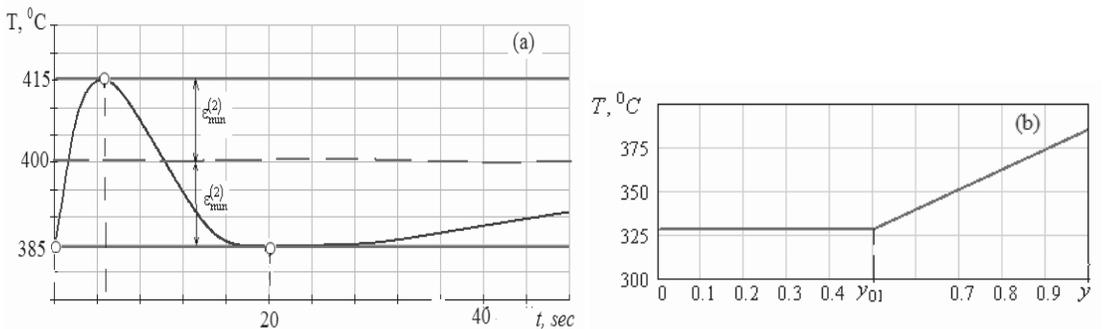
The problem in formula (8) has been solved using alternance properties of optimal variation of temperature in the die hole vs. time during extrusion process under constant speed [4].

Some computational results in absolute values of temperature T and time t are presented on Picture 3. The diagram (a) shows the variation of temperature in the die whole vs time of optimal pressing of aluminium alloy billet for the following process data: $V_p = 450 \text{ mm/min}$; diameter of

billet is 250 mm; billet length is 430 mm; $\theta_{3ad}^* = 400^{\circ}C$. It is clear visible that the maximum deviation of temperature in the die hole $\theta_3(l_{dh}, 1, t)$ from maximum admissible value θ_{3ad}^* at each moment of pressing time does not exceed $15^{\circ}C$ that is in fairly good correspondence with technological requirements. The optimal temperature profile along the billet length before pressing stage is shown by the diagram (b) for the fixed value $y_{01} = 0.5$ and $\theta_{01} = \theta_{02}$ in formula (7). So if it necessary to optimise pressing operation, it can be reasonable to provide an initial positive drop of temperature along the billet length (in the direction of extrusion). In this case the problem of technological line optimisation can be solved under condition of gradient billet heating.



Picture 2 – Temperature profiles along a billet length before the pressing operation

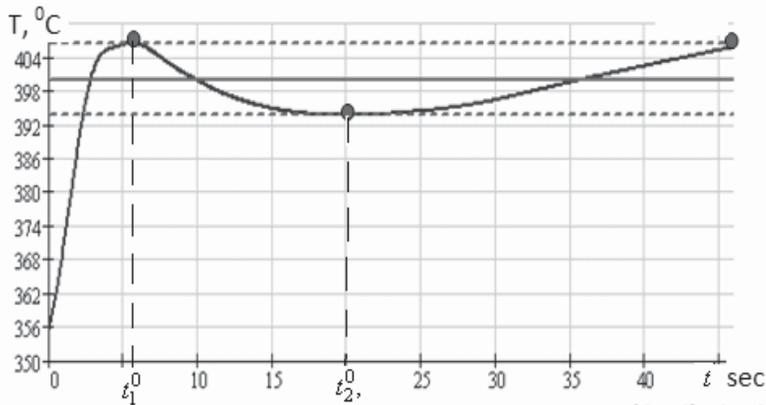


Picture 3 – Results of pressing stage optimization

2.2 Combined optimisation of heating and pressing stages

Let us consider the problem of combined optimisation for the whole technological line "induction heating - extrusion pressing". The criterion in formula (8) could be considered as the cost function and the powers $P_1(t)$, $P_2(t)$ of autonomously controlled sections could be chosen as parameters

to be optimised. The following problem has been solved for optimisation a technological line as a whole. It is necessary to select such controls by section powers $P_1(t)$ and $P_2(t)$ restricted by pre-assigned values $P_{1\max}$; $P_{2\max}$ that provide a final press-product of desired length and hold overall cost criterion in formula (8) at extreme value. The results of problem solution are presented on Picture 4. It is shown that the maximum deviation of temperature in the die hole $\theta_3(l_{dh}, 1, t)$ from maximum admissible value θ_{3ad}^* at each moment of pressing time does not exceed $7^{\circ}C$ that is even better than the result of optimisation of pressing stage given above [2].



Picture 4 – Temperature in the die hole vs time during pressing of aluminium alloy cylindrical billet

Conclusions

A significant effect is achieved both for local and combined optimisation of steady state operation modes of technological line for metal hot forming with induction pre-heating. It is necessary to underline that the described approach leads to determining capabilities of a control algorithm out of the framework of traditional “servicing of” technology and ensures active participation in technology formation. This system approach deals with optimisation of induction heating installation based on requirements to the end product quality under condition of non-fixed temperature distribution within the billet after the heating stage.

Literature

- [1] E. Rapoport, Yu. Pleshivtseva. Optimal Control of Induction Heating Processes. DK6039, CRC Press/Taylor & Francis Group, 6000 Broken Sound Parkway, NW Suite, 300. Boca Raton, FL 33487 (USA), 2007, 349 p.
- [2] Pleshivtseva Yu.E., Afinogentov A.A. Optimal control of energy-technological processes in production complexes. Moscow: Izvestiya vysshikh uchebnykh zavedenij, Electromechanics, 2008, №3. p. 51-55.
- [3] Zimin L.S., Rapoport E.Ya., and Kondrashov S.V. Determination of temperature field of aluminium alloy cylindrical billet during axially symmetrical extrusion. Algoritmy i programmy, 1986, № 1, pp. 64.
- [4] Afinogentov A.A. Simulation of cylindrical billets hot forming process. Vestnik of Samara State Technical University. Volume: Physical and mathematical sciences, 2007, № 2(15), pp. 170 – 172.
- [5] Pleshivtseva Yu., Rapoport E., Afinogentov A., Yu. Shemyakin, Nacke B., Nikanorov A. Application of optimal control theory for optimisation of metal hot forming lines with induction pre-heating, Proc.: International Seminar “Heating by Electromagnetic Sources” HES-10, Padua, Italy, May 19-21, 2010 (in print).

NUMERICAL SIMULATION OF MULTI-PHYSICS DYNAMIC PROCESSES IN INDUCTION HEATING SYSTEMS GRANTED BY GERMAN ACADEMIC EXCHANGE SERVICE (DAAD)

E. Baake¹, K. Blinov¹, S. Korshikov², O. Sharapova²

¹Institute of Electrotechnology, Leibniz University of Hannover ,
D-30167 Hannover, Wilhelm-Busch-Str. 4, Germany
etp@etp.uni-hannover.de

phone: +49 511 762-3248, fax: +49 511 762-3275

²Samara State Technical University SamGTU
443100, Samara, Molodogvardeyskaya str., 244, Russia
usat@samgtu.ru

phone: +7 846 332-4234, fax: +7 846 278-4440

Key words: *numerical simulation, induction through heating, transient mode, temperature distribution, heating by rotation, DC magnetic field, thermal stresses*

Abstract

В статье представлены результаты научно-исследовательских работ аспирантов Института Электротехнологий Университета им. Лейбница г. Ганновер (Германия) и Самарского государственного технического университета, проводимых в рамках программы совместных научных исследований и поддержанных Германской службой академических обменов DAAD. В статье представлены два проекта в области численного моделирования температурных полей и полей термонапряжений в процессах традиционного индукционного нагрева в проходных индукционных печах и индукционного нагрева алюминиевой заготовки, вращающейся в магнитном поле постоянного тока.

Introduction

The content of this paper is based on the results of recent research activities carried out by young Phd-students in the frame of the close co-operation between the Institute of Electrotechnology (ETP) of the Leibniz University of Hannover, Germany, and the Samara State Technical University (SamGTU) on the one hand and the St. Petersburg Electrotechnical University (LETI) on the other hand.

The successful and fruitful co-operation between the ETP in Hannover and SamGTU in Samara as well as LETI in St. Petersburg has started already approximately 20 years ago. Many scientific projects and an intensive exchange of young scientists and Professors from both sides have been done during this time. Due to the financial and organisational help of the German Academic Exchange Service (DAAD) the support of young Russian student, PhD-students as well experienced post-doc researchers could be improved considerable in the last years. For example by the “Michael-Lomonosov-Programme” highly qualified scientists from St. Petersburg and Samara have been supported in order to work in scientific cooperation projects. Up to now more than 30 students and PhD-students could receive an important financial support in Russia and the unique possibility to stay and work in Hannover in the frame the “Leonhard Euler” Programme. As an example of the scientific work carried out by the Russian servitors supported by the DAAD this paper describes two projects, one in the field of numerical modelling of induction through heaters for forging in transient operation modes and the other in the field of simulation of aluminium billet heating by rotation in superconducting coils. The authors of this paper highly appreciate the support of the DAAD without this it would be not possible to realize the work and the results described as follows.

1 Numerical modelling of induction through heaters for forging in transient operation modes

Induction through heaters for forging are usually designed to provide the required characteristics in so called “quasi” steady-state modes where temperature field in the heater is constant or it is periodically changed with step-by-step movement of the workpiece. Optimization of operation in “quasi” steady-state modes is usually realized on the step of the heater design. This task requires numerical modelling of induction through heating process in “quasi” steady-state operation mode only [1].

Nevertheless, in industrial practice the induction through heaters operate under various controlled and uncontrolled disturbances. The heater regime can be under changing more than a half of total operation time. Transients because of controlled excitations include the system reaction on initial start of the heater, restart after technological brakes, changing the workpiece specification or production rate of the installation. The heater reacts on uncontrolled disturbances by more or less slight continuous transients in operation mode. Special group of numerical models is required to simulate any kind of transient modes in induction heating systems. In transient models thermal analysis must include time additionally to one, two or three space coordinates. Transient modelling of induction through heating should be based on special algorithm which provides coupling between electromagnetic and thermal analysis.

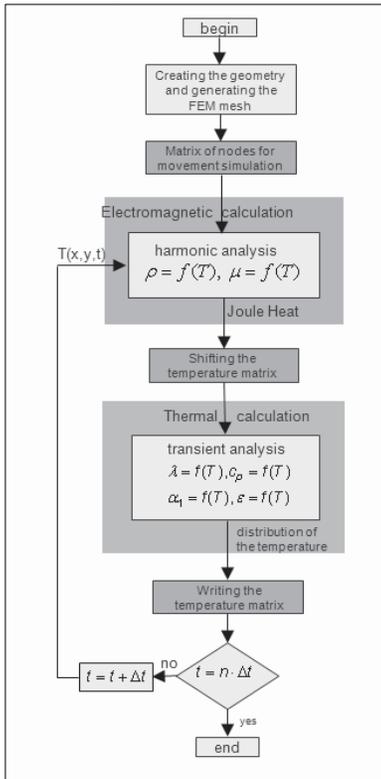


Figure 1 – The algorithm for coupling of electromagnetic and thermal analysis in transient modes with the workpiece movement

One 2D transient numerical model of induction through heater for cylindrical billets with continuous movement has been developed at the Institute of Electrotechnology, Leibniz University of Hannover according to an algorithm shown in Figure 1. For numerical simulation continuously running physical process is replaced by big enough number of time steps. Electromagnetic and thermal analysis is carried out at each time step of modelling. The Joule heat distribution in the workpiece, calculated in electromagnetic analysis, is used as an excitation for thermal analysis, which provides simulation of thermal process during the running time step. For electromagnetic analysis at each time step, temperature dependent electro-physical material properties are corrected according to temperature distribution in the workpiece after the previous time step. Thermal analysis includes simulation of thermal losses by convection and radiation from all open surfaces of the calculated system. Heat flux by radiation is calculated taking into account to view angles. Input of voltage or current of induction coil in electromagnetic analysis at each time step individually allows simulating various transient as well as “quasi” steady-state operation modes of the heater. The model is realized using Finite Element Method on the basis of commercial program package ANSYS and it can be also used for simulation of induction through heaters with different control systems.

One robust way to implement the workpiece movement is based on shifting the temperature field before each time step outside the thermal analysis [2]. This approach can be applied only for induction heating systems with continuous movement of endless workpieces with constant cross-section. The algorithm requires uniform numerical mesh in the workpiece in the direction of motion. Speed of the workpiece is taken into account via value of time step.

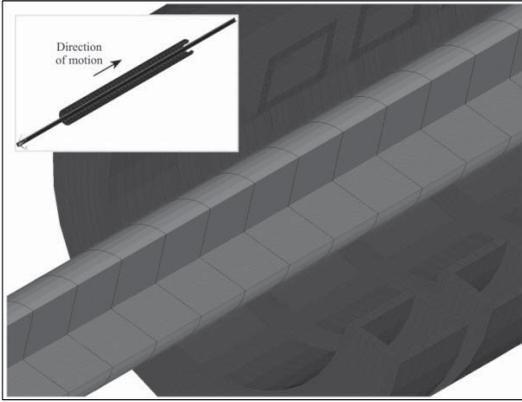


Figure 2 – Geometry and FEM mesh of induction coil and workpiece in the simulated through heater for forging

The induction coil current has been kept constant.

The described approach has been successfully applied to investigate numerically several induction heating systems. One example of simulated induction through heater for forging is shown in Figure 2. An endless cylindrical steel workpiece with diameter of 34 mm, to be heated in longitudinal magnetic flux, is transported through one solenoid induction coil of 37 windings with length of 1.4 m. As it is typically made in induction heaters for forging, the coil is casted to refractory for better thermal insulation and protection. Numerical mesh in the workpiece is uniform in the direction of motion.

The modelled heater has been tested in transient as well as in “quasi” steady-state modes of operation. Temperature field in the workpiece and refractory in the reached after the first start “quasi” steady-state mode is shown in Figure 3.

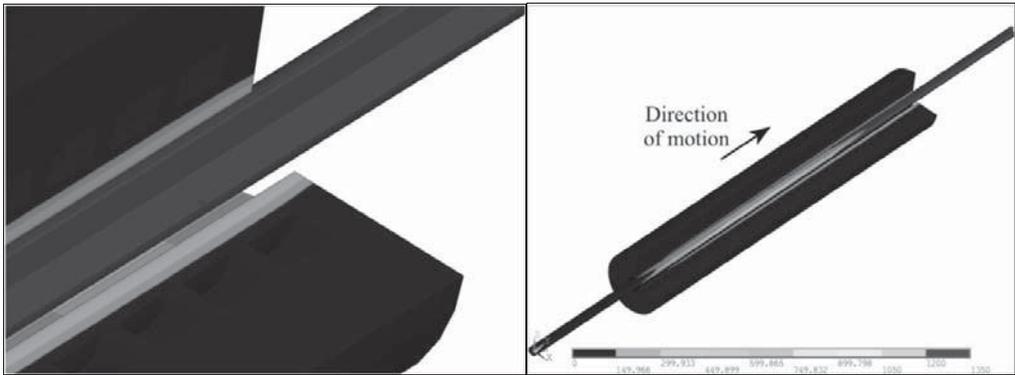


Figure 3 – Temperature field in the workpiece and refractory in “quasi” steady-state mode



Figure 4 – Temperature distribution along the workpiece during the first start of the heater

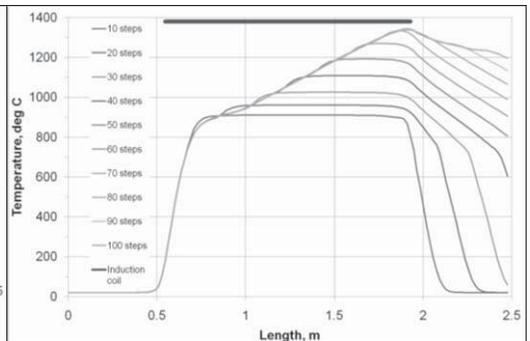


Figure 5 – Temperature distributions at the surface and in the centre of the workpiece in “quasi” steady-state mode

The modelled heater has been investigated during its first start with cold workpiece and refractory. Dynamics of surface temperature distribution along the workpiece is shown in Figure 4. One can see that the transient mode of first start is finished inside the induction coil after 80 time steps. All electrical data of the heater during its transient operation are available as well.

Temperature distributions at the surface and in the centre of the workpiece in “quasi” steady-state mode are presented in Figure 5. The results, received using the developed model, have been confirmed by comparison with numerical simulation using the model for steady-state mode.

2 Numerical simulation of thermal stresses fields in process of Alu-billet induction heating by rotation

Induction heating of aluminum billets by rotating in DC magnetic field produced by superconductive coils is an innovative technique. Aluminum billets are heated from ambient temperature up to 500°C prior to extrusion. The primary type of furnace used for preheating billet before extrusion is induction heater. The choice of heater depends on many factors. The most important criterion is final cost of products and product quality [3].

The power rating of the induction heaters ranges from several hundred kilowatts up to dozen megawatts. Aluminum is a low-resistive metal that makes possible to apply low frequencies. Utilizing low frequencies at 50 - 60 Hz leads to such benefits as low capital cost of equipment and low energy consumption. From other hand low electrical resistive metals are known to have a low coil electrical efficiency. The efficiency of conventional induction heaters does not exceed 50 - 60 % because 40 - 50 % of total power is transformed into heat in the copper windings and removed by the cooling water. To improve the process efficiency an innovative induction technique has been proposed. A novel approach is based on generating a magnetic field by DC current in superconductive coils. Rotating of the billet in the magnetic field leads to the eddy current induction and heat generation in the billet. The principle of rotating induction billet heater, where aluminum billet in diameter 215 mm rotates with the set frequency in a magnetic field of a direct current, that produced by the coil with configuration shown in Figures 6-8. Theoretically this approach should increase the electrical efficiency of the aluminum billet heater up to 90 % [3].

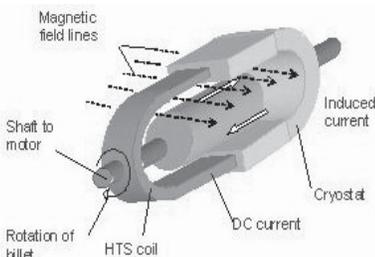


Figure 6 – Principle of rotating induction billet heater

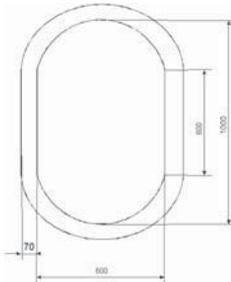


Figure 7 – Coil design

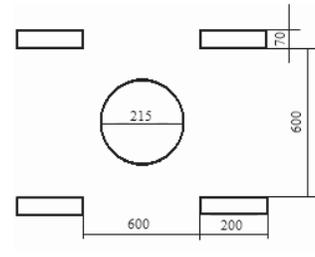


Figure 8 – Cross-section of billet and coil

For efficient operation of technological line “heater-hot forming equipment”, at the end of heating process it is necessary to provide a required final temperature distribution within the heated workpiece taking into account the constraint on the admissible thermal stress value. To obtain this purpose the adequate models of the temperature and thermal stresses fields should be developed and their dependence on various process parameters should be explored.

In order to investigate thermal stresses fields during heating process the numerical 2D ANSYS model has been used. The model allows to simulate the DC magnetic field, the rotation of the billet inside the field and the eddy currents generation inside the aluminum billets. As the results of

simulation the distribution of electric current and heat density, temperature and thermal stresses fields within the billet cross-section could be obtained. The model mesh within the billet cross-section is shown on Figure 9.

Investigations of thermal stresses fields were carried out for the constant values of parameters characterizing mechanical properties of material, such as Young's modulus, Poisson's ratio and others.

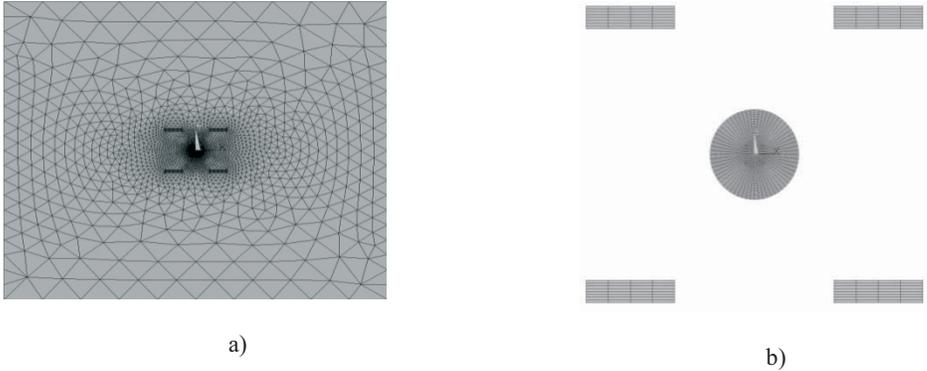


Figure 9 – 2D mesh of induction heating model:

a) Mesh of induction system and ambient area, b) Mesh of billet and coils cross-sections

Figure 10 shows that the temperature during heating process monotonously increases. The maximum temperature is observed at the billet surface and the minimum one is in the billet center. The temperatures at the surface and in the middle and center points differ not more than by 80 °C. At the same time the temperatures in all points exceed maximum admissible value of 500 °C at the final stage of heating process. Therefore, the constraint on maximum admissible temperature during heating process should be taken into account.

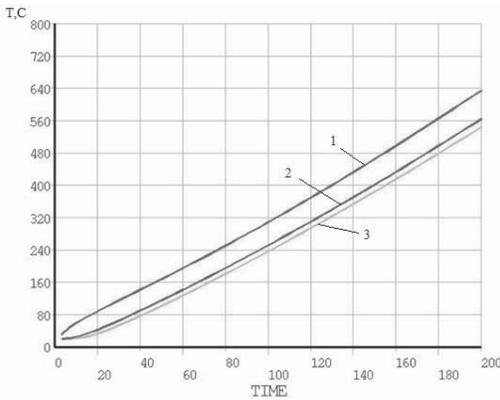


Figure 10 – Time-temperature history for heating process (1- surface; 2-middle point; 3-center)

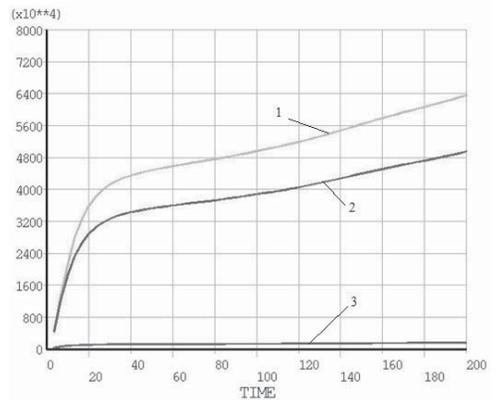


Figure 11 – Thermal stresses vs time during heating process (1- center; 2-middle point; 3- surface)

The maximum thermal stresses are observed in the center of the billet, and the minimum one is on the surface (Fig. 11). The difference between thermal stresses at the surface and in the middle and center points increases dramatically at the initial stage of heating and grows continuously during the whole heating process. It is clear visible that during heating process the thermal stresses in the middle and center points exceed maximum admissible value of 30 MPa. So, the constraint on maximum admissible thermal stress during heating process should be considered as well.

From the above parametric studies it follows that the main features of innovative induction heating technology are in good correspondence with the well known regularities of typical induction heating processes. The numerical 2D model of aluminum billet heating by rotation in DC magnetic field could be used for control and optimization of the process taking into account the main technological constraints.

Acknowledgements

This work has been supported by the German Academic Exchange Service (DAAD) and the Russian Foundation for Fundamental Research.

Literature

- [1] Pleshivtseva, Yu.; Rapoport E.; Efimov A.; Nacke B.; Nikanorov A.: Optimal design and control of induction heaters for forging industry. Proceedings of International Seminar "Heating by Electromagnetic Sources" HES-07, Padua, Italy, June 19-22, 2007. – pp.251-258
- [2] Galunin, S.; Zlobina, M.; Blinov, K.; Nikanorov, A.; Zedler, T.; Nacke, B.: Numerical analysis of coupled physics for induction heating of movable workpieces. Proceedings of International Scientific Colloquium Modelling for Electromagnetic Processing (MEP2008), Hannover, Germany, October 27 – 29, 2008, pp.59-64
- [3] Zlobina, M.; Nacke B.; Nikanorov, A.: Adaptive Induction System for Heating of Aluminium Billet by Rotation in DC Magnetic Field. Proceedings of International Scientific Colloquium Modelling for Electromagnetic Processing (MEP2008), Hannover, Germany, October 27 – 29, 2008, pp.349-354

MATHEMATICAL SIMULATION OF PRESSING STAGE IN THE “HEATER - HOT FORMING EQUIPMENT” LINE¹

Yu.V. Shemyakin

Samara State Technical University
443100, Samara, Molodogvardeyskaya Str., 244, Russia

Key words: numerical simulation, induction heating, pressing, system approach, special software, finite element method, optimisation, local and combined criteria

Abstract

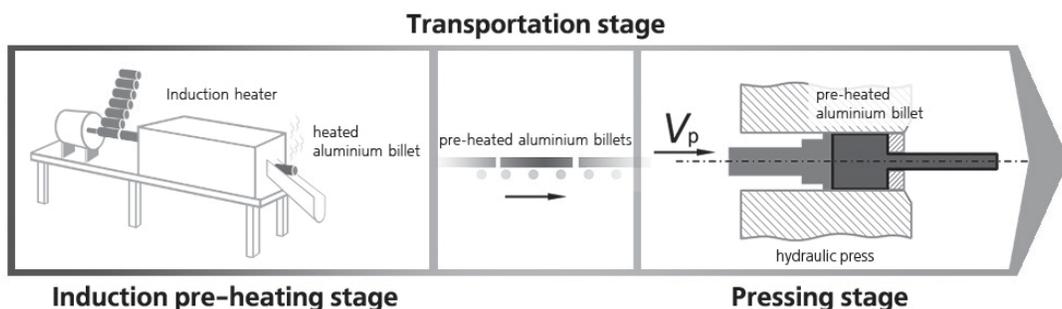
Предлагается подход к численному моделированию процесса прессования предварительно нагретой цилиндрической заготовки из алюминиевого сплава. Разработанная двумерная нелинейная модель может быть использована как в процедуре локальной оптимизации процесса прессования, так и для комплексной оптимизации технологической линии «нагрев – обработка металла давлением» по совокупным критериям качества.

Introduction

Nowadays the major share of all heavy industries bases on production systems consisting of the heating systems and the equipment for metal hot forming which could be represented as united technological lines [1].

Typical structure of technological line “heating – metal treatment by pressure” includes three technological stages: metal heating, transportation of heated billets, hot forming of pre-heated billets. Representative example of the “heating – metal treatment by pressure” line binds the induction mass heating as the one of the most widespread methods of heating at the first stage of the technological line and the pressing of aluminum alloy cylindrical billets at the final stage (Picture 1).

From point of view of **system approach** the whole sequence of the interrelated technological operations should be taken into account.



Picture 1 – Structure of technological line for metal pressing with induction pre-heating

Modern industry development demands very strict typical technological requirements to all stages and the number of such requirements increases constantly. That is why the optimization of the industrial lines by the certain quality criteria becomes vitally important.

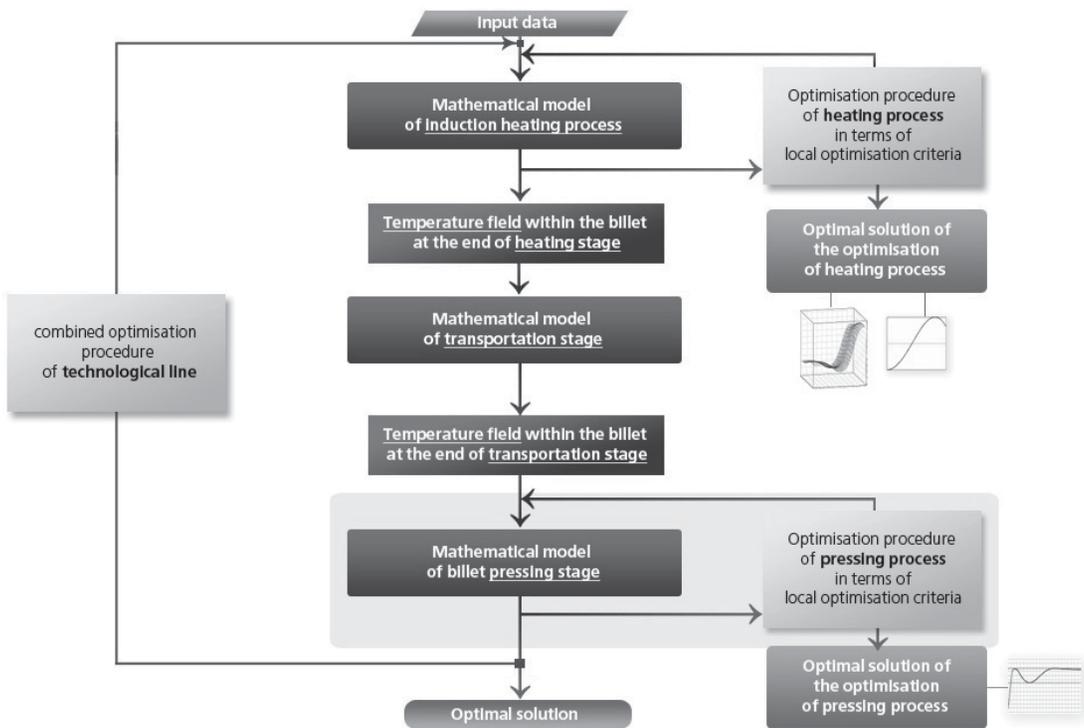
The global goal of the research is the optimisation of the complete technological line in terms of local and combined optimization criteria. Mathematical modeling that is used in the optimisation pro-

¹ This research was supported by German Academic Exchange Service (DAAD) within the bounds of Leonard Euler grant.

cedures is one of the major factors in the successful design of both induction heating and pressing processes [2].

Interconnected mathematical models of the mentioned technological stages are considered to realize the optimisation procedures. Initial temperature distribution at the beginning of the billet transportation stage represents final temperature distribution at the end of the heating stage at the same time the temperature distribution before the hot forming represents final temperature distribution at the end of the billet transportation stage. These facts are imperative to combine workpiece temperature states during all stages of the “induction heating–metal hot forming” technological cycle [3, 4].

From one side the stages of induction heating and pressing could be optimized separately with regard to local optimization criteria and from the other side the whole technological line could be optimized in terms of combined optimization criteria (Picture 2) [4].



Picture 2 – Main approaches of optimisation of “heating – metal treatment by pressure” line

The main goal of paper is the mathematical simulation of pressing stage in the «heater – hot forming line» for pre-heated cylindrical aluminium alloy billets.

1 Numerical axisymmetrical 2d model of pressing

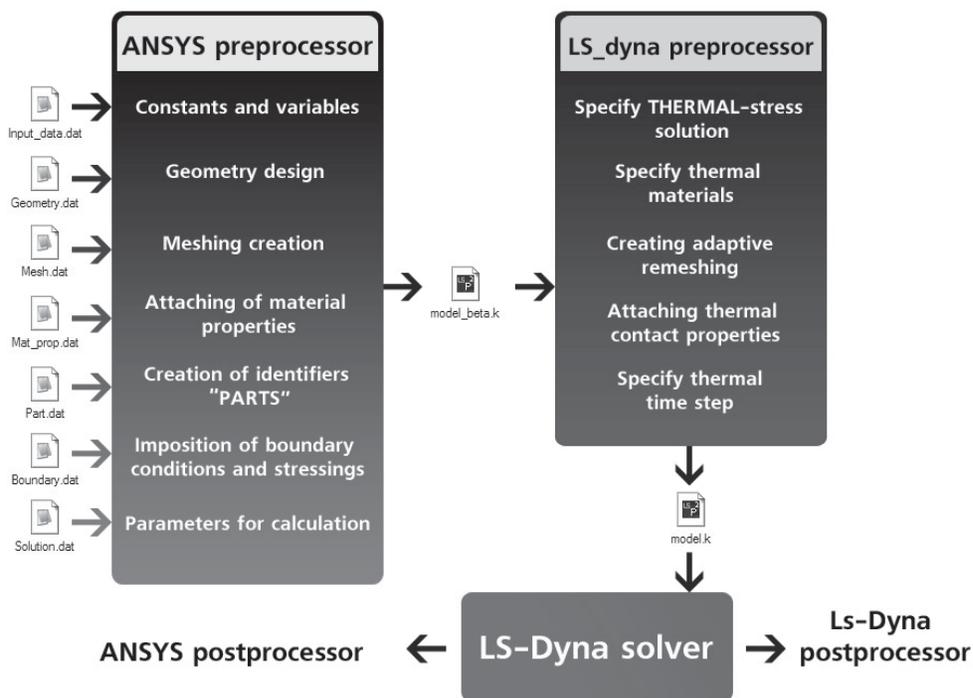
As far as the temperature field within the billet is the output data for the optimisation procedures it is necessary to provide the simulation of temperature field evolving over pressing stage. The temperature distribution during the pressing process stage could be described with the two-dimensional Fourier heat conductivity Equation for cylindrical billet taking into account internal sources of heating due to deformation of treated billet and additional heat flow due to the contact friction. Computation of temperature distribution evolving over extrusion process represents problem that cannot be solved easily. One of difficulties deals with necessity to solve previously a highly complicated prob-

lem of mechanics of continua in order to define velocity field of metal flow and spatial distribution of plastic deformation energy that should be taken into account in heat transfer equation [4]. Mentioned equation can be solved numerical only. Numerical coupled thermal-stress simulation of pressing process applying the finite elements method is suggested here.

Due to comparative analysis of special software LS-Dyna has been chosen for simulation. LS-Dyna program complex is used to solve multi-physics problems including solid mechanics, heat transfer, and fluid dynamics either as separate phenomena or as coupled physics, e.g., thermal stress or fluid structure interaction.

As far as, Ls-Dyna is integrated into ANSYS as a solver, it's possible to use a lot of commands and data directly through the ANSYS interface. At the same time some necessary commands for coupled thermal-stress analysis are not available using only ANSYS interface. That is why to actualize the pressing stage simulation the LS-Dyna Pre-Postprocessor has been used in addition.

The flow chart (Picture 3) shows specific approach to simulation of pressing cylindrical aluminum billets through the extrusion die hole using ANSYS\ Ls-Dyna in terms of coupled thermal-stress analysis.

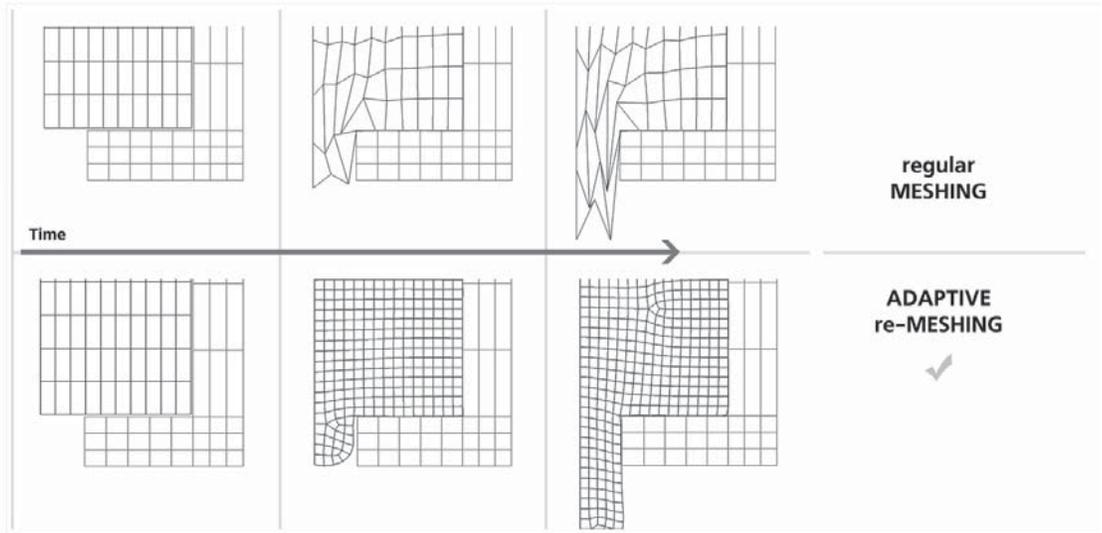


Picture 3 – Simulation approach to modeling of pressing process

It's possible to use ANSYS for specifying constants and variables, design geometry and mesh, attaching material properties and making the specific identifiers called "PARTs". ANSYS allows setting initial conditions and stressing and making it ready for solving in Ls-Dyna. Material type 004_MAT_ELASTIC_PLASTIC_THERMAL was applied for modeling pressing of aluminium billets due to the temperature dependence of material properties such as corresponding yield stress and Young's modulus that must be considered. But it is necessary to specify much more option to make it complete for the computation. That is why it's required to generate the Ls-Dyna keyword input file (*.k-file), but instead of starting the computation via LS-Dyna, the input file has to be uploaded into LSPrePost program (Ls-Dyna's pre&post processor) for making required modifications.

It's valid to change the solution type to the coupled thermal-stress (*CONTROL_SOLUTION = 2) through the LSPrePost interface as well as to specify the thermal material properties (*MAT_THERMAL ISOTROPIC) and the thermal contacts for the operating RARTs (*CONTACT_2D_AUTOMATIC_SURFACE_TO_SURFACE_(THERMAL)). Then after the setting of thermal time-step (*CONTROL_THERMAL_TIMESTEP) and applying the ADAPTIVE 2D re-meshing (*CONTROL_ADAPTIVE) the final k-file have to be generated and sent to computation with LS-Dyna solver.

The option of adaptive re-meshing should be turned on and set properly because the regular mesh is destroyed during computation and does not allow obtaining adequate results (Picture 4).



Picture 4 – Comparison of regular mesh and adaptive re-meshing

Several time steps have to be executed before the implementation of adaptive re-meshing because the adaptive re-meshing procedure interpolates and extrapolates node data and it's possible to get wrong answers if the time converged solution wasn't performed before the adaptivity.

COMPUTATIONAL RESULTS

The hot pressing of aluminum alloy cylindrical billet through the extrusion die hole with the specified material properties has been investigated. Table 1 presents the initial data for computation such as material properties [5], geometry of the billet and pressing settings.

Table 1 – Initial data for computation

2024 aluminum alloy material properties

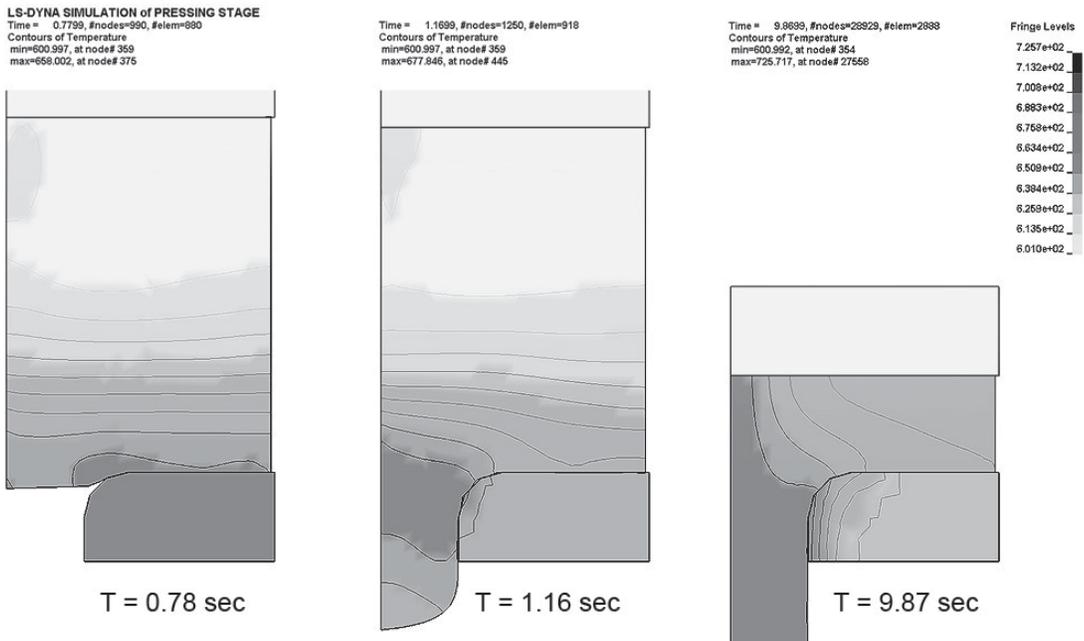
Property	Value
Yield stress	0.345 GPa
Young's modulus	71.0 GPa
Tangent modulus	0.46 GPa
Density	2923 kg/m³
Poisson ratio	0.334
Heat capacity	904 J/kg K
Thermal conductivity	222 W/m K

The end of Table 1

Geometry of the billet and pressing settings

Diameter	0.15 m
Length	0.5 m
Pressing speed	0.008 m/sec

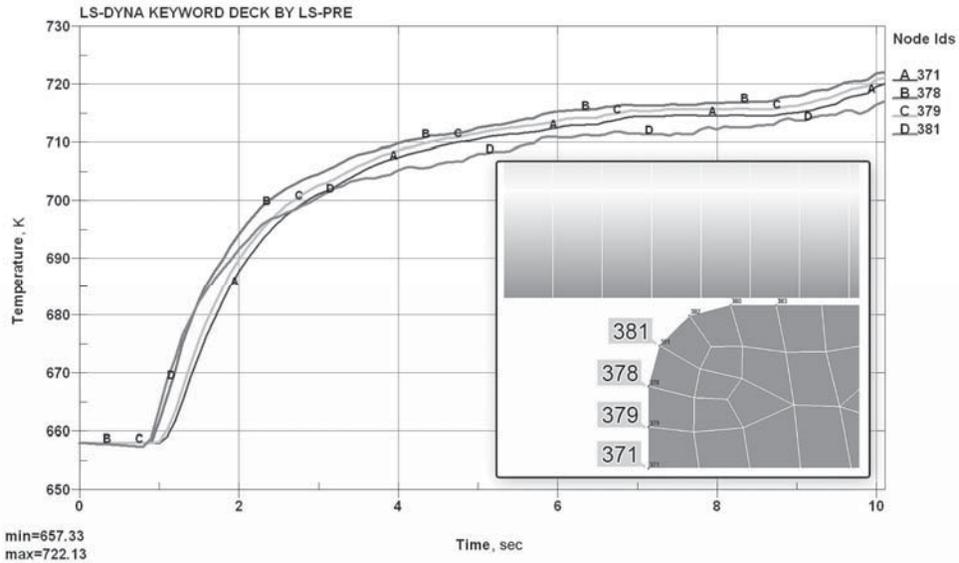
Suggested approach allows receiving dynamic nonlinear simulation and presents the temperature distribution of the billet with the required accuracy at each time step (Picture 5). This simulation model describes also the heat-transfer processes between operating parts: billet/extrusion die and billet/punch.



Picture 5 – Simulation of pressing of pre-heated cylindrical aluminium billets

It's shown that this method provides qualitative characteristics that are in good correspondence with basic physical characteristics and known regularities of physical processes regulations.

Using ASCII Database for the data output with TPRINT option turned on it is possible to monitor the temperature in the die hole (the hottest place) that is necessary for being used in the optimisation procedure mentioned above. The plot described the temperature variation at the nodes of the border of the extrusion die is given (Picture 6).



Picture 6 – Temperature variation at the extrusion die hole

Proposed approach for the simulation of pressing stage in the “induction heating–metal hot forming” technological line is useful to be applied in the optimisation procedures for both local and combined optimisation criteria.

Conclusion

- The new simulation approach and numerical model of pressing stage which is in fairly correspondence with known physical processes regulations was developed and presented at this paper;
- This axisymmetrical two-dimensional model allows to simulate numerically the temperature and stress fields evolving over pressing of aluminium alloy cylindrical billet.
- There are big potentials in application of this simulation method for optimization of a technological line “induction heater – hot working equipment”. So the subsequent work should be aimed at verification and experimental model validation and adjustments of the model to be used in optimisation procedures such as model simplifications; development of connection interfaces and data exchanging files modules; optimisation of model mesh, time-step, thermal time-step, accuracy, etc.

Literature

- [1] Rapoport E.Ya. (1993). Optimization of induction heating of metals. Moscow: Metallurgia, 279 p.
- [2] Zimin L.S., Rapoport E.Ya., and Kondrashov S.V. (1986). Determination of temperature field of aluminium alloy cylindrical billet during axially symmetrical extrusion. *Algoritmy i programmy*, № 1, p. 64.
- [3] Rapoport E., Pleshivtseva Yu. (2006). Optimal Control of Induction Heating Processes. DK6039, CRC Press/Taylor & Francis Group, (USA), 349 p.
- [4] Pleshivtseva Yu., Rapoport E., Afinogentov A., Yu. Shemyakin, Nacke B., Nikanorov A. Application of optimal control theory for optimisation of metal hot forming lines with induction pre-heating, Proc.: International Seminar “Heating by Electromagnetic Sources” HES-10, Padua, Italy, May 19-21, 2010.in print
- [5] Military Handbook - MIL-HDBK-5H: Metallic Materials and Elements for Aerospace Vehicle Structures (Knovel Interactive Edition) U.S. Department of Defense (2003), 1721 p.

ИССЛЕДОВАНИЕ ЭНЕРГЕТИЧЕСКИХ ХАРАКТЕРИСТИК ИНДУКЦИОННОЙ СИСТЕМЫ С ХОЛОДНЫМ ТИГЛЕМ

А.Н. Шатунов, И.В. Позняк, А.Ю. Печенков, А.И. Максимов, В.В. Кичигин
Санкт-Петербургский Государственный Электротехнический университет «ЛЭТИ»
197376, Санкт-Петербург, ул. Профессора Попова, 5, Россия
root@post.etu.spb.ru
тел: +7 (812) 346-44-87, факс: +7 (812) 346-27-58

Ключевые слова: *индукционная система, холодный тигель, плавка оксидов, поверхностный эффект, математическое моделирование*

Abstract

Investigation of energetical parameters of the induction system with slitted water-cool copper crucible for melting of oxide materials and present dependences of the powerlosses in the crucible's section and efficiency of the induction system via characteristics of the melt. The results of the calculation were received with using mathematical modeling in the package ANSYS.

Введение

Индукционные установки с холодным тиглем для плавки оксидов используются в промышленности десятки лет [1-3]. Однако полного представления о происходящих электромагнитных явлениях в секциях тигля до сих пор нет. В частности, недостаточно изучена зависимость потерь в секциях тигля от параметров загрузки, что может вносить погрешности при проектировании индукционных установок и управлении технологическими процессами.

В статье исследуются энергетические характеристики индукционной системы с разрезным водоохлаждаемым медным тиглем для плавки оксидных материалов на частоте 1,76 МГц и приводятся зависимости мощности электрических потерь в секциях тигля и электрического КПД индукционной системы от характеристик загрузки. Результаты расчетов получены с использованием математического моделирования в программном пакете ANSYS.

1 Геометрия

Эскиз индукционной системы представлен на рис. 1. Разрезной водоохлаждаемый тигель представляет собой набор секций из Π образных трубок, собранный вокруг основания. Фиксация трубок осуществляется с помощью медного хомута. Хомут и секции электроизолированы друг от друга, что приводит к отсутствию перетекания токов с секций на хомут и обратно. Для защиты основания тигля от воздействия высокотемпературного расплава в конструкции индукционной системы предусмотрена медная водоохлаждаемая пластина, которая исполняет роль дна тигля.

2 Численный анализ индукционной системы

Для анализа индукционной системы с холодным тиглем была разработана трехмерная математическая модель электромагнитного поля относительно векторного магнитного потенциала. Модель реализована в виде программы на языке APDL [4] в коммерческом пакете ANSYS v.11. [5]

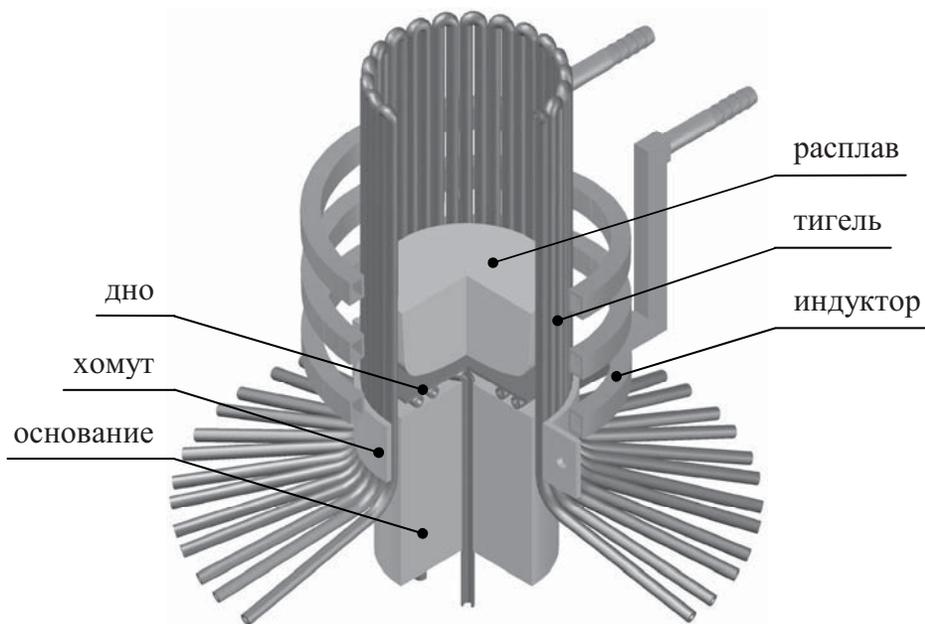


Рисунок 1 – Эскиз индукционной системы с холодным тиглем

В статье приводятся результаты исследований системы, параметры которой представлены в таблице 1. По сравнению с рисунком 1 в электромагнитную систему индукционной печи были внесены следующие допущения:

- в расчете не учитывались дно тигля и хомут;
- индуктор одновитковый;
- нижние торцы индуктора и загрузки располагаются на одном уровне;
- середина высоты индуктора совпадает с серединой высоты трубки тигля;
- секции тигля представлены в виде одиночных трубок;

Таблица 1 – Параметры исследуемой системы

Элемент системы	Удельное сопротивление, Ом·см	Радиус, см	Высота, см
Индуктор	2,0E-6	6,7 (внутр.)	8,0
Загрузка	5,13E-4 ÷ 15,1	3,4 (внеш.)	1,0; 2,0; 3,0; 4,0; 5,6; 7,0
Тигель	2,0E-6	3,5 (внутр.)	20,0

Исследовались энергетические характеристики индукционной системы при различной степени заполнения тигля. Было исследовано шесть вариантов высот загрузки от 1 до 8 см (см. таблицу 1). Расчеты проводились для различного удельного сопротивления загрузки в диапазоне степени проявления поверхностного эффекта в загрузке от 0,1 до 40.

Индукционная система с холодным тиглем характеризуется круговой симметрией, что учитывалось при построении твердотельной модели. На левой части рисунка 2 показан сектор

индукционной системы, охватывающий одну трубку холодного тигля и по половине воздушного зазора между трубками. Периодичность в системе задавалась граничными условиями. На рисунке 2 справа представлена оптимальная конечно-элементная сетка, которая содержала 650 тысяч элементов и 460 тысяч узлов.

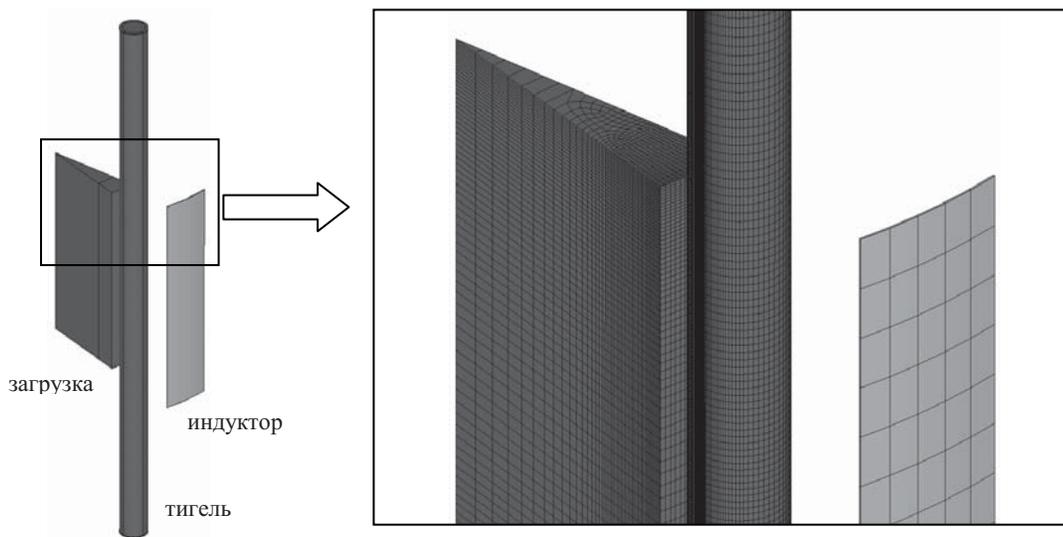


Рисунок 2 – Твердотельная модель и фрагмент конечно-элементной сетки индукционной системы

3 Результаты

В результате расчетов были получены мощности электрических потерь в индукторе, загрузке и трубках тигля. Степень загруженности печи характеризуется отношением высот загрузки и индуктора $H_{\text{загр}}/H_{\text{инд}}$. Степень проявления поверхностного эффекта в загрузке – отношением радиуса загрузки к глубине проникновения электромагнитного поля в загрузку $R_{\text{загр}}/\Delta_{\text{загр}}$.

По результатам расчетов построены зависимости, характеризующие потери в секциях холодного тигля. На рисунке 3 приведены зависимости потерь в секциях тигля для каждой из высот загрузки, приведенные к потерям в тигле без загрузки. Расчет проводился при одинаковом токе индуктора. Можно видеть, что в области $R_{\text{загр}}/\Delta_{\text{загр}} = 0,1 \div 5$ при увеличении степени проявления поверхностного эффекта в загрузке наблюдается резкое падение потерь в тигле. Скорость падения растет при увеличении высоты загрузки. При $R_{\text{загр}}/\Delta_{\text{загр}} > 5$ и максимальной загрузке влияние загрузки на потери в тигле минимально. Однако в области малых высот загрузки наблюдается тенденция к увеличению потерь. Минимум потерь в секциях тигля зависит от высоты загрузки и располагается в интервале $R_{\text{загр}}/\Delta_{\text{загр}} = 5 \div 10$. При уменьшении высоты загрузки минимум потерь сдвигается в область менее яркого поверхностного эффекта. При высоте загрузки меньше четверти высоты индуктора и ярком поверхностном эффекте в загрузке потери в тигле превышают потери в тигле без загрузки. При высоте загрузки близкой к высоте индуктора ($H_{\text{загр}}/H_{\text{инд}} = 0,88$) потери на 37% меньше потерь в пустом тигле.

Анализ представленных зависимостей позволил сделать вывод о наличии нескольких факторов, влияющих на величину потерь в секциях тигля.

- 1) Степень проявления поверхностного эффекта в загрузке. Начинает действовать при $R_{\text{загр}}/\Delta_{\text{загр}} = 1$ в виде резкого падения потерь в трубках тигля. При $R_{\text{загр}}/\Delta_{\text{загр}} < 1$ ток с внеш-

ней стороны трубки концентрируется под индуктором, а затем растекается по всей высоте трубки. При повышении $R_{\text{загр}}/\Delta_{\text{загр}}$ с внутренней стороны трубки ток концентрируется по высоте загрузки. При ярко выраженном поверхностном эффекте падение замедляется, т.к. ток уже максимально стянут к загрузке, поэтому наблюдается минимум потерь. При дальнейшем увеличении $R_{\text{загр}}/\Delta_{\text{загр}}$ ток, сконцентрированный на поверхности загрузки слабеет, а ток в трубке тигля снова начинает растекаться по высоте трубки, отсюда повышение потерь.

- 2) Плотность тока с внутренней стороны трубки зависит от высоты загрузки. Проявляется более слабым падением потерь при малых высотах загрузки. Перетекающий с внешней стороны трубки тигля ток концентрируется по высоте загрузки. Таким образом, чем меньше высота загрузки, тем выше плотность тока с внутренней стороны трубки и больше потери.

Сумма двух факторов определяет, при какой степени проявления поверхностного эффекта в загрузке будет наблюдаться минимум потерь в трубках тигля.

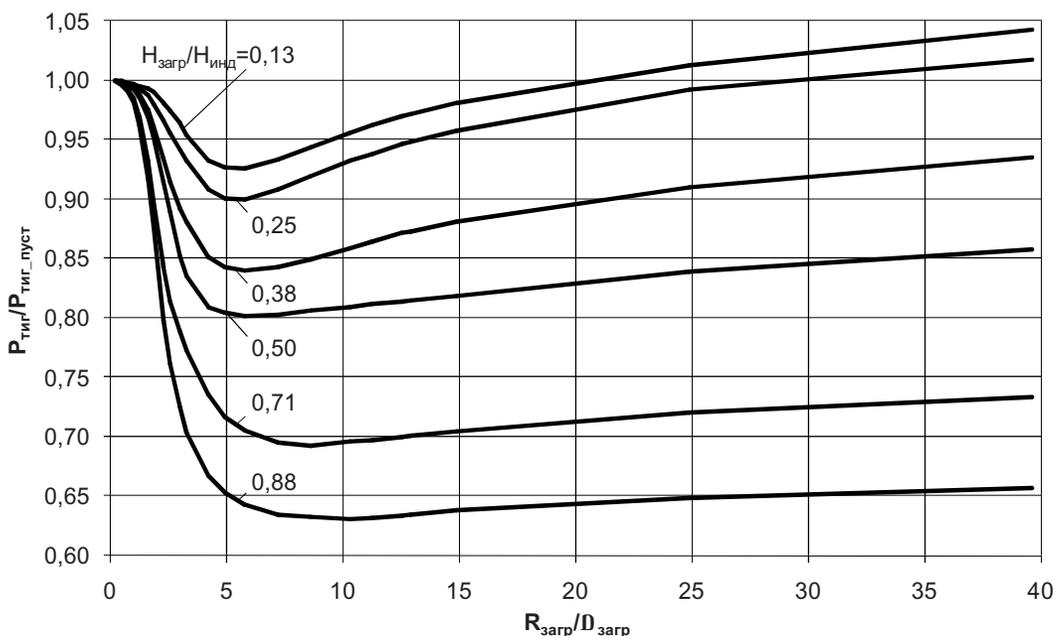


Рисунок 3 – Потери в холодном тигле, приведённые к потерям в тигле без загрузки

На рисунке 4 приведены зависимости электрического КПД индукционной системы с холодным тиглем, построенные по результатам расчетов. Их анализ показывает, что, в отличие от КПД индукционной системы без холодного тигля, КПД имеет максимум и уменьшается при повышении проявления поверхностного эффекта в загрузке. При неярком поверхностном эффекте КПД мал. В области с $R_{\text{загр}}/\Delta_{\text{загр}} = 2,5 \div 5$ достигается максимум КПД. При понижении высоты загрузки максимум сдвигается в область более яркого поверхностного эффекта. От высоты загрузки КПД зависит слабо. В зоне минимальных потерь в секциях тигля (при $R_{\text{загр}}/\Delta_{\text{загр}} = 5 \div 10$) КПД снижается от максимума на $2 \div 8\%$ в зависимости от высоты загрузки.

Таким образом, одним из критериев при проектировании индукционных систем с холодным тиглем является условие $R_{\text{загр}}/\Delta_{\text{загр}} = 5 \div 10$. В этом случае система имеет близкий к максимальному электрический КПД. В то же время потери в тигле близки к минимальным и слабо зависят от $R_{\text{загр}}/\Delta_{\text{загр}}$, следовательно, меньше всего влияют на режим индукционной системы.

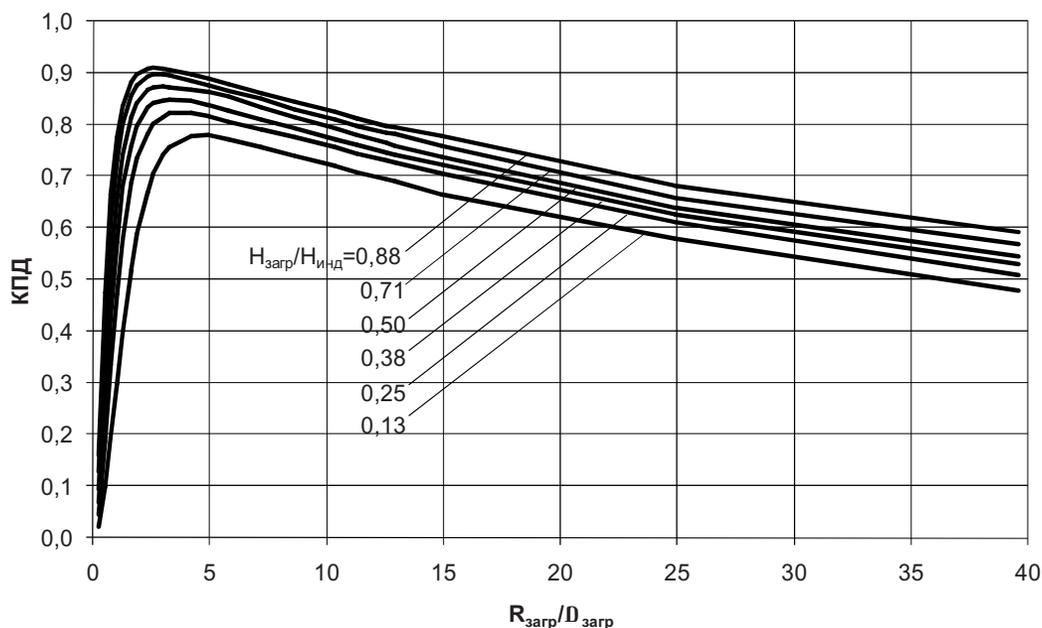


Рисунок 4 – Электрический КПД индукционной системы

Заключение

Потери в секциях холодного тигля значительно зависят от степени проявления поверхностного эффекта в загрузке и высоты загрузки. Соотношение мощности в загрузке и мощности в тигле часто не позволяет пренебречь выявленным явлением уменьшения потерь в тигле при наличии загрузки. Поэтому эту зависимость необходимо учитывать при проектировании индукционных установок с холодным тиглем.

Благодарности

Исследование было проведено при использовании оборудования и программного обеспечения, приобретенного в рамках Инновационного образовательного проекта Санкт-Петербургского государственного электротехнического университета "ЛЭТИ" – "Программа подготовки специалистов для приоритетных высокотехнологичных отраслей инновационной экономики страны" за 2009 год.

Список литературы

- [1] Петров, Ю.Б. Холодные тигли / Ю.Б.Петров, Д.Г.Ратников. – М.: Металлургия, 1972.
- [2] Петров, Ю.Б. Индукционные печи для плавки оксидов / Ю.Б.Петров, И.А.Канаев. – Л.: Политехника, 1991.
- [3] Кузьминов, Ю.С. Тугоплавкие материалы из холодного тигля / Ю.С.Кузьминов, Е.Е.Ломонова, В.В.Осико. – М.: Наука, 2004.
- [4] ANSYS parametric design language (APDL) guide. Release 11.0.
- [5] ANSYS Basic Analysis Guide. ANSYS Release 11.0, Canonsburg, ANSYS Inc. Houston, 2007.

ТЕХНОЛОГИЯ ПОЛУЧЕНИЯ МУЛЬТИКРИСТАЛЛИЧЕСКОГО КРЕМНИЯ НА ОСНОВЕ ИНДУКЦИОННОЙ ПЛАВКИ В ХОЛОДНОМ ТИГЛЕ

В.В. Кичигин, И.В. Позняк, А.Ю. Печенков, А.И. Максимов, А.Н. Шатунов

Санкт-Петербургский Государственный Электротехнический университет «ЛЭТИ»

197376, Санкт-Петербург, ул. Профессора Попова, 5, Россия

root@post.etu.spb.ru

тел: +7 (812) 346-44-87, факс: +7 (812) 346-27-58

Ключевые слова: мультикристаллический кремний, направленная кристаллизация, холодный тигель, индукционная печь, численный анализ, электромагнитное поле

Abstract

This paper includes research results of directional solidification technology of multicrystalline silicon in induction furnace with cold crucible for photovoltaic cells. Those investigations are carrying out at Institute of Electrotechnology (ETP) Hannover University together with Department of electrotechnology (ETPT) Saint-Petersburg electrotechnical university. For studying of starting heating and technological parameters of the melting and crystallization technology was designed and produced induction furnace with square section cold crucible. Dimension of the cold crucible external side is 14 cm. Induction system is working with transistor generator at 100 kHz. Crystallization process is realized at argon atmosphere. Thermal and some electrical parameters during tests from start to steady-state mode are shown too in the paper.

Введение

Одним из экологически чистых способов производства электроэнергии является преобразование солнечного света в электричество. Во многих странах действуют правительственные программы поддержки развития наземной фотоэнергетики. В частности, по программе Министерства энергетики США рост наземной фотоэнергетики за период с 2000 по 2020 год должен составить от 7 до 15 ГВт/год. В этом случае примерно 15% энергии, вырабатываемой для применения в жилых домах, будет получаться при помощи солнечных батарей. В настоящее время широкому использованию солнечных батарей препятствует высокая стоимость фотоэлектрических преобразователей (ФЭП). Наиболее перспективными для широкомасштабного наземного применения являются кремниевые ФЭП. Коэффициент полезного действия промышленно выпускаемых кремниевых ФЭП находится на уровне 14 - 17,5%, а ресурс работы достигает 10 лет. В структуре себестоимости производимых в настоящее время кремниевых ФЭП около 50% составляет цена исходного материала - монокристаллического кремния электронного качества. Поэтому существует актуальная задача уменьшения стоимости кремния для изготовления фотоэлектрических преобразователей для наземной энергетики.

Полный цикл производства монокристаллического кремния включает в себя следующие основные этапы: получение, металлургического кремния карботермическим восстановлением минерального кварцевого сырья в мощных электродуговых печах, хлорирование металлургического кремния и глубокая очистка хлорсиланов в ректификационных установках, водородное восстановление в реакторах осаждения, выращивание монокристаллов кремния в ростовых установках. При этом стоимость килограмма монокристаллического кремния с содержанием примесей менее 10^{-6} масс. % в 60 – 120 раз превышает стоимость килограмма металлургического кремния с содержанием примесей около 0,5 масс. %.

Столь высокая чистота кремния является излишней, так как для изготовления высокоэффективных ФЭП наземного применения кремний может содержать электрически активных

примесей в количестве до $10^{-3} - 10^{-4}$ масс. %. Кремний такой степени чистоты называется солнечным кремнием.

В настоящее время для снижения стоимости изготовления фотоэлектрических преобразователей при их изготовлении используются отходы полупроводникового кремния, образующиеся при производстве изделий микроэлектроники. Однако, ожидается, что в ближайшее время потребности в кремнии для фотоэлектрических преобразователей превысят то количество материала, которое может быть получено из отходов полупроводникового производства, в 2-4 раза. Это связано с тем, что рост промышленного производства солнечных элементов значительно опережает темпы роста производства кремния для микроэлектроники. Переход на сырье солнечного качества позволяет снизить себестоимость ФЭП.

В настоящее время в мире признано, что для производства слитков кремния для ФЭП наземного применения наиболее экономически выгодной является плавка поликристаллического кремния солнечного качества с последующей направленной кристаллизацией расплава. Продуктами подобного передела кремния являются слитки квадратного или прямоугольного сечения со столбчатой структурой. Размеры столбчатых кристаллов значительно превышают размеры поликристаллов, образующихся при спонтанной кристаллизации расплава, поэтому подобная структура получила название мульткристаллической. Чем больше будут размеры бездефектных мульткристаллов кремния в слитке, тем выше будет в последующем коэффициент полезного действия ФЭП.

Мульткристаллические блоки кремния получают по методу Бриджмена-Стокбаргера, либо в одной и той же печи, где осуществляется и плавка кремния и направленная кристаллизация расплава перемещением его через зоны роста кристаллов, либо сначала плавку проводят в отдельной печи, а затем разливают расплав в кварцевый тигель, который помещается во вторую печь, где и осуществляется направленная кристаллизация расплава. Масса загрузки не превышает 420 кг расплава. И в том и другом случае недостатками являются необходимость в использовании одноразовых тиглей из высокочистого кварца, стоимость которых нельзя не принимать во внимание при расчете себестоимости мульткристаллического слитка кремния и периодичность процессов.

Наиболее радикальным путем преодоления этих недостатков является технологическая схема процесса получения слитков кремния солнечного качества, совмещающая незагрязняющий индукционный метод нагрева с незагрязняющим способом плавки в гарнисаже и непрерывно-последовательной кристаллизацией расплава в слиток квадратного или прямоугольного поперечного сечения характеризующегося крупноблочной мульткристаллической структурой.

Таблица 1 – Техничко-экономические характеристики различных методов получения мульткристаллического кремния

Метод	Сечение слитка	Скорость кристаллизации	Эффективность ФЭП	Производительность по площади ФЭП	Количество кремния на 1 Вт энергии
	см	мм/мин	%	м ² /день	грамм
Кристаллизация в кварцевом тигле	80 × 80	0,1 – 0,6	15	140	7
Кристаллизация в индукционной печи с холодным тиглем	35 × 35	1,2 – 1,5	15	500	7

На сегодняшний день только две фирмы освоили промышленную технологию направленной непрерывно-последовательной кристаллизации солнечного кремния в индукционной печи с холодным тиглем: Sumko Solar Corporation (Япония) [1] и Emix (Франция) [2]. Преимущество технологии непрерывно-последовательной направленной кристаллизации кремния в индукци-

онной печи с холодным тиглем перед направленной кристаллизацией в кварцевых тиглях становится ясной из сравнения технико-экономических характеристик, приведенных в таблице 1. Технологическая схема выращивания мультикристаллического кремния методом направленной кристаллизации с использованием индукционной печи с холодным тиглем показана на рисунке 1.

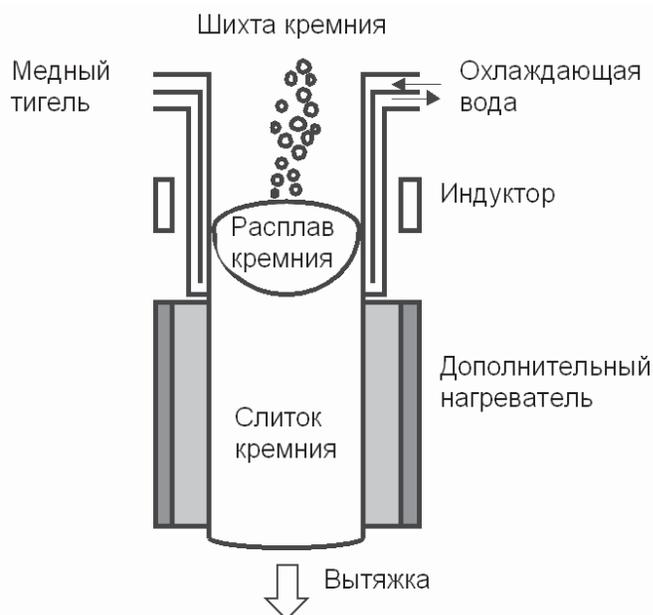


Рисунок 1 – Технологическая схема выращивания мультикристаллического кремния методом направленной кристаллизации с использованием индукционной печи с холодным тиглем

Результаты исследований

Целью данной работы было исследование изменения энергетических характеристик индукционной печи с холодным тиглем от стартового нагрева шихты кремния до выхода на стационарный режим плавки. Плавка производилась в индукционной системе на установке Института Электронагрева г. Ганновер, работающей на частоте около 100 кГц. Старт осуществлялся на графитовом основании, которое нагревало шихту кремния до температуры плавления. Для определения тепловых потоков расплав удерживался в стационарном состоянии при температуре 1660 °С, при этом измерялись электрические потери в тигле и в дне. На рисунке 2 представлены изменения электрических потерь в тигле, в дне и индукторе во время эксперимента.

Для оценки электрического КПД и сравнения с энергетическими параметрами существующей установки, работающей на частотах 30 кГц [3], была оценена мощность в расплаве на основе измеренных электрических потерь в стационарном режиме и сразу после выключения источника питания (рисунок 3). Тепловой поток от расплава в стенку тигля был определен через отношение измеренной мощности в расплаве к площади боковой поверхности полученного слитка. Тепловой поток от расплава в стенку холодного тигля составил 60 Вт/см².

На рисунке 4 показаны распределения электрических мощностей между компонентами индукционной системы при плавке кремния в холодном тигле для установки компании SUMCO SOLAR Corporation [1] и установки, работающей на частоте 100 кГц.

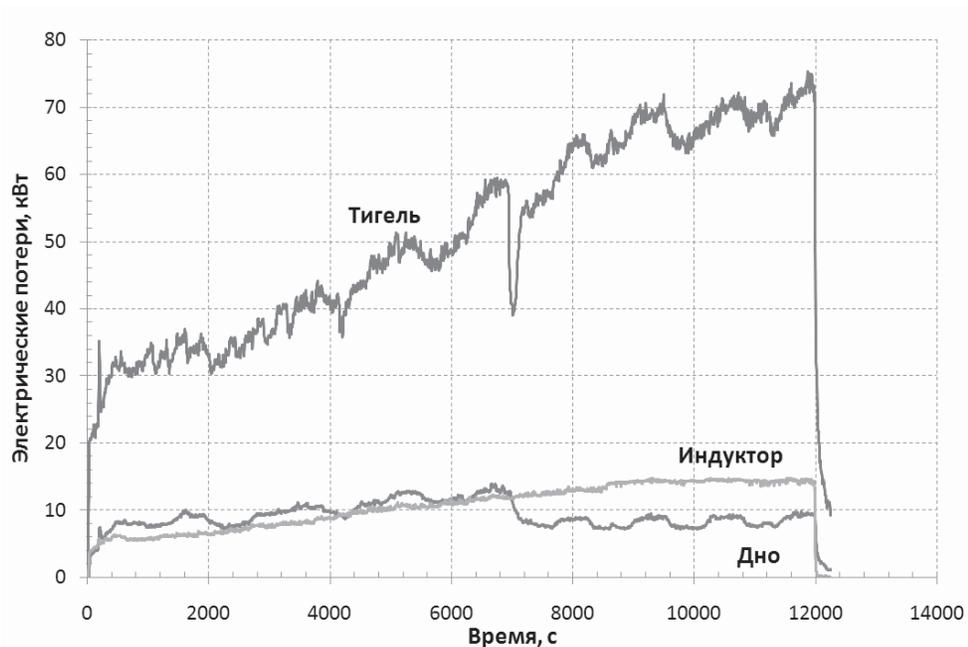


Рисунок 2 – Электрические потери индукционной системы

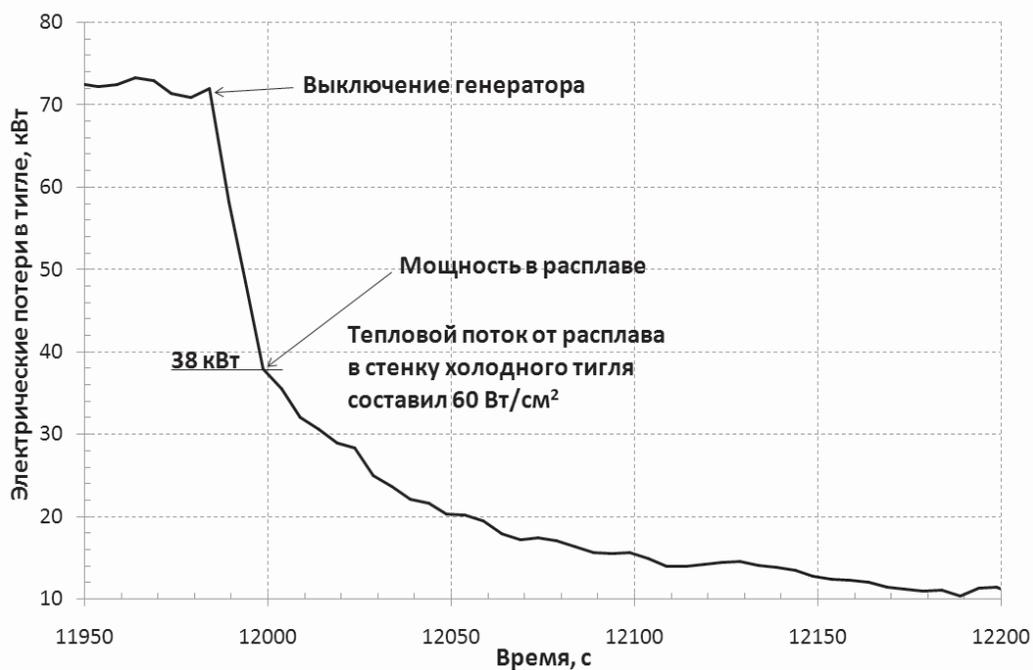


Рисунок 3 – Электрические потери в тигле после выключения генератора

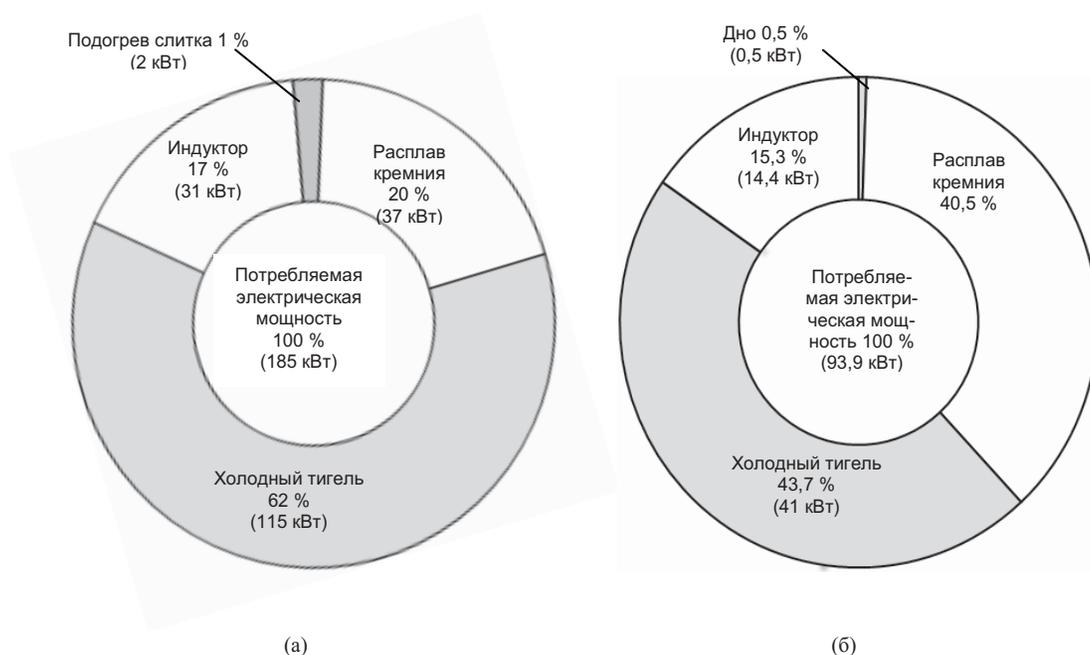


Рисунок 4 – Распределение электрической мощности при плавке кремния в холодном тигле для установки SUMCO SOLAR company (а) и для установки, работающей на частоте 100 кГц (б)

Заключение

Полученные результаты на основе проведенных тестов по плавке кремния с целью получения слитка с мультикристаллической структурой в индукционной печи с холодным тиглем показывают, что плавка на более высоких частотах по сравнению с данными существующих установок, работающими на 20-30 кГц, имеет большую эффективность.

Благодарности

Эксперименты были проведены при использовании оборудования и программного обеспечения Института Электронагрева г. Ганновер.

Список литературы

- [1] Kaneko, K., Nakagawa, K., Tomohiro Onizuka, Ken-ichi Sasatani, Nobuyuki Kubo, Michio Kida. Recent Status on Electromagnetic Casting. SUMCO SOLAR Corporation, 260-100 Funoo, Kainan, Wakayama, 642-0001, Japan, 2-7.
- [2] Nonstop shop. Emix produces silicon ingots using a new crystallization process. PHOTON International. 2006, #12, p. 32.
- [3] Kaneko, K., Kawamura, R. and Misawa, T. Present status and future prospects of electro-magnetic casting for silicon solar cells. Reserch & Development Center, Sumitomo SiTix Corporation Higashima-cho 1, Amagasaki, Hyogo, 660, Japan, first WCPEC; Dec. 5-9, 1994; Hawaii.

ЭЛЕМЕНТЫ СИНЕРГЕТИЧЕСКОЙ ТЕОРИИ УПРАВЛЕНИЯ ДИНАМИЧЕСКИМИ СИСТЕМАМИ С РАСПРЕДЕЛЕННЫМИ ПАРАМЕТРАМИ

Э.Я. Рапопорт

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
cscmp@iccs.ru
тел: +7 (846) 332-39-27, факс: +7 (846) 333-27-70

Ключевые слова: объекты с распределенными параметрами, синергетическая теория управления, макропеременная, притягивающие многообразия, агрегированные регуляторы

Abstract

Methods of synergetic control theory are used for analytical design of nonlinear aggregated regulators of distributed parameter systems with different controlling actions.

Введение

Одним из перспективных направлений на путях создания современной физической теории управления [1] является синергетический подход [2, 3], реализующий концепцию целевой самоорганизации синтезируемых систем и основанный на создании в их структуре желаемых инвариантных притягивающих многообразий (искусственных аттракторов), на которых естественные свойства объекта максимально согласуются с целями задачи управления.

Уравнения инвариантных многообразий описываются в терминах агрегированных макропеременных, являющихся аналогом главных степеней свободы («параметров порядка»), выделяемых в процессах самоорганизации природных динамических систем [4]. Соответствующий выбор макропеременных в форме некоторых функций фазовых координат, определенных на уравнениях модели объекта, позволяет решить задачу аналитического конструирования агрегированных регуляторов (АКАР) [2, 3], реализующих двухэтапную организацию процессов управления – близкие к оптимальным по переменным в фазовом пространстве критериям качества, асимптотически устойчивые движения модели объекта сначала к притягивающему многообразию, а затем вдоль него к равновесному состоянию системы. В настоящей работе исследуются возможности распространения синергетического подхода на представляющие самостоятельный интерес задачи синтеза бесконечномерных систем управления объектами с распределенными параметрами (ОРП) [5-7].

1 Типовые модели ОРП

В достаточно общем случае типовая модель ОРП может быть представлена нелинейным пространственно-одномерным уравнением в частных производных параболического типа

$$(1) \quad A(Q) \frac{\partial Q}{\partial t} = C(Q) \frac{\partial^2 Q}{\partial x^2} + B(Q) \frac{\partial Q}{\partial x} + B_1(Q) \left(\frac{\partial Q}{\partial x} \right)^2 + D(Q) Q + f^*(x, t, Q, U(x, t))$$
$$x_0 < x < x_1, t > 0$$

относительно управляемой функции состояния $Q(x, t)$, изменяющейся в зависимости от пространственной координаты $x \in [x_0, x_1]$ и времени t , с начальными и граничными условиями

$$(2) \quad Q(x, 0) = Q_0(x), \quad x \in [x_0, x_1];$$

$$(3) \quad \alpha_1(Q(x_1, t))Q(x_1, t) + \alpha_2(Q(x_1, t)) \frac{\partial Q(x_1, t)}{\partial x} = f_1^*(t, Q(x_1, t), U_1(t));$$

$$(4) \quad \frac{\partial Q(x_0, t)}{\partial x} = 0,$$

внутренним распределенным или(и) граничным сосредоточенным управляющими воздействиями, соответственно $U(x, t)$ и $U_1(t)$, при заданных достаточно гладких функциях f^* , f_1^* и зависимостях от Q коэффициентов $A, B, B_1, C, D, \alpha_1$ и α_2 [5-9].

В частности, к виду (1) приводится фундаментальное уравнение теплопроводности и диффузии в декартовых и цилиндрических координатах

$$c\rho \left(\frac{\partial Q}{\partial t} + V \operatorname{grad} Q \right) = \operatorname{div}(\lambda \operatorname{grad} Q) + f^*$$

в движущейся со скоростью $V \geq 0$ среде с заданными характеристиками c, ρ, λ [9].

В линейном приближении при $A, B, C, D, \alpha_1, \alpha_2 = \text{const}$, $B_1 = 0$;

$$f^*(x, t, Q(x, t), U(x, t)) \equiv U(x, t); \quad f_1^*(x, t, Q(x_1, t), U_1(t)) \equiv U_1(t)$$

применение к уравнениям объекта (1) – (4) конечного интегрального преобразования по пространственному аргументу с ядром, равным его собственным функциям $\varphi_n(\mu_n, x)$, $n = 1, 2, \dots$, и весовой функцией $r(x)$, где μ_n^2 – собственные числа, приводит к описанию ОРП бесконечной линейной системой обыкновенных дифференциальных уравнений

$$(5) \quad A \frac{d\bar{Q}_n(\mu_n, t)}{dt} = -\mu_n^2 \bar{Q}_n(\mu_n, t) + \bar{U}_n(\mu_n, t) + d_n U_1(t), \quad n = 1, 2, \dots$$

$$\bar{Q}_n(\mu_n, 0) = \bar{Q}_0(\mu_n)$$

для временных мод $\bar{Q}_n(\mu_n, t)$ разложения $Q(x, t)$ в сходящийся в среднем бесконечный ряд по ортонормированной системе $\varphi_n(\mu_n, x)$ [8, 10]:

$$(6) \quad Q(x, t) = \sum_{n=1}^{\infty} \bar{Q}_n(\mu_n, t) \varphi_n(\mu_n, x); \quad \bar{Q}_n(\mu_n, t) = \int_{x_0}^{x_1} Q(x, t) \varphi_n(\mu_n, x) r(x) dx.$$

Здесь d_n – известные постоянные коэффициенты, $\bar{Q}_0(\mu_n)$ – заданные значения $\bar{Q}_n(\mu_n, 0)$, и $\bar{U}_n(\mu_n, t)$ – временные моды разложения распределенного управления $U(x, t)$ в ряд вида (6):

$$(7) \quad U(x, t) = \sum_{n=1}^{\infty} \bar{U}_n(\mu_n, t) \varphi_n(\mu_n, x); \quad \bar{U}_n(\mu_n, t) = \int_{x_0}^{x_1} U(x, t) \varphi_n(\mu_n, x) r(x) dx.$$

Аналогичным образом можно получить точное описание в подобной форме нелинейной модели (1) – (4), только теперь уже бесконечной системой нелинейных дифференциальных уравнений, путем разложения $Q(x, t)$ в бесконечный ряд по физически родственной системе собственных функций $\varphi_n(\mu_n, x)$ ее линейного приближения.

Запишем краевую задачу (1) – (4) в следующем виде:

$$(8) \quad A_0 \frac{\partial Q}{\partial t} = C_0 \frac{\partial^2 Q}{\partial x^2} + B_0 \frac{\partial Q}{\partial x} + D_0 Q + F(x, t), \quad x_0 < x < x_1, \quad t > 0;$$

$$(9) \quad \alpha_{10}Q(x_1, t) + \alpha_{20} \frac{\partial Q(x_1, t)}{\partial x} = F_1(x_1, t), \quad t > 0;$$

$$\frac{\partial Q(x_0, t)}{\partial x} = 0; \quad Q(x, 0) = Q_0(x), \quad x \in [x_0, x_1],$$

где

$$(10) \quad F(x, t) = f^*(x, t, Q(x, t), U(x, t)) + (C(Q) - C_0) \frac{\partial^2 Q(x, t)}{\partial x^2} +$$

$$+ (B(Q) - B_0) \frac{\partial Q(x, t)}{\partial x} + B_1(Q) \left(\frac{\partial Q(x, t)}{\partial x} \right)^2 + (D(Q) - D_0) Q(x, t) -$$

$$- (A(Q) - A_0) \frac{\partial Q(x, t)}{\partial t};$$

$$(11) \quad F_1(t) = f_1^*(t, Q(x_1, t), U_1) - (\alpha_1(Q(x_1, t)) - \alpha_{10}) Q(x_1, t) -$$

$$- (\alpha_2(Q(x_1, t)) - \alpha_{20}) \frac{\partial Q(x_1, t)}{\partial x},$$

и $A_0, B_0, C_0, D_0, \alpha_{10}, \alpha_{20}$ - некоторые (например, полученные путем усреднения коэффициентов в (1) - (4)) константы.

Для любых кусочно-непрерывных функций F и F_1 , если считать их заданными в (10), (11), модель ОРП (8), (9) становится линейной и описывается, подобно (5), в модальном представлении бесконечной системой уравнений

$$(12) \quad A_0 \frac{d\bar{Q}_n(\mu_n, t)}{dt} = -\mu_n^2 \bar{Q}_n(\mu_n, t) + \bar{F}_n(\mu_n, t) + d_n F_1(t), \quad n = 1, 2, \dots;$$

$$\bar{Q}_n(\mu_n, 0) = \bar{Q}_0(\mu_n).$$

Здесь $\bar{F}_n(\mu_n, t)$ - моды разложения $F(x, t)$ в ряд по $\varphi_n(\mu_n, x)$:

$$(13) \quad \bar{F}_n(\mu_n, t) = \int_{x_0}^{x_1} F(x, t) \varphi_n(\mu_n, x) r(x) dx, \quad n = 1, 2, \dots$$

Подставляя $Q(x, t)$ в виде (6) в выражения (10), (11) для $F(x, t)$ и $F_1(t)$ и интегрируя в (13), получим $\bar{F}_n(\mu_n, t)$ и $F_1(t)$ в форме вполне определенных нелинейных зависимостей от $\bar{Q} = (\bar{Q}_n)$, $n = 1, 2, \dots$, и мод \bar{f}_n^* функции f^* , заданных своими разложениями в сходящиеся бесконечные ряды, что в результате приводит к преобразованию соотношений (12) в замкнутую нелинейную систему уравнений ОРП. В частности, в типичном случае $f^*(x, t, Q(x, t), U(x, t)) \equiv U(x, t)$; $f_1^*(t, Q(x_1, t), U_1(t)) \equiv U_1(t)$ эта система

$$(14) \quad A_0 \frac{d\bar{Q}_n(\mu_n, t)}{dt} = -\mu_n^2 \bar{Q}_n(\mu_n, t) + \Phi_n(\bar{Q}) + \bar{U}_n(\mu_n, t) + d_n U_1(t);$$

$$\bar{Q}_n(\mu_n, 0) = \bar{Q}_0(\mu_n), \quad n = 1, 2, \dots$$

отличается от линейной модели (5) лишь заданными указанным образом нелинейными, непрерывно дифференцируемыми по своим аргументам функциями $\Phi_n(\bar{Q})$

$$(15) \quad \Phi_n(\bar{Q}) = \bar{F}_n - \bar{U}_n + (F_1 - U_1) d_n, \quad n = 1, 2, \dots$$

В отклонениях $z_n = \bar{Q}_n - \bar{Q}_n^*$; $u(x,t) = U(x,t) - U^*(x,t)$, $\bar{u}_n = \bar{U}_n - \bar{U}_n^*$; $u_1 = U_1 - U_1^*$ от равновесного состояния $Q^*(x,t)$; $U^*(x,t)$; \bar{Q}_n^* , \bar{U}_n^* , $n = 1, 2, \dots$; U_1^* в установившемся режиме работы ОРП уравнения возмущенного движения объекта принимают теперь, вместо (14), следующий вид:

$$(16) \quad A_0 \frac{dz_n}{dt} = -\mu_n^2 z_n + f_n(\mathbf{z}) + \bar{u}_n(\mu_n, t) + d_n u_1(t), \quad n = 1, 2, \dots;$$

$$z_n(0) = z_n^0,$$

где $\mathbf{z} = (z_n)$, $n = 1, 2, \dots$; $f_n(\mathbf{z}) = \Phi_n(\bar{\mathbf{Q}}^* + \mathbf{z}) - \Phi_n(\bar{\mathbf{Q}}^*)$, $f_n(0) = 0$; $z_n^0 = \bar{Q}_n(\mu_n, 0) - \bar{Q}_n^*(\mu_n, 0)$, и при $f_n(\mathbf{z}) = 0$ уравнения (16) соответствуют линейной модели (5).

Принципиальной особенностью системы (16) по сравнению с аналогичным описанием объектов с сосредоточенными параметрами является ее бесконечный порядок. Можно показать [11, 12], что при обычно выполняющихся в прикладных задачах (в виду сходимости рядов в (6) и (7)) формальных требованиях к поведению функций $f_n(\mathbf{z})$ и алгоритмов управления $\bar{u}_n(\mathbf{z}), u_1(\mathbf{z})$ с обратными связями, в частности, при выполнении усиленных условий Коши-Липшица система (16) имеет единственное решение, которое с любой требуемой точностью аппроксимируется решением «укороченной» системы, образуемой достаточно большим конечным числом N первых из уравнений (16) при $z_n = 0 \quad \forall n > N$, т.е. оказывается допустимой конечномерная аппроксимация модели (16) при $n = \overline{1, N}$, $N < \infty$.

2 Задача АКАР для распределенных объектов с граничными управляющими воздействиями

Система уравнений модели ОРП с граничным управляющим воздействием $u_1(t)$, подчиненным заданному ограничению

$$(17) \quad u_{1\min} \leq u_1(t) \leq u_{1\max}; \quad u_{1\min}, u_{1\max} = const,$$

принимает вид (16) при $\bar{u}_n(\mu_n, t) = 0$.

В целях учета ограничений (17) дополним структуру объекта интегратором [3]

$$(18) \quad T_0 \frac{dz_0}{dt} = u_0(t), \quad T_0 = const$$

с условным управлением $u_0(t)$ на его входе и достаточно малой постоянной времени T_0 , представляя при этом $u_1(t)$ следующей типичной нелинейной ограниченной функцией дополнительной фазовой координаты z_0 [3], удовлетворяющей для всех $\eta = const > 0$ неравенствам (17) при любых значениях z_0 :

$$(19) \quad u_1 = L_1 + L_2 \text{th}(\eta z_0), \quad L_1 = \frac{u_{1\max} + u_{1\min}}{2}; \quad L_2 = \frac{u_{1\max} - u_{1\min}}{2}.$$

В результате ОРП описывается в расширенном фазовом пространстве (\mathbf{z}, z_0) системой уравнений

$$(20) \quad T_0 \frac{dz_0}{dt} = u_0;$$

$$A_0 \frac{dz_n}{dt} = -\mu_n^2 z_n + f_n(\mathbf{z}) + d_n (L_1 + L_2 \text{th}(\eta z_0)).$$

Согласно методологии синергетического подхода [2, 3], требуется найти алгоритм управления с обратной связью $u_1(\mathbf{z}, z_0)$, обеспечивающий перевод объекта (20) из произвольного начального состояния сначала в окрестность притягивающего многообразия

$$(21) \quad \psi(\mathbf{z}, z_0) = 0, ,$$

определяемого априори фиксируемой формой зависимости агрегированной макропеременной ψ от фазовых координат, а затем последующее асимптотически устойчивое движение вдоль этого многообразия к точке равновесия $z = 0$ с требуемыми качественными показателями переходного процесса.

Выбор траектории движения макропеременной к многообразию (21) в виде решений дифференциального уравнения

$$(22) \quad T \frac{d\psi}{dt} + R(\psi) = 0, \quad T = const,$$

принадлежащих подсемейству устойчивых экстремалей, минимизирующих сопровождающий квадратичный функционал качества

$$(23) \quad I = \int_0^{\infty} \left[R^2(\psi) + T^2 \left(\frac{d\psi}{dt} \right)^2 \right] dt,$$

обеспечивает асимптотическую устойчивость в целом этого движения при любых $T > 0$ для произвольных непрерывных и дифференцируемых функций $R(\psi)$ в (22), удовлетворяющих условиям $R(0) = 0, \psi R(\psi) > 0 \forall \psi \neq 0$ [2, 3].

Задача обеспечения асимптотической устойчивости движения вдоль многообразия (21) должна быть решена в процессе аналитического конструирования агрегированных регуляторов.

Вычисляя полную производную $\frac{d\psi}{dt}$ на уравнениях объекта (20) и подставляя результаты дифференцирования в (22), получим основное функциональное уравнение

$$(24) \quad T \left\{ \frac{\partial \psi}{\partial z_0} \frac{1}{T_0} u_0 + \frac{1}{A_0} \sum_{n=1}^{\infty} \frac{\partial \psi}{\partial z_n} \left[-\mu_n^2 z_n + f_n(\mathbf{z}) + d_n(L_1 + L_2 \text{th}(\eta z_0)) \right] \right\} + R(\psi) = 0$$

для отыскания условного входного воздействия $u_0(\mathbf{z}, z_0)$, определяющего в силу (18), (19) искомый алгоритм $u_1(\mathbf{z}, z_0)$, на котором притягивающее многообразие (21) становится инвариантным [2, 3].

Центральным моментом при реализации данного подхода является выбор конкретных форм макропеременной, для которых многообразие (21) можно рассматривать в роли желаемого целевого множества, наилучшим образом согласующего естественные свойства объекта управления и технические требования к синтезируемым системам.

При этом изначально задаваемая неоднозначным образом подынтегральная функция критерия (23) может обеспечить (за счет ее уточнения путем надлежащего выбора зависимостей ψ от \mathbf{z} и z_0) переменный в фазовом пространстве характер сопровождающего функционала качества, в соответствии с типовыми многокритериальными показателями эффективности функционирования реальных объектов управления в конкретных прикладных задачах АКАР [2, 3].

Полагая $R(\psi) = \psi$ в (22), выберем следующую агрегированную макропеременную в (21) – (24):

$$(25) \quad \psi(\mathbf{z}, z_0) = \text{th} \left[L_1 + L_2 \text{th}(\eta z_0) + \sum_{n=1}^{\infty} \beta_n z_n \right]$$

с априори неизвестными параметрами $\beta_n = \text{const}$, определяемыми далее, в частности, из условий асимптотически устойчивого движения объекта вдоль многообразия (21) к точке $\mathbf{z} = 0$ и обеспечивающими при этом сходимость ряда в (25).

В таком случае получаем на основании (24) функциональное уравнение

$$(26) \quad \frac{T}{\text{ch}^2 P} \left\{ \frac{L_2 \eta}{T_0 \text{ch}^2(\eta z_0)} u_0 + \frac{1}{A_0} \sum_{n=1}^{\infty} \beta_n [-\mu_n^2 z_n + f_n(\mathbf{z}) + d_n(L_1 + L_2 \text{th}(\eta z_0))] \right\} + \text{th} P = 0,$$

$$P = L_1 + L_2 \text{th}(\alpha z_0) + \sum_{n=1}^{\infty} \beta_n z_n,$$

разрешая которое относительно u_0 , найдем алгоритм для условного управления

$$(27) \quad u_0 = -T_0 \frac{\text{ch}^2(\eta z_0)}{\eta L_2} \left\{ \frac{1}{2T} \text{sh}(2P) + \frac{1}{A_0} \sum_{n=1}^{\infty} \beta_n [-\mu_n^2 z_n + f_n(\mathbf{z}) + d_n(L_1 + L_2 \text{th}(\eta z_0))] \right\},$$

определяющий с учетом соотношений (18), (19) искомую структуру регулятора $u_1(\mathbf{z}, z_0)$ в рассматриваемой задаче АКАР с полным измерением состояния объекта (20), где здесь и далее можно, согласно сказанному в п.1, ограничиться учетом достаточно большого, но конечного числа N мод $z_n, n = \overline{1, N}$.

Выбор макропеременной в форме (25) определяет при этом характер процесса управления, близкий к оптимальному по переменным в фазовом пространстве наиболее представительным для прикладных задач критериям качества.

Действительно, на достаточно большом удалении от притягивающего многообразия, где

$$(28) \quad P \neq 0; |\psi| = |\text{th} P| \approx 1; \frac{d\psi}{dt} \approx 0,$$

сопровождающий функционал (23) с макропеременной (25) в условиях $R(\psi) = \psi$ на соответствующем временном интервале практически превращается в центральный критерий быстрого действия, отвечая типичным требованиям минимизации времени процесса перевода объекта в равновесное состояние. При этом управляющее воздействие u_1 принимает, согласно (18), (19), только свои предельно допустимые значения $u_{1\max}$ или $u_{1\min}$. В окрестности притягивающего многообразия при малых P , где $\text{th} P \approx P$, движение оказывается оптимальным по стандартному квадратичному критерию качества, опять описываемому в этой области фазового пространства функционалом (23).

При движении вдоль многообразия (21), где $\text{th} P = 0 \Rightarrow P = 0$ при достаточно малых отклонениях от установившегося режима получаем, согласно (21), (25), (19), независимо от вида функций $f_n(\mathbf{z})$ в уравнениях объекта линейный закон управления:

$$(29) \quad u_{1\psi}(\mathbf{z}) = - \sum_{m=1}^{\infty} \beta_m z_m.$$

В итоге, приходим с весьма малой погрешностью к значительно более простому по сравнению с (27), реализуемому простыми инженерными средствами алгоритму управления на всем протяжении переходного процесса:

$$(30) \quad u_1(\mathbf{z}) = \begin{cases} -\sum_{m=1}^{\infty} \beta_m z_m, & u_{1\min} \leq -\sum_{m=1}^{\infty} \beta_m z_m \leq u_{1\max}; \\ u_{1\max}, & -\sum_{m=1}^{\infty} \beta_m z_m \geq u_{1\max}; \\ u_{1\min}, & -\sum_{m=1}^{\infty} \beta_m z_m \leq u_{1\min}, \end{cases}$$

с ограниченным выходом пропорционального регулятора (29), аналогичному приближенному результату в классической линейно-квадратичной задаче динамической оптимизации с ограничением на управляющее воздействие [13].

Движение замкнутой системы вдоль притягивающего многообразия в окрестности начала координат $\mathbf{z} = 0$ описывается уравнениями модели объекта (16) при $\bar{u}_n = 0$ с регулятором (29):

$$(31) \quad \begin{aligned} A_0 \frac{dz_n}{dt} &= -\mu_n^2 z_n + f_n(\mathbf{z}) - d_n \sum_{m=1}^{\infty} \beta_m z_m, \quad n = 1, 2, \dots \\ z_n(0) &= z_n^0. \end{aligned}$$

Полагая здесь априори неизвестным вектор $\boldsymbol{\beta}^N = (\beta_m), m = \overline{1, N}$, первых N коэффициентов обратных связей и выбирая остальные параметры $\beta_m, m > N$, в зависимости от $\boldsymbol{\beta}^N$ из условий сходимости ряда в (29) (например, полагая $\beta_m = 0, \forall m > N$), получаем отмеченную в п.1 возможность ограничиться оценкой поведения системы (31) по достаточно большому, но конечному числу N первых ее уравнений.

Для линейных конечномерных, управляемых относительно $\mathbf{z}^N = 0, \mathbf{z}^N = (z_1, \dots, z_N)$, систем вида (31) при $f_n(\mathbf{z}^N) = 0$ всегда можно найти вектор $\boldsymbol{\beta}^N$, при котором достигается заданное расположение корней характеристического полинома с отрицательной вещественной частью и, в частности, обеспечивается асимптотически устойчивое движение к точке $\mathbf{z}^N = 0$ [14, 15], и следовательно, достигаются требуемые качественные показатели процесса перевода линейного бесконечномерного объекта (31) в равновесное состояние $\mathbf{z} = 0$ при $f_n(\mathbf{z}) = 0 \quad \forall n = 1, 2, \dots$

Аналогичные выводы можно сделать применительно к линейному приближению нелинейной модели (31) при $f_n(\mathbf{z}) \neq 0$:

$$(32) \quad A_0 \frac{dz_n}{dt} = -\mu_n^2 z_n + \sum_{m=1}^{\infty} \left(\frac{\partial f_n(\mathbf{z})}{\partial z_m} \right)_{\mathbf{z}=0} z_m - d_n \sum_{m=1}^{\infty} \beta_m z_m, \quad n = 1, 2, \dots;$$

что гарантирует, согласно критерию Ляпунова [16], асимптотическую устойчивость положения равновесия нелинейной системы (31).

Реализация алгоритма управления (30) требует использования в замкнутом контуре соответствующих наблюдателей состояния вектора \mathbf{z} или \mathbf{z}^N . В простейшем частном варианте

$$(33) \quad \beta_m = K \varphi_m(\mu_m, x_c), x_c \in [x_0, x_1], \quad m = 1, 2, \dots$$

будем иметь в (29), согласно (6):

$$(34) \quad u_{1\psi}(\mathbf{z}) = -K \sum_{m=1}^{\infty} z_m(t) \varphi_m(\mu_m, x_c) = -K [Q(x_c, t) - Q^*(x_c)].$$

В данном случае алгоритм управления (30) сводится, согласно (34), к построению соответствующей системы автоматической стабилизации выхода объекта $Q(x_c, t)$, непосредственно измеряемого в некоторой точке x_c области пространственного распределения управляемой функции состояния $Q(x, t)$, если коэффициент K передачи пропорционального регулятора обеспечивает требуемые показатели качества управляемого процесса. Во многих типичных ситуациях такой выбор K становится возможным [6], однако, решение этого вопроса в каждой конкретной задаче требует специального анализа.

3 Задача АКАР для распределенных объектов с распределенными управляющими воздействиями

Система уравнений модели ОРП с внутренним управлением $u(x, t)$ принимает вид (16) при $u_1(t) = 0$ с подлежащими определению автономными модальными управлениями $\bar{u}_n(\mathbf{z})$ для каждой моды z_n , удовлетворяющими ограничениям

$$(35) \quad \bar{u}_{n \min} \leq \bar{u}_n \leq \bar{u}_{n \max}, \quad n = 1, 2, \dots,$$

где все предельно допустимые значения $\bar{u}_{n \min}$, $\bar{u}_{n \max}$ должны быть найдены путем декомпозиции связывающего их, в соответствии с (7), условия

$$(36) \quad u_{\min} \leq \sum_{n=1}^{\infty} \bar{u}_n \varphi_n(\mu_n, x) \leq u_{\max}; \quad u_{\min}, u_{\max} = \text{const},$$

накладываемого на реализуемое пространственно-временное воздействие $u(x, t)$ при заданных величинах u_{\min} и u_{\max} .

Применительно к данной ситуации с бесконечномерным векторным управлением $\bar{\mathbf{u}} = (\bar{u})$, $n = 1, 2, \dots$, необходимо, в отличие от (21), осуществить перевод объекта в окрестность теперь уже пересечения параллельной совокупности [2, 3] бесконечного числа притягивающих многообразий, определяемых выбором соответствующих макропеременных ψ_i , $i = 1, 2, \dots$

Распространение методики решения задачи АКАР для граничных управляющих воздействий по схеме (17) – (27) с представлением управлений \bar{u}_n в виде (19) и макропеременных в форме, подобной (25):

$$(37) \quad \begin{aligned} \psi_i(\mathbf{z}, z_{0i}) &= \text{th}[L_{1i} + L_{2i} \text{th}(\eta z_{0i}) + \beta_i z_i], \quad i = 1, 2, \dots \\ L_{1i} &= \frac{u_{i \max} + u_{i \min}}{2}; \quad L_{2i} = \frac{u_{i \max} - u_{i \min}}{2} \end{aligned}$$

приводит, в первом приближении, к типовому алгоритму вида (30):

$$(38) \quad \bar{u}_m(\mathbf{z}) = \begin{cases} -\beta_m z_m, & \bar{u}_{m \min} \leq -\beta_m z_m \leq \bar{u}_{m \max}; \\ \bar{u}_{m \max}, & \beta_m z_m \geq \bar{u}_{m \max}; \\ \bar{u}_{m \min}, & \beta_m z_m \leq \bar{u}_{m \min}; \end{cases}$$

для всех $m = 1, 2, \dots$, близкому к оптимальному по быстродействию при больших отклонениях системы от равновесного режима и осуществляющему независимо от вида функции $f_n(\mathbf{z})$ в (16) автономное пропорциональное регулирование каждой фазовой переменной

$$(39) \quad \bar{u}_{m\psi}(\mathbf{z}) = -\beta_m z_m, \quad m = 1, 2, \dots$$

при движении по пересечению многообразий (37).

Отсюда следует, что решение задачи АКАР сводится в данном случае к построению совокупности автономных контуров независимого регулирования достаточно большого числа мод распределенного выхода ОРП.

Так же, как и в случае с граничным управляющим воздействием, асимптотическая устойчивость движения объекта с управлениями (39) к точке $\mathbf{z} = 0$, описываемого системой уравнений

$$(40) \quad A_0 \frac{dz_n}{dt} = -\mu_n^2 z_n + f_n(\mathbf{z}) - \beta_n z_n, \quad n = 1, 2, \dots$$

$$z_n(0) = z_n^0,$$

обеспечивается соответствующим выбором коэффициентов обратных связей β_n , стесняемых дополнительным требованием сходимости ряда

$$(41) \quad u(x, t) = \sum_{n=1}^{\infty} \bar{u}_n(\mathbf{z}) \varphi_n(\mu_n, x),$$

восстанавливающего реализуемое распределенное управление в форме (7) по найденному алгоритму (39).

В линейном случае $f_n(\mathbf{z}) = 0$, $n = 1, 2, \dots$, с диагональной матрицей коэффициентов вывод об асимптотической устойчивости решений (40) заведомо справедлив для любых $\beta_n \geq 0$.

Самостоятельный интерес при реализации алгоритма (38) представляет специфическая проблема вычисления предельно допустимых значений $\bar{u}_{m \min}$, $\bar{u}_{m \max}$ модальных управлений для всех $m = 1, 2, \dots$ по заданным ограничениям на $u(x, t)$ в (36) [7, 17]. Во многих случаях программное управление $u(x, t)$ повторяет характер пространственного распределения известного управляющего воздействия $U^*(x, t)$ установившегося режима и определяется по его структуре в форме параметрически заданных явных зависимостей от своих аргументов. Анализ разложений этих зависимостей в бесконечный ряд (7) на предельно допустимых, согласно (36), величинах $u(x, t)$ позволяет найти возможные в этих условиях амплитуды $\bar{u}_{m \min}$, $\bar{u}_{m \max}$ гармоник этого разложения и осуществить, тем самым, необходимую декомпозицию ограничения (36) по модам $u(x, t)$ [7].

В типичном для приложений варианте представления $u(x, t)$ в форме [8]:

$$(42) \quad u(x, t) = \sum_{i=1}^s v_i(t) g_i(x), \quad s \geq 1,$$

с изначально фиксируемыми, исходя из требований технической реализации, зависимостями $g_i(x)$ от пространственной координаты (например, задаваемыми в классе кусочно-постоянных функций), задача сводится к синтезу $s \geq 1$ сосредоточенных управлений $v_i(\mathbf{z})$, $i = \overline{1, s}$, стесняемых заданными ограничениями

$$(43) \quad v_{i \min} \leq v_i \leq v_{i \max}, \quad i = \overline{1, s}.$$

Распространение на этот случай методики аналитического конструирования агрегированных регуляторов граничного управления ОРП по схеме (17) – (27) с использованием параллельной совокупности s притягивающих многообразий вида (21), (25) приводит к приближенному алгоритму управления, подобному (30):

$$(44) \quad v_i(\mathbf{z}) = \begin{cases} -\sum_{m=1}^{\infty} \beta_{im} z_m, & v_{i\min} \leq -\sum_{m=1}^{\infty} \beta_{im} z_m \leq v_{i\max}; \\ v_{i\max}, & -\sum_{m=1}^{\infty} \beta_{im} z_m \geq v_{i\max}; \\ v_{i\min}, & -\sum_{m=1}^{\infty} \beta_{im} z_m \leq v_{i\min}, \end{cases} \quad i = \overline{1, s},$$

где опять все коэффициенты обратных связей $\beta_{im}, i = \overline{1, s}; m = 1, 2, \dots$, могут быть выбраны из условий обеспечения требуемых качественных показателей переходного процесса.

Аналогично (33), (34), в случае

$$(45) \quad \beta_{im} = K_i \varphi_m(\mu_m, x_{ic}), x_{ic} \in [x_0, x_1], i = \overline{1, s}; m = 1, 2, \dots;$$

$$(46) \quad v_{i\psi}(\mathbf{z}) = -K_i \sum_{m=1}^{\infty} z_m(t) \varphi_m(\mu_m, x_{ic}) = -K_i [Q(x_{ic}, t) - Q^*(x_{ic})]; i = \overline{1, s},$$

этот алгоритм сводится к построению системы регулирования выхода объекта в s точках $x_{ic} \in [x_0, x_1]$ с обратными связями по непосредственно измеряемым величинам $Q(x_{ic}, t)$, если требуемые коэффициенты K_i передачи пропорциональных регуляторов обеспечивают асимптотическую устойчивость равновесного состояния системы.

Список литературы

- [1] Красовский А.А. Проблемы физической теории управления // Автоматика и телемеханика, 1990, №11, с.3-28.
- [2] Колесников А.А. Последовательная оптимизация нелинейных агрегированных систем управления. – М.: Энергоатомиздат, 1987.
- [3] Колесников А.А. Синергетическая теория управления. – М.: Энергоатомиздат, 1994.
- [4] Хакен Г. Синергетика. Иерархия неустойчивостей в самоорганизующихся системах и устройствах. – М.: Мир, 1985.
- [5] Бутковский А.Г. Методы управления системами с распределенными параметрами. – М.: Наука, 1975.
- [6] Рапопорт Э.Я. Структурно-параметрический синтез систем автоматического управления с распределенными параметрами // Изв. РАН. Теория и системы управления, 2006. №4, с. 47-60.
- [7] Рапопорт Э.Я. Оптимальное управление системами с распределенными параметрами. – М.: Высшая школа, 2009.
- [8] Рапопорт Э.Я. Структурное моделирование объектов и систем управления с распределенными параметрами. – М.: Высшая школа, 2003.
- [9] Самарский А.А., Вабишевич П.Н. Вычислительная теплопередача. – М.: УРСС, 2003.
- [10] Плешивцева Ю.Э., Рапопорт Э.Я. Метод последовательной параметризации управляющих воздействий в краевых задачах оптимального управления системами с распределенными параметрами // Известия РАН. Теория и системы управления, 2009. №3. С. 22-33.
- [11] Валеев Г.К., Жаутыков О.А. Бесконечные системы дифференциальных уравнений. – Алма-Ата: Наука Казахской ССР, 1974.
- [12] Коваль В.А. Спектральный метод анализа и синтеза распределенных управляемых систем. – Саратов: СГТУ, 1997.
- [13] Летов А.М. Динамика полета и управление. – М.: Наука, 1969.
- [14] Андреев Ю.Н. Управление конечномерными линейными объектами. – М.: Наука, 1976.
- [15] Воронов А.А. Введение в динамику сложных управляемых систем. – М.: Наука, 1985.
- [16] Ким Д.П. Теория автоматического управления. Т.2. – М.: Физматлит, 2007.
- [17] Черноусько Ф.Л., Афанасьевский И.М., Решмин С.А. Методы управления нелинейными механическими системами – М.: Физматлит, 2006.

СИНТЕЗ НЕЛИНЕЙНЫХ СИСТЕМ КОМБИНИРОВАННОГО УПРАВЛЕНИЯ С БИНАРНО-ОПЕРАТОРНЫМИ СВЯЗЯМИ

О.С. Гавриленко, В.С. Елсуков, В.И. Лачин

Южно-Российский Государственный Технический Университет
(Новочеркасский политехнический институт)
346428, Новочеркасск, Ростовская обл., ул. Просвещения, 132, Россия
olga_gavrilenko@inbox.ru, elsvs@mail.ru, vlachin@mail.ru
тел: +7 (8635) 25-59-27, факс: +7 (8625) 25-57-39

Ключевые слова: система автоматического управления, комбинированное управление, синтез, агрегированный закон

Abstract

The method of designing of the aggregated laws of the combined management is presented by nonlinear objects, the example of synthesis of nonlinear system of automatic control is considered.

Введение

Синтез законов управления нелинейными объектами в условиях ограниченной неопределенности является достаточно актуальной задачей [1-5]. В настоящей статье данная задача решается путем синтеза такой системы автоматического управления, составляющие закона управления которой формируются с помощью дифференциальных бинарно-операторных связей.

1 Постановка задачи

Рассматривается класс нелинейных объектов, которые можно описать уравнением вида:

$$(1) \quad y^{(n)}(t) = k(y, \dot{y}, \dots, y^{(n-1)}, t)u(t) - \varphi(y, \dot{y}, \dots, y^{(n-1)}, \lambda, t),$$

где $y(t)$ – выходная величина; $u(t)$ – управляющее воздействие; $\lambda = \lambda(t)$ – неизмеряемое внешнее возмущающее воздействие; $k(y, \dot{y}, \dots, y^{(n-1)}, t)$, $\varphi(y, \dot{y}, \dots, y^{(n-1)}, \lambda, t)$ – неизвестные ограниченные непрерывно-дифференцируемые нелинейные функции, скорость изменения которых является тоже ограниченной, т.е.

$$(2) \quad \begin{cases} |dk(\dots)/dt| \leq v_1 = const, \\ |d\varphi(\dots)/dt| \leq v_2 = const, \end{cases}$$

причем $k(\dots) = 0$ при $y = \dot{y} = \dots = y^{(n-1)} = 0$ и $k_1 \leq k(\dots) \leq k_2$.

Указанные функции представляют собой неопределенные характеристики объекта управления, причем зависимость их от времени t отражает действие параметрических возмущений.

Необходимо найти для объекта (1) такой закон управления $u(t) = u(y, \dot{y}, \dots, y^{(n)})$, чтобы движение САУ удовлетворяло, например, линейному дифференциальному уравнению:

$$(3) \quad \left(1 + \sum_{i=1}^n \tau_i p^i\right) y(t) = g,$$

где g – сигнал задания, $g = \text{const}$; τ_i $i = \overline{1, n}$ – неопределенные постоянные коэффициенты, которые подлежат определению путем решения задачи на условный экстремум функционала:

$$(4) \quad J = \int_0^{\infty} [y(0) - y(t)] dt \rightarrow 0, \quad y(0) > 0$$

$$y^{(i)}(0) = 0 \quad (i = \overline{1, n-1})$$

при ограничении

$$(5) \quad |A(p)|_{p=j\omega} \geq 1$$

где $A(p)$ – характеристический полином линеаризованной системы; $p = d/dt$.

2 Алгоритм структурного синтеза

Из заданного уравнения движения системы (3) можно выразить требуемый закон изменения старшей (n -й) производной выходной величины объекта (1):

$$(6) \quad F_n(t) = \frac{1}{\tau_n} \left\{ g - \left(1 + \sum_{i=1}^{n-1} \tau_i p^i \right) y(t) \right\}.$$

Подставляя полученное выражение в уравнение объекта (1) вместо $y^{(n)}(t)$ можно найти искомый закон управления в общем виде, который с учетом условий его реализации можно представить следующим образом [8]:

$$(7) \quad u(t) = \hat{k}^{-1}(t) \left[\frac{1}{\tau_n} \{ W_{И}(p)g - W_{Д}(p)y(t) + \tau_n \hat{\varphi}(t) \} \right],$$

где $\hat{k}^{-1}(t)$, $\hat{\varphi}(t)$ – оценки обратного значения текущего коэффициента усиления и эквивалентного аддитивного возмущения $\varphi(\dots)$ старшей производной выходной величины объекта (1). В закон управления (7) введены дифференцирующий фильтр с передаточной функцией $W_{Д}(p) = \left(1 + \sum_{i=1}^{n-1} \tau_i p^i \right) / (1 + \mu p)^n$ для реализации форсирующего звена и инерционный фильтр с передаточной функцией $W_{И}(p) = 1 / (1 + \mu p)^n$ и достаточно малой постоянной времени μ для согласования каналов формирования сравнимых сигналов.

Для получения оценок $\hat{k}^{-1}(t)$ и $\hat{\varphi}(t)$ можно учесть следующее. Если продифференцировать по времени обе части уравнения (1), то получим: $y^{(n+1)} = k(\dots) \cdot du/dt + u(t) \cdot dk(\dots)/dt - d\varphi(\dots)/dt$. В силу ограничений (2) в синтезируемой системе можно создать две разнотемповые составляющие скорости изменения старшей производной выходной величины: одна, медленно меняющаяся, создается составляющими $u(t) \cdot dk(\dots)/dt$ и $d\varphi(\dots)/dt$, а другую, быстро меняющуюся, можно создать составляющей $k(\dots) \cdot du/dt$. Тогда оценку $\hat{k}^{-1}(t)$ можно сформировать с помощью операции деления:

$$(8) \quad \hat{k}^{-1}(t) = | \dot{u}(t) | / | y^{(n+1)}(t) |.$$

Реализация операции (8) может быть осуществлена при помощи дифференциальной бинарно-операторной обратной связи со следующим алгоритмом оператора [8]:

$$(9) \quad \hat{k}^{-1}(t) = \frac{1}{\sigma_0 p} \left\{ W_{Д1}(p)u(t) - \hat{k}^{-1}(t)W_{Д2}(p)y(t) \right\} + \hat{k}^{-1}(0),$$

где $\hat{k}^{-1}(0) = \sqrt{k_1^{-1}k_2^{-1}}$, $W_{Д1}(p)$ – передаточная функция дифференцирующего фильтра, $W_{Д1}(p) = \tau_n \sigma p / ((1 + \sigma p)(1 + \mu p)^n)$, с помощью которого осуществляется приближенная реализация производной сигнала $u(t)$; σ – постоянный коэффициент, значение которого будет определено при параметрическом синтезе; $W_{Д2}(p)$ – передаточная функция дифференцирующего фильтра, $W_{Д2}(p) = \tau_n \sigma p^{(n+1)} / ((1 + \sigma p)(1 + \mu p)^n)$, который служит для получения $(n+1)$ -й производной выходной величины.

Если подключить ко входу объекта (1) умножитель и подать на один его вход сигнал оценки (9), то другой вход последнего будет представлять собой вход преобразованного объекта управления. Обозначим через u_0 входной сигнал преобразованного объекта, охваченного мультипликативной обратной связью (9). Введение данной связи приводит к тому, что коэффициент усиления прямой цепи преобразованного объекта становится равным единице.

Тогда с учетом ограничения функции $\varphi(\dots)$, ее оценку $\hat{\varphi}(t)$ можно получить с помощью применения еще одной бинарно-операторной связи. Но в отличие от известных технических решений [7, 8], сигнал этой связи нами предложено формировать в виде двух составляющих. Одна из них пропорциональна сигналу ”дифференциальной вилки”, т.е. сигналу рассогласования между требуемым и текущим значениями старшей производной выходной величины.

$$\varepsilon = \tau_n W_{И}(p)F_n(t) - W_{Д3}(p)y(t),$$

где $W_{Д3}(p)$ – передаточная функция дифференцирующего фильтра, $W_{Д3}(p) = \tau_n p^n / (1 + \mu p)^n$, который служит для получения n -й производной выходной величины.

Другая $\varphi_{И}$, равна интегралу от суммы двух сигналов. Один из них – это сигнал ε . Другой, знакопеременный пропорционален модулю измеренного косвенным способом эквивалентного возмущения $\varphi(\dots)$ старшей производной выходной величины. При этом знак указанного сигнала определяется полярностью сигнала ”дифференциальной вилки”, т.е.

$$(10) \quad \tau_n \hat{\varphi}(t) = \varepsilon + \frac{1}{\sigma p} \left\{ \varepsilon + |\varphi_{И}(t)| \text{sign} \varepsilon \right\},$$

где $\varphi_{И}(t) = \tau_n W_{И}(p)u_0 - W_{Д3}(p)y(t)$.

Структурная схема САУ, соответствующая уравнениям (1), (7), (9) и (10) приведена на рисунке 1. На нем пунктирной линией обозначен контур стабилизации коэффициента усиления объекта управления с мультипликативной обратной связью по сигналу (9).

$$(12) \tau_1 = 2^n (\sigma + n\mu); \tau_{j+1} = \begin{cases} \tau_1^2 / 2 & \text{при } j = 1; \\ \tau_j^2 / 2\tau_{j-1} & \text{при } j = 2, n-1. \end{cases}$$

4 Пример синтеза

Рассмотрим задачу синтеза САУ для нелинейного объекта второго порядка, аналогичного приведенному в [3], но имеющего неопределенный коэффициент усиления:

$$(13) \ddot{y}(t) = 5 - \varphi(y, \dot{y}, \dots, y^{(n-1)}, \lambda, t) + k(y, \dot{y}, \dots, y^{(n-1)}, t)u(t),$$

где в качестве неопределенных функций $\varphi(\dots)$ и $k(\dots)$ приняты следующие:

$$\varphi(y, \dot{y}, \dots, y^{(n-1)}, \lambda, t) = y^3(t) + \lambda,$$

$$k(y, \dot{y}, \dots, y^{(n-1)}, t) = 1 + y(t), \quad 1 \leq k(\dots) \leq 20.$$

Формируем требуемый закон движения системы

$$(14) (1 + \tau_1 p + \tau_2 p^2)y(t) = g.$$

Из уравнения (14) выражаем требуемый закон изменения второй производной выходной величины объекта (13)

$$(15) F_2(t) = \frac{1}{\tau_2} \{g - (1 + \tau_1 p)y(t)\}.$$

Подставляя выражение (15) в уравнение объекта (13) находим искомый закон управления в общем виде, который с учетом условий его реализации можно представить следующим образом:

$$(16) u(t) = \hat{k}^{-1}(t) \left[\frac{1}{\tau_2} \left\{ \frac{1}{(1 + \mu p)^2} g - \frac{1 + \tau_1 p}{(1 + \mu p)^2} y(t) - \tau_2 5 + \tau_2 \hat{\varphi}(t) \right\} \right].$$

где $\hat{k}^{-1}(t)$, $\hat{\varphi}(t)$ – оценки обратного значения коэффициента усиления и эквивалентного возмущения $\varphi(\dots)$ старшей производной выходной величины объекта (13).

Для получения оценки $\hat{k}^{-1}(t)$ применяем дифференциальную бинарно-операторную связь с аналогичным (9) алгоритмом:

$$(17) \hat{k}^{-1}(t) = \frac{1}{\sigma_0 p} \left\{ \left| \frac{\tau_2 \sigma p}{(1 + \sigma p)(1 + \mu p)^2} u(t) \right| - \hat{k}^{-1}(t) \left| \frac{\tau_2 \sigma p^3}{(1 + \sigma p)(1 + \mu p)^2} y(t) \right| \right\} + \hat{k}^{-1}(0),$$

где $\hat{k}^{-1}(0) = \sqrt{1/20} \approx 0,22$, $\sigma = 4\mu$.

Оценку $\hat{\varphi}(t)$ получаем с помощью еще одной дифференциальной операторной связи. Сигнал это дифференциальной связи формируем в виде суммы двух составляющих:

$$(18) \tau_n \hat{\varphi}(t) = \varepsilon + \frac{1}{\sigma p} \{ \varepsilon + |\varphi_H(t)| \text{sign} \varepsilon \},$$

где $\varphi_{И}$ – измеренное косвенным способом эквивалентное возмущение $\varphi(\dots)$,

$$\varphi_{И}(t) = \frac{\tau_2}{(1 + \mu p)^2} u_0 - \frac{\tau_2 p^2}{(1 + \mu p)^2} y(t);$$

$$\varepsilon = \frac{\tau_2}{(1 + \mu p)^2} F_2(t) - \frac{\tau_2 p^2}{(1 + \mu p)^2} y(t).$$

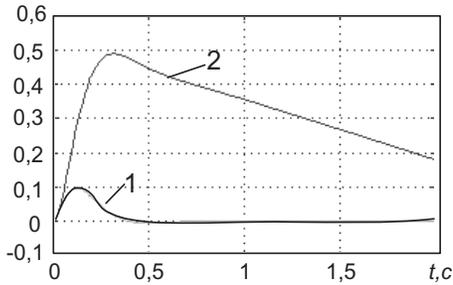


Рисунок 2 – Переходные процессы

Параметры управляющего устройства синтезированной системы были выбраны следующими: $\mu=0.005$ с; $\sigma=0.02$ с; $\tau_1=0.12$ с; $\tau_2=0.0072$ с². Исследование системы (13), (16)-(18) выполнялось методом компьютерного моделирования. Исследовалась реакция синтезированной системы на единичное ступенчатое изменение возмущающего воздействия λ . На рисунке 2 представлены реакции синтезированной системы (кривая-1) и известной системы [8] (кривая-2). Из сравнения полученных кривых видно, что динамическая точность управления синтезированной системы гораздо

выше динамической точности управления известной системы [8]. В частности, интегральная ошибка в известной системе на порядок больше, чем в синтезированной системе.

Список литературы

- [1] Morse A.S. High-order parameter tuners for adaptive control of nonlinear systems// Systems, Models and Feedback: Theory and Applications/ Eds A. Isodori, T.J. Tarn. Birkhauser, 1992.
- [2] Никифоров В.О. Адаптивное и робастное управление с компенсацией возмущений. – СПб.: Наука, 2003.
- [3] Воронов К.В., Королева О.И., Никифоров В.О. Робастное управление нелинейными объектами с функциональными неопределенностями// Автоматика и телемеханика, 2001, №2. – С.112-121.
- [4] Емельянов С.В., Коровин С.К. Новые типы обратных связей: Управление при неопределенности. – М.: Наука; Физматлит, 1997.
- [5] Востриков А.С. Синтез систем регулирования методом локализации. – Новосибирск: Изд-во НГТУ, 2007.
- [6] Фрер Ф., Орттенбургер Ф. Введение в электронную технику регулирования. – М.: Энергия, 1973.
- [7] Елсуков В.С., Лачин В.И. Синтез агрегированных законов управления нелинейными объектами с функциональными неопределенностями // Проблемы управления и моделирования в сложных системах: труды X международной конференции (Самара, 23-25 июня 2008 г.) / Под ред.: акад. Е.А. Федосова, акад. Н.А. Кузнецова, проф. В.А. Виттиха. – Самара: Самарский научный центр РАН, 2008. – С.87-94.
- [8] Елсуков В.С., Лачин В.И. Синтез нелинейных систем управления с комбинированной знакопеременной обратной связью // Проблемы управления и моделирования в сложных системах: труды XI международной конференции (Самара, 22-24 июня 2009 г. Самара, Россия) / Под ред.: акад. Е.А. Федосова, акад. Н.А. Кузнецова, проф. В.А. Виттиха. – Самара: Самарский научный центр РАН, 2009. – С.103-109.

СРАВНИТЕЛЬНЫЙ МНОГОКРИТЕРИАЛЬНЫЙ АНАЛИЗ ЭФФЕКТИВНОСТИ ОПЕРАЦИОННОЙ ДЕЯТЕЛЬНОСТИ ПРОМЫШЛЕННЫХ ПРЕДПРИЯТИЙ

Н.В. Дилигенский¹, М.В. Цапенко¹, А.Н. Давыдов²

¹Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
cscmp@iccs.ru

тел./факс: +7 (846) 333-27-70

²Самарский государственный технический университет
443010, Самара, ул. Молодогвардейская, 244, Россия
nich@samgtu.ru

тел: +7 (846) 278-43-16

Ключевые слова: сравнительная эффективность, многокритериальная оценка, *Data Envelopment Analysis*, экспертные оценки, метод анализа иерархий

Abstract

The problem of construction of efficiency industrial companies is considered. This research is using methods *Data Envelopment Analysis* and *Analytical Hierarchy Process*.

Введение

В условиях конкурентного взаимодействия и нестабильной внешней среды актуальной является задача сравнительного многокритериального анализа эффективности операционной деятельности промышленных предприятий на основе некоторых критериев эффективности.

Существуют различные способы получения сравнительных многокритериальных оценок эффективности. Одни основаны на свёртывании частных критериев с учётом их значимости, определяемой, как правило, на основе экспертных оценок. Другие подходы наряду с экспертными оценками дополнительно используют модельные конструкции, например, *метод анализа иерархий* Томаса Саати (МАИ). Ряд способов основан на применении только формальных математических подходов – аппарата математического программирования – метод *Data Envelopment Analysis* (DEA).

В целом, конструктивное применение этих методов позволяет получать интегральные оценки эффективности на основе свертывания локальных критериев по тем или иным алгоритмам.

1 Формализованный способ многокритериального оценивания сравнительной эффективности

В последнее время в западной литературе широкое распространение для оценки сравнительной эффективности разнообразных сложных производственных региональных систем получил метод *Data Envelopment Analysis* [1], название которого на русский язык, можно перевести как анализ «упаковки» (охвата) данных. Метод имеет обширную библиографию, состоящую из более чем 1000 статей, монографий и диссертаций и применяется для анализа организаций, процессов, производств и операций в различных сферах деятельности: страховании [2], банковском деле [3], нефтегазовой отрасли [4], розничной торговле [5], сфере услуг [6], здравоохранении [7], образовании [8] и прочее.

В русской научной литературе использование метода DEA носит фрагментарный характер [9], [10], и название метода в [10] переводится как анализ среды функционирования.

Базовая модель методологии DEA разработана в 1957 году американским исследователем Фареллом (Farrel M. J.) [11] и усовершенствована в 1978 году Чарнсом, Купером, Родсом (Charnes, Cooper, Rhodes) [12]. Впоследствии, в 1984 году данный подход был развит Банкиром, Чарнсом, Купером (Banker, Charnes, Cooper) [13]. В литературе эти методы носят названия, соответственно, CCR и ВСС моделей. Основная идея метода CCR состоит в способе формирования оценки сравнительной эффективности сложных многомерных систем.

В постановке DEA полагается, что исследуемая сложная система (СС) характеризуется m входами и k выходами – рисунок. 1.

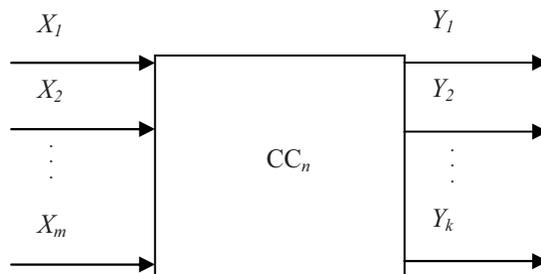


Рисунок 1 – Многомерная сложная система (СС)

Выходные величины Y_1, Y_2, \dots, Y_k выбираются таким образом, чтобы каждая из них характеризовала фактор, играющий положительную роль в суммарном показателе эффективности f исследуемой системы:

$$(1) \frac{\partial f(Y_1, Y_2, \dots, Y_k)}{\partial Y_i} > 0, i = 1, 2, \dots, k.$$

Таковыми выходными показателями могут быть самые различные величины, характеризующие разнообразные аспекты деятельности: производственно-технологические (объем и качество конечной продукции, надёжность, долговечность), управленческие (устойчивость, управляемость, наблюдаемость), экономические (прибыль, доход, рентабельность), экологические (чистота производств) и другие.

В качестве входных величин X_1, X_2, \dots, X_m берутся затраты на деятельность, уменьшение которых приводит к повышению показателя суммарной эффективности:

$$(2) \frac{\partial f(X_1, X_2, \dots, X_m)}{\partial X_j} < 0, j = 1, 2, \dots, m.$$

Основным классом входных характеристик являются различные ресурсы систем – сырьевые, материальные, информационные, капитальные, финансовые, трудовые, энергетические и другие.

На основе содержательного выбора m входов $X_1, X_2, X_3, \dots, X_m$ и k выходов $Y_1, Y_2, Y_3, \dots, Y_k$, совокупность которых с позиции исследователя даёт достаточно полную и адекватную характеристику систем, структура комплексного показателя эффективности исследуемой системы, в базовом варианте метода DEA, формируется следующим образом:

$$(3) f = \frac{u_1 \cdot Y_1 + u_2 \cdot Y_2 + \dots + u_k \cdot Y_k}{v_1 \cdot X_1 + v_2 \cdot X_2 + \dots + v_m \cdot X_m}.$$

В (3) u_i ($i = 1, 2, \dots, k$) – положительные весовые коэффициенты, характеризующие относительный вклад каждого из выходных факторов Y_i в суммарный коэффициент эффективности f . Соответственно, v_j ($j = 1, 2, \dots, m$) – веса входных величин X_j .

В формуле (3) веса u_i, v_j , в общем случае, являются произвольными, неизвестными, и от них требуется лишь положительность – $u_i \geq 0, v_j \geq 0$.

В качестве методологии нахождения численных показателей комплексной эффективности каждого из n ($n = 1, 2, \dots, N$) объектов в методе ССР полагается, что величины всех показателей комплексной эффективности f конечны, и осуществляется ранжировка этих значений на числовом интервале $[0, 1]$, исходя из условия максимизации показателей эффективности (3) для каждой из систем.

В этом случае, задача отыскания для n -ой системы обобщённого показателя эффективности f_n и соответствующего ему набора весовых коэффициентов u_{in} и v_{jn} записывается следующим образом.

Найти максимум функционала:

$$(4) \quad f_n = \max_{u_{in}, v_{jn} \in G} \frac{u_{1n} \cdot Y_{1n} + u_{2n} \cdot Y_{2n} + u_{3n} \cdot Y_{3n} + \dots + u_{kn} \cdot Y_{kn}}{v_{1n} \cdot X_{1n} + v_{2n} \cdot X_{2n} + v_{3n} \cdot X_{3n} + \dots + v_{mn} \cdot X_{mn}},$$

при наличии системы ограничений, определяющих область значений G весов u_{in} и v_{jn} :

$$(5) \quad \left\{ \begin{array}{l} \frac{u_{1n} \cdot Y_{11} + u_{2n} \cdot Y_{21} + u_{3n} \cdot Y_{31} + \dots + u_{kn} \cdot Y_{k1}}{v_{1n} \cdot X_{11} + v_{2n} \cdot X_{21} + v_{3n} \cdot X_{31} + \dots + v_{mn} \cdot X_{m1}} \leq 1, \\ \frac{u_{1n} \cdot Y_{12} + u_{2n} \cdot Y_{22} + u_{3n} \cdot Y_{32} + \dots + u_{kn} \cdot Y_{k2}}{v_{1n} \cdot X_{12} + v_{2n} \cdot X_{22} + v_{3n} \cdot X_{32} + \dots + v_{mn} \cdot X_{m2}} \leq 1, \\ \dots\dots\dots \\ \frac{u_{1n} \cdot Y_{1n} + u_{2n} \cdot Y_{2n} + u_{3n} \cdot Y_{3n} + \dots + u_{kn} \cdot Y_{kn}}{v_{1n} \cdot X_{1n} + v_{2n} \cdot X_{2n} + v_{3n} \cdot X_{3n} + \dots + v_{mn} \cdot X_{mn}} \leq 1, \\ u_{in} \geq 0; v_{jn} \geq 0, \end{array} \right.$$

$$i = \{1, 2 \dots k\}; j = \{1, 2 \dots m\}; n = \{1, 2 \dots N\}.$$

В (5) величины X_{jn} ($j = 1, 2 \dots m$) и Y_{in} ($i = 1, 2 \dots k$) являются, соответственно, численными значениями входа X_j и выхода Y_i для n -ой системы.

Система соотношений (4), (5) для $n = \{1, 2, \dots, N\}$ определяет N задач математического программирования. Решение каждой n -ой задачи ($n = 1, 2, \dots, N$) для n -ой системы даёт значение n -го показателя эффективности f_n , ранжированное на единичном интервале $[0, 1]$, и соответствующий ему набор весовых коэффициентов $\{u_{1n}, u_{2n}, \dots, u_{kn}\}$ и $\{v_{1n}, v_{2n}, \dots, v_{mn}\}$, максимизирующий функционал (4). Вообще говоря, эти веса u_{in} и v_{jn} будут различными для различных n -ых систем.

Отметим, что в общем случае, решение задачи многокритериальной оптимизации (4), (5) имеет не единственное, а целое множество решений, удовлетворяющих заданным ограничениям (5) [14]. При этом, одной и той же относительной эффективности может отвечать некоторое множество решений как в пространстве входных и выходных параметров, так и в пространстве весов функционала.

2 Экспертная модель многокритериального оценивания сравнительной эффективности

В широком круге задач обобщённого оценивания являются неизвестными сами первичные локальные показатели эффективности. Возможным подходом в такой ситуации является

применение разработанного в начале 1970 года американским математиком Томасом Саати метода анализа иерархий (*Analytical Hierarchy Process*) [15], [16].

Этот метод получил широкое распространение и активно применяется в различных отраслях для решения прикладных задач многокритериального оценивания на основе обработки субъективных экспертных суждений в форме парных сравнений. Метод применяется к различным по своей сущности и характеристикам системам и объектам. Так известны постановки и решения задач ранжирования и выбора возможных рыночных альтернатив (продуктов и услуг), проведения маркетинговых исследований, определения значимости совокупности факторов, стратегий развития, распределения ресурсов, и т.д.

В основе метода лежит процедура декомпозиции сложной проблемы – представление её в виде структурированного набора компонент или критериев, взаимосвязи между которыми формируются в иерархическом варианте представления. При этом вершиной иерархии является общая цель – желаемое состояние системы. Следующий уровень представляет собой детализацию общей цели в виде набора критериев, компонент или сил, оказывающих влияние на достижение обозначенного результата. На самом нижнем уровне иерархии представлены возможные альтернативы, степень приоритетности (значимости) которых требуется оценить.

В методе осуществляется процедура парного сравнения отдельных компонент иерархии между собой. Результаты экспертного оценивания представляются в виде набора положительных обратносимметричных матриц парных сравнений вида:

$$(6) \quad A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

В (6) каждый элемент матрицы a_{ij} определяет субъективное мнение эксперта относительно значимости оцениваемой i -ой компоненты иерархии относительно j -ой.

Матрицы вида (6) составляются для сравнения важности элементов каждого нижестоящего уровня иерархии относительно вышестоящего. Оценки элементов иерархии производятся в соответствии с девятибалльной шкалой относительной важности, при этом для оценивания n компонент необходимо сформировать $n(n-1)/2$ суждений. На основе матриц экспертных суждений вычисляются локальные приоритеты (значимости факторов) по каждой отдельной матрице, и критерии качества экспертных оценок.

Локальные приоритеты формализуются в методе МАИ в виде нормализованного главного собственного вектора матрицы. Широко распространённым приближённым способом нахождения собственного вектора является подход, основанный на нахождении среднего геометрического. В соответствии с ним компонента собственного вектора i -ой строки находится по формуле:

$$(7) \quad b_i = \sqrt[n]{a_{i1} \times a_{i2} \times \dots \times a_{in}}.$$

Затем все компоненты b_n собственного вектора нормализуются на единицу путем их деления на сумму $\sum b_i$ – и эти значения определяют локальную значимость относительно оцениваемых элементов.

Критерии качества работы экспертов отыскиваются на основе расчёта показателей согласованности суждений для каждой матрицы парных сравнений по максимальному собственному значению матрицы.

Если качество суждений экспертов неприемлемо – оценки несогласованны между собой – то следует пересматривать суждения экспертов по соответствующей матрице.

Завершающим этапом метода является нахождение интегральных обобщённых оценок значимости альтернатив. Процедура свертывания локальных приоритетов заключается в нахождении взвешенных сумм по всем элементам одного уровня, учитывающих весовые коэф-

фициенты (вектора приоритетов) вышестоящего уровня иерархии. Так в случае с трёхуровневой иерархией в соответствии с весовыми коэффициентами критериев второго уровня производится свёртка по каждой из альтернатив третьего уровня. Найдя, таким образом, значения глобальных приоритетов, делается окончательный вывод о сравнительной значимости оцениваемых альтернатив.

3 Анализ эффективности операционной деятельности региональных энергетических производств

В ряде работ [17-19] для более полного, всестороннего учёта разнообразных факторов при оценивании сложных систем применялись одновременно две взаимодополняющие методологии DEA и МАИ.

Рассмотрим способ формирования комплексных оценок функционирования сложных систем основанный на совместном применении методологий DEA и МАИ для анализа сравнительной эффективности региональных энергетических производств России.

Исходные данные для анализа возьмём на основе материалов ежегодного рейтинга Рейтингового агентства «Эксперт РА» [20]. Для определения оценки комплексной эффективности, выделим ряд основных статистических показателей, характеризующих деятельность крупнейших энергетических предприятий: прибыль после налогообложения (чистая прибыль), объем реализации и темп роста, в качестве производного показателя возьмем рентабельность – (отношение чистой прибыли к объему реализации).

Численные значения основных и производных показателей деятельности 35 региональных энергетических компаний представлены в таблице 1.

Таблица 1 – Значения локальных критериев эффективности

№№	Компания	Чистая прибыль, млн руб.	Объем реализации, млн руб.	Темп прироста, %
1.	Концерн "Энергоатом"	10159.4	151675.1	22.7
2.	"РусГидро"	-19955.4	107669.8	33.3
3.	"ФСК ЕЭС"	4465.5	68485.0	11.6
4.	МОЭСК	7190.1	65331.9	35.1
5.	"Башкирэнерго"	3315.2	61906.1	14.9
6.	ИНТЕР РАО ЕЭС	2839.6	60894.1	36.3
7.	"Татэнерго"	5073.4	56964.8	28.9
8.	ОГК-1	7457.0	49898.0	11.2
9.	"МРСК Центра"	2785.7	43726.6	31.9
10.	"Энел ОГК-5"	1630.4	42813.0	27.9
11.	ТГК-9	796.6	41863.3	26.0
12.	"Иркутскэнерго"	5509.3	41829.3	71.0
13.	"Волжская ТГК"	264.4	38821.2	18.7
14.	ОГК-4	6078.7	38013.3	20.9
15.	Дальневосточная генерирующая компания	-1890.1	35209.0	9.3
16.	"Тюменьэнерго"	1080.4	35193.8	26.7
17.	"МРСК Центра и Приволжья"	1428.6	31686.2	нет данных
18.	ТГК-4	-9719.9	29222.8	18.7
19.	"МРСК Сибири"	898.7	27157.1	нет данных
20.	"Новосибирскэнерго"	-114.0	24652.0	12.7
21.	"МРСК Урала"	861.4	23777.7	нет данных

Окончание таблицы 1

22.	ТГК-2	-4783.9	23162.8	11.4
23.	"МРСК Волги"	808.1	22614.5	нет данных
24.	ТГК-10	2489.3	22282.3	-8.0
25.	"Кузбассэнерго"	-2141.6	20925.5	11.8
26.	"Ленэнерго"	1539.6	20827.8	36.4
27.	ТГК-6	46.5	20106.7	77.1
28.	"Кубаньэнерго"	601.9	17840.1	21.3
29.	"МРСК Северо-Запада"	598.2	15954.2	нет данных
30.	ТГК-11	775.9	15925.8	21.7
31.	ТГК-5	723.4	15354.3	20.2
32.	ТГК-13	-1221.6	15036.2	25.4
33.	"Якутскэнерго"	88.8	13854.9	1.2
34.	СО ЕЭС	-2623.6	13348.5	33.6
35.	"МРСК Юга"	486.0	13092.6	нет данных

В качестве критерия оценки сравнительной эффективности по методологии DEA была взята линейная комбинация относительных показателей – рентабельности и темпа прироста:

$$(8) f = \frac{I}{Q} \cdot v_1 + T \cdot v_2,$$

где I – чистая прибыль;

Q – объём реализации;

T – темп прироста;

v_1, v_2 – весовые коэффициенты.

В соответствии с методом DEA была проведена постановка и найдено решение соответствующих задач линейного программирования (4), (5). Результаты оценивания представлены на рисунке 2 (номер по горизонтальной шкале соответствует порядковому номеру предприятия в таблице 1).

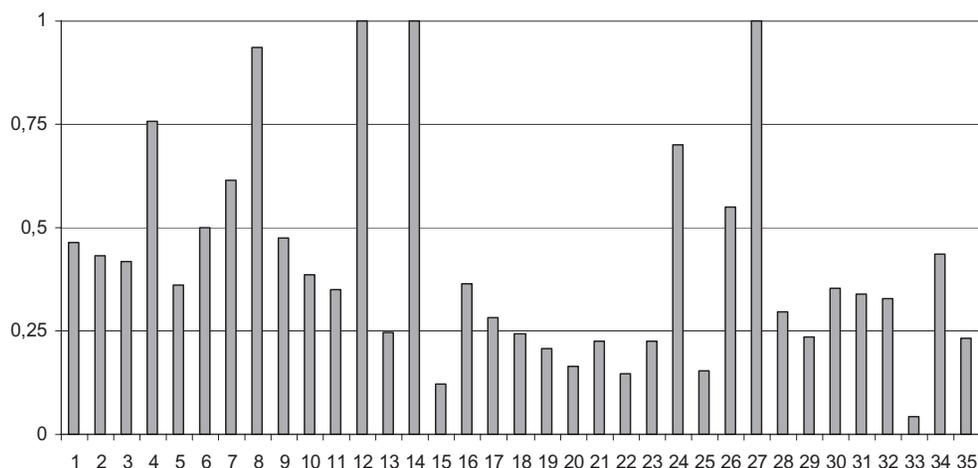


Рисунок 2 – Результаты DEA-оценки

Результаты формализованной оценки позволяют разбить все энергетические производств из анализируемой совокупности на три группы: первая группа – с значением критерия системной эффективности менее 0,25 – 12 предприятий; вторая группа – значение критерия эф-

фактивности от 0,25 до 0,5 – 15 предприятий; третья группа – системная эффективность более 0,5 – 8 предприятий.

Отметим, что три энергетические компании получили максимальную оценку сравнительной эффективности: «Иркутскэнерго», ОГК-4 и ТГК-6. ОГК-4 имеет максимальную рентабельность – 16%, «Иркутскэнерго» при достаточно высоком значении рентабельности – 13% имеет высокий темп прироста – 71%, ТГК-6 при самых низких значениях рентабельности имеет максимальный темп прироста (77%) из всей анализируемой группы предприятий.

Выделим из данной совокупности предприятия, системная эффективность которых превышает 0,5 (третья группа). Отметим, что отдельный расчет DEA-оценок только для этой группы из восьми предприятий дает результаты, абсолютно совпадающие с оцениванием, проведенным по всей совокупности производственных объектов, что дополнительно подтверждает наличие эффективного фронта, состоящего из этих организаций.

Для выделенных таким образом компаний – таблица 2 – проведем альтернативное сравнение экспертным методом МАИ.

Таблица 2 – Значения локальных критериев эффективности и системная оценка для третьей группы

№	Компания	Чистая прибыль, млн руб.	Объем реализации, млн руб.	Рентабельность, %	Темп прироста, %	Значения DEA-оценки
1.	МОЭСК	7190.1	65331.9	11	35.1	0.7562
2.	"Татэнерго"	5073.4	56964.8	9	28.9	0.6135
3.	ОГК-1	7457.0	49898.0	15	11.2	0.9346
4.	"Иркутскэнерго"	5509.3	41829.3	13	71.0	1.0000
5.	ОГК-4	6078.7	38013.3	16	20.9	1.0000
6.	ТГК-10	2489.3	22282.3	11	-8.0	0.6986
7.	"Ленэнерго"	1539.6	20827.8	7	36.4	0.5499
8.	ТГК-6	46.5	20106.7	0.2	77.1	1.0000

В соответствии с алгоритмом метода построим иерархическую структуру задачи – рисунок 3.



Рисунок 3 – Иерархия системы критериев оценки энергетических производств

В соответствии с алгоритмом метода анализа иерархий проведем парные сравнения и математическую обработку экспертных суждений. В таблицах 3 и 4 представлены матрицы парных сравнений и нормированные экспертные оценки значимости.

Таблица 3 – Матрица парных сравнений и экспертные оценки по критерию рентабельность

Рентабельность	МОЭСК	"Татэнерго"	ОГК-1	"Иркутскэнерго"	ОГК-4	ТГК-10	"Ленэнерго"	ТГК-6	Норм. оценки значимости
МОЭСК	1	2	1/4	1/2	1/5	1	4	8	0,0863
"Татэнерго"	1/2	1	1/6	1/4	1/7	1/2	2	7	0,0501
ОГК-1	4	6	1	2	1/2	4	8	9	0,2457
"Иркутскэнерго"	2	4	1/2	1	1/3	2	6	8	0,1492
ОГК-4	5	7	2	3	1	5	9	9	0,3362
ТГК-10	1	2	1/4	1/2	1/5	1	4	9	0,0876
"Ленэнерго"	1/4	1/2	1/8	1/6	1/9	1/4	1	8	0,0320
ТГК-6	1/8	1/7	1/9	1/8	1/9	1/9	1/8	1	0,0128

Таблица 4 – Матрица парных сравнений и экспертные оценки по критерию темп прироста

Темп прироста	МОЭСК	"Татэнерго"	ОГК-1	"Иркутскэнерго"	ОГК-4	ТГК-10	"Ленэнерго"	ТГК-6	Норм. оценки значимости
МОЭСК	1	2	3	1/4	2	4	1	1/5	0,0929
"Татэнерго"	1/2	1	3	1/5	2	3	1/2	1/6	0,0657
ОГК-1	1/3	1/3	1	1/7	1/2	2	1/5	1/8	0,0313
"Иркутскэнерго"	4	5	7	1	6	8	4	1/2	0,2732
ОГК-4	1/2	1/2	2	1/6	1	3	1/2	1/6	0,0513
ТГК-10	1/4	1/3	1/2	1/8	1/3	1	1/4	1/9	0,0240
"Ленэнерго"	1	2	5	1/4	2	4	1	1/5	0,0990
ТГК-6	5	6	8	2	6	9	5	1	0,3626

Экспертное оценивание энергетических компаний было проведено на основе количественных значений критериев рентабельности и темпа прироста. Результаты качественно коррелируют между собой, например, при максимальном значении рентабельности у ОГК-4 (16%), экспертная оценка этой компании максимальна – 0,34 (таблица 3), минимальная рентабельность у ТГК-6, и её экспертная оценка – 0,01 (таблица 3), при этом для этой организации характерен максимальный темп прироста (77%), что отражает экспертная оценка значимости – 0,36 (таблица 4).

Графически результаты экспертного оценивания представлены на диаграммах рисунков 4, 5.



Рисунок 4 – Экспертные оценки значимости по критерию рентабельность

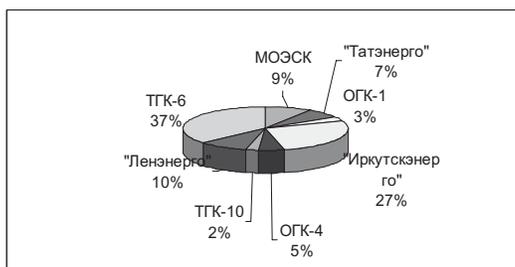


Рисунок 5 – Экспертные оценки значимости по критерию темп прироста

Так же следует отметить хорошие значения параметров качества работы экспертов – индекса согласованности и отношения согласованности. Так для матрицы парных сравнений по критерию рентабельности значение отношения согласованности составляет 4,84%, а для оценок по критерию темп прироста – 3,15%. При этом критическим значением для признания экспертных оценок несогласованными является величина 10% и более.

Для получения итоговых оценок проведём синтез глобальных приоритетов – найдём результирующие экспертные оценки значимости с учетом введенных критериев. Для этого необходимо определить весовые коэффициенты вклада критериев в системный показатель эффективности. Экспертным методом данные коэффициенты получили значения 0,6 – для рентабельности и 0,4 – для темпа прироста.

Результаты синтеза глобальных приоритетов в виде взвешенной суммы значимостей объектов по критериям с учетом их весов представлены в таблице 5.

Таблица 5 – Синтез глобальных приоритетов

№ №	Компания	Рентабельность	Темп прироста	Синтезированные оценки оценки МАИ	Значения DEA- оценки	Нормированная DEA- оценка	Отклонения
		0,6	0,4				
1	2	3	4	5	6	7	8
1.	МОЭСК	0,0863	0,0929	0,0889	0.7562	0,1154	0,0265
2.	"Татэнерго"	0,0501	0,0657	0,0564	0.6135	0,0936	0,0373
3.	ОГК-1	0,2457	0,0313	0,1599	0.9346	0,1426	-0,0173
4.	"Иркутскэнерго"	0,1492	0,2732	0,1988	1.0000	0,1526	-0,0462
5.	ОГК-4	0,3362	0,0513	0,2223	1.0000	0,1526	-0,0697
6.	ТГК-10	0,0876	0,0240	0,0622	0.6986	0,1066	0,0445
7.	"Ленэнерго"	0,0320	0,0990	0,0588	0.5499	0,0839	0,0251
8.	ТГК-6	0,0128	0,3626	0,1528	1.0000	0,1526	-0,0001

Как видно, синтезированные экспертные оценки имеют высокую корреляцию с полученными ранее системными DEA-оценками (таблица 5, 6-я колонка), так три предприятия, имеющие максимальную DEA-оценку (№№ 4, 5, 8) получили соответствующие максимальные экспертные оценки.

Сопоставим полученные аналитические оценки эффективности по DEA методологии и экспертные по методу МАИ. Для этого проведём нормирование системных DEA-оценок эффективности (таблица 5, 7-я колонка) и рассмотрим их отклонения от экспертных оценок (таблица 5, 8-я колонка). Результаты анализа отклонений представлены на диаграмме рисунка 6.

Как видно из диаграммы рисунка 6 для ТГК-6, имеющей максимальную абсолютную DEA-оценку, нормированные значения эффективностей практически совпадают. Небольшие отклонения с тенденцией завышения экспертных оценок имеют предприятия – лидеры, имеющие единичные («Иркутскэнерго» и ОГК-4) или близкие к единице (ОГК-1) абсолютные DEA-оценки.

Для оставшихся четырёх предприятий экспертные оценки несколько ниже формальных DEA-оценок, в среднем на 0,03 относительно приведённой к единице суммарной эффективности.

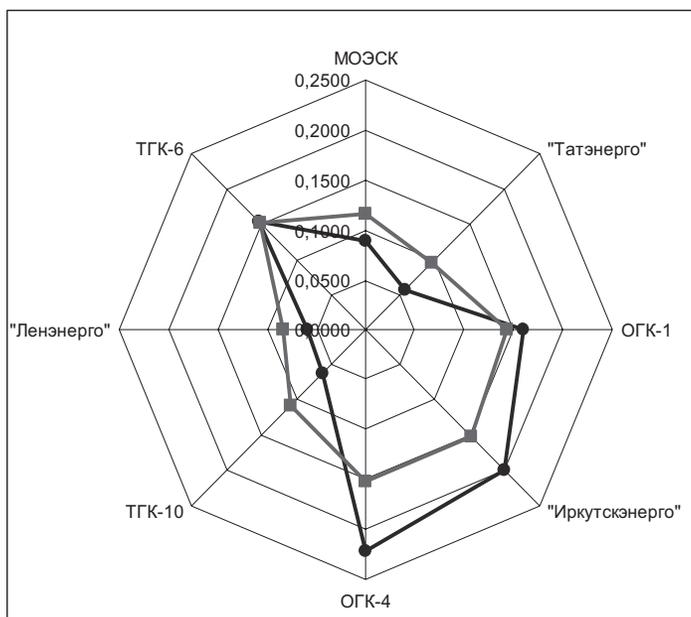


Рисунок 6 – Сравнение экспертных (круглые метки) и формализованных нормированных DEA-оценок (квадратные метки) эффективности

Выводы

Полученные результаты показывают возможности применения аналитического метода DEA и экспертных способов анализа эффективности сложных производственных систем МАИ к формированию частных показателей качества и свёртыванию их в интегральные критерии для оценки обобщённой эффективности.

Метод Data Envelopment Analysis позволил получить комплексные оценки эффективности, провести системную ранжировку и определить наиболее совершенные производства.

Метод анализа иерархий позволил синтезировать частные субъективные мнения экспертов, оценить качество их работы и получить интегральные оценки эффективности.

К ключевым особенностям метода МАИ относится то, что, во-первых, весовые коэффициенты оценочных критериев определяются экспертами, а не задаются априори. Во-вторых, рассматриваемый подход позволяет получать оценки на разных уровнях иерархии, что существенно облегчает задачу системного оценивания в случае большого количества объектов сравнения.

Сопоставление полученных различными методами оценок выявило их близость и отражает возможности совместного использования этих различных по сути методов.

Список литературы

- [1] HANDBOOK ON DATA ENVELOPMENT ANALYSIS edited by: William W. Cooper, Lawrence M. Seiford, Joe Zhu // Kluwer Academic Publishers, 2004 – 593 с.
- [2] Fukuyama, Hirofumi and William L. Weber, (2001), "The efficiency and productivity change of non-life insurance companies in Japan", Pacific Economic Review 6(1), pp. 129-146.
- [3] Aleskerov, Fuad, Hasan Ersel and Osman Reha Yolalan, (1999), "Multicriterial method for evaluating bank branch performance", Discussion Paper Series No: 97-01, Yapi Kredi Bank.

- [4] Thompson, Russell G., Parakramaweera Sunil Dharmapala, Juan Díaz, María Gonzáles-Lima and Robert M. Thrall, (1996), "DEA multiplier analytic center sensitivity with an illustrative application to independent oil companies", *Annals of Operations Research* 66, pp. 163-177.
- [5] Cook, Wade D., (2000), "Site selection for retail facilities using DEA", Presented at: North American Productivity Workshop [in DEA and weight restriction], Schenectady, United States of America.
- [6] Rouse, Paul, (2001), "Pricing public health care services using DEA: Methodology versus politics", Presented at: (OR 43) 43th Annual Conference of the Operational Research Society [in Data envelopment analysis], Bath, United Kingdom, 9/4 to 9/6.
- [7] Ozcan, Yasar A., (1995), "Efficiency of hospital service production in local markets: The balance sheet of U.S. medical armament", *Socio-Economic Planning Sciences* 29(2), pp. 139-150.
- [8] Breu, T. M. and Raymond L. Raab, (1994), "Efficiency and Perceived Quality of the Nation's Top 25 National Universities and national liberal arts colleges: An application of data envelopment analysis to higher education", *Socio-Economic Planning Sciences* 28(1), pp. 33-45
- [9] Выгон Г.В., Поманский А.Б. Анализ связи технологической эффективности и рыночной капитализации компаний // *Экономика и математические методы*, № 2, том 36, 2000, с. 79-87.
- [10] Кривоножко В.Е., Пропой А.И., Сеньков Р.В., Родченков И.В., Анохин П.М. Анализ эффективности функционирования сложных систем. // *Автоматизация проектирования*, №1, 1999. <<http://www.osp.ru/ap/1999/01/02.htm>>
- [11] Farrell M.J. The Measurement of Productive Efficiency // *Journal of the Royal Statistical Society, Series A (General)*, Vol. 120, Part III, 1957, 253 – 281 p.
- [12] Charnes A., Cooper W.W., Rhodes E. Measuring the Efficiency of Decision Making Units // *European Journal of Operational Research*, Vol. 2, 1978, pp. 429 – 444.
- [13] Banker R.D., Charnes A, Cooper W.W. Some Models for Estimating Technical and Scale Efficiency in Data Envelopment Analysis // *Management Science*, Vol. 30, No. 9, 1984, pp. 1078 –1092.
- [14] Сиразетдинов Т.К. Методы решения многокритериальных задач синтеза технических систем. – М.: Машиностроение, 1988.
- [15] Саати Т. Принятие решений. Метод анализа иерархий. М.: Радио и связь, 1993. – 320 с.
- [16] Саати Т., Кернс К. Аналитическое планирование. Организация систем. М.: Радио и связь, 1991.–224с.
- [17] Changhoon Shin, Jaeyoung Song, Hyunchan Kim DEA/AHP Model for Measuring Restricted and Effective Efficiency with Expert's Knowledge // *Informs&Korms* – 1140 – Department of Logistic Engineering, Korea Maritime University – Seoul 2000.
- [18] Z. Sinuany-Stern, A. Mehrez, Y.Hadad, An AHP/DEA methodology for ranking decision making units // *International Transactions in Operational Research* #7, 2000, pp. 109-124.
- [19] T. Yang, C.A. Kuo, A hierarchical AHP/DEA methodology for the facilities layout design problem // *European Journal of Operational Research* #147, 2003, pp. 128-136.
- [20] Рейтинг «Эксперт 200» // Рейтинговое агентство «Эксперт РА», <www.raexpert.ru>.

ПРЯМЫЕ И ОБРАТНЫЕ ЗАДАЧИ РАСПРОСТРАНЕНИЯ ТЕПЛА В ПОЧВЕ

Ф.Д. Михайлов

Сельскохозяйственный факультет университета Сельчук
42075, Конья, Кампус, Турция
farizm@selcuk.edu.tr
тел: +9 (0332)-223-29-34, факс: +9 (0332) 241-01-08

Ключевые слова: почва, температурный режим, тепловые свойства почвы, коэффициент температуропроводности, моделирование теплопереноса в почве

Abstract

The purpose of this study is to develop the average integral solutions of the heat transfer in soil, based on solving inverse problems, and in the rationale of the experimental methods on the basis of this decision, for determining the thermal conductivity of the soil.

Введение

Ученые разных дисциплин все большее внимание уделяют экологии и энергетике почвообразования. Несомненно, что солнечная энергия, приходящая на земную поверхность, а также атмосферные осадки играют основную роль в почвообразовании. Эти же факторы и определяют тепловой и водный режимы уже сложившихся почв. Совместное изучение тепловых и водных режимов почвы явиться новым звеном в энергетике почвообразования [1].

Температурный режим почвы играет значительную роль в процессах жизнедеятельности растений. От температуры существенно зависят скорости химических реакций, протекающих в почве и корнях, передвижение влаги в почве и в растении, газообмен в почве, растворимость минеральных солей и ряд других факторов.

Для всестороннего знания тепловых свойств почвы необходимо иметь также данные, с помощью которых можно было бы найти значения тепловых характеристик для данного состава и состояния почвы. Появившиеся в последнее время подобные исследования показывают, что для каждого типа почвы требуется отдельное определение тепловых свойств, желательно как в естественных, так и в лабораторных условиях.

Основными тепловыми свойствами почвы являются коэффициенты теплоемкости, температуропроводности (K), и теплопроводности. Знание этих характеристик почвы может приблизить разрешение такой острой проблемы современности как прогноз теплового режима почв. При обсуждении температурного режима почвы удобно пользоваться температуропроводностью, т.к. именно этот параметр характеризует «распространение» измеряемой величины-температуры [7].

К определению коэффициента K почвы, на основе использования решения упрощенного уравнения теплопереноса, посвящено немало работ. Все эти методы, основаны на анализе хода температуры почвы на разных глубинах и во времени, причем все они предполагают либо наличие периодического хода температуры в определенный период, либо реальный ход температуры.

Цель данной работы заключалась в разработке среднеинтегрального метода, позволяющего определить коэффициент температуропроводности почвы.

Постановка задачи

Нахождение температурного поля в многофазной капиллярно-пористой системе, каковой является почва, с учетом всех процессов кондуктивной теплопроводности, конвекции в порах, переноса влаги, является исключительно сложной задачей. Известно, что при моделировании почвенных процессов, протекающих в системе «почва-растения-атмосфера», основными этапами являются *идентификация* и *реализация* модели теплопереноса. Поэтому при системном подходе к анализу температурного режима почвы упрощают как уравнение теплопереноса, так и постановку краевых условий.

Решение прямой задачи

Решение различных прямых задач теплопереноса в почвах, подробно описаны в работах [2, 4, 6-9]. Так, например, для решения как прямой задачи теплопереноса в почве (прогноза переноса тепла в почве), так и обратной задачи (определения коэффициента K по данным полевых или лабораторных экспериментов), уравнение теплопереноса принимают в виде [1, 2, 4, 6-10]:

$$(1) \quad \frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial x^2}$$

и рассматривают его аналитические решения, полученные без начального и при периодическом граничном условии на поверхности, т.е.:

$$(2) \quad \text{при } x=0: \quad T(x,t) = T_0 + T_a \cdot \cos(\omega t + \varepsilon),$$

а также при условии, что температура почвы на нижней границе (в бесконечности) постоянна, т.е.:

$$(3) \quad \text{при } x \rightarrow \infty: \quad T(x,t) = T_0.$$

Здесь T – температура почвы в точке x в момент времени t ; k – коэффициент теплопроводности, T_0 – средне суточная (или годовая) температура деятельной поверхности почвы; T_a – амплитуда колебаний температуры деятельной поверхности почвы; $\omega = 2\pi / \tau_0$ – круговая суточная (или годовая) частота; τ_0 – период (*длина*) волны, выраженный в сутках или в годах; ε – сдвиг фазы, зависящий от начала отсчета времени.

Решение задачи (1)-(3), в безразмерных переменных имеет вид [4, 7, 8-9]:

$$(4) \quad T(y, \tau) = T_0 + \Phi_a(y, b) \cdot \cos[\bar{\omega}\tau + \varepsilon - \psi(y, b)]$$

где $y = x/L$, $\tau = \kappa t/L^2$, $b = L\sqrt{\omega/2\kappa}$ и

$$(5) \quad \Phi_a(y, b) = T_a \cdot e^{-by}, \quad \psi(y) = by.$$

Однако при выполнении практических расчетов нет возможности [2] в качестве исходных данных задать величины температуры почвы на бесконечности, так как они неизвестны. Поэтому в таких случаях вместо (3) задается условие на нижней границе в виде:

$$(6) \quad \text{при } x \rightarrow L: \quad \partial T(x,t) / \partial x = 0,$$

которое более реально описывает процесс теплопереноса.

При такой постановке граничного условия на нижней границе получены решения задачи теплопереноса, который также имеет вид (4), где

$$(7) \quad \Phi_a(y, b) = T_a \sqrt{\frac{\text{ch}(z) + \cos(z)}{\text{ch}(2b) + \cos(2b)}}, \psi(y, b) = \arctan \frac{P_2}{P_1}.$$

$$\left. \begin{aligned} P_1(y) &= \text{ch}(z) \cos(by) + \text{ch}(by) \cos(z) \\ P_2(y) &= \text{sh}(z) \sin(by) + \text{sh}(by) \sin(z) \end{aligned} \right\}, z = 2b(1-y)$$

Важным является также использование средней температуры почвы мощностью $[0, L]$, поскольку как и другие почвенные характеристики, среднепрофильная температура варьирует в меньшей степени, чем температуры на определенной глубине [3-4]. Поэтому её удобнее использовать при определении (κ) на основе экспериментов в полевых и лабораторных условиях. Для этого проинтегрируем (7) в промежутке $0 \leq y \leq 1$ и получим среднеинтегральное решение в следующем виде:

$$(8) \quad \bar{T}(\tau, b) = \int_0^1 T(y, \tau) dy = T_0 + M_a(b) \cdot \cos[\bar{\omega}\tau + \varepsilon - \hat{\psi}(b)],$$

где $M_a(b)$ и $\hat{\psi}(b)$ имеют вид:

$$M = \frac{T_a \sqrt{\text{sh}^2(2b) + \sin^2(2b)}}{\sqrt{2} b [\text{ch}(2b) + \cos(2b)]} \quad \hat{\psi} = \arctan \left[\frac{\sin(2b) - \text{sh}(2b)}{\sin(2b) + \text{sh}(2b)} \right]$$

Решение обратной задачи

На основе решения (8) выведена формула, которая позволяет определить *усредненный коэффициент K* для почвенного слоя $[0, L]$ для произвольного периода времени τ_0 в следующем виде:

$$(9) \quad \frac{[\bar{T}(t_3) - \bar{T}_1(t_1)]^2 + [\bar{T}(t_4) - \bar{T}(t_2)]^2}{4T_a^2} = \frac{\text{sh}^2(2b) + \sin^2(2b)}{2b^2 [\text{ch}(2b) + \cos(2b)]^2} = \Omega(b).$$

Здесь \bar{T} – усредненные значения температуры почвенного слоя $[0, L]$ в момент времени $t_i = i\tau_0 / 4$. На рисунке 1 приведен график зависимости функции $\Omega(b)$ для величин b , равных 0; 0,1; 0,2; ..., 4,0.

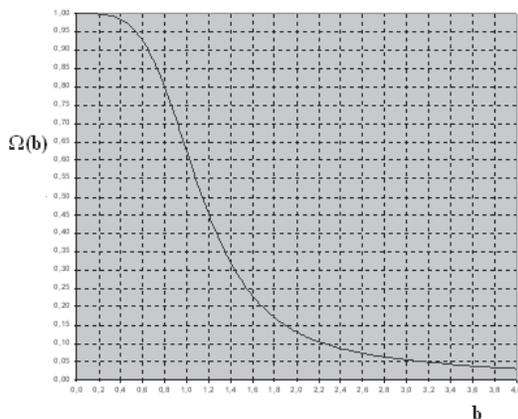


Рисунок 1 – График функции $\Omega(b)$ при граничных условиях $\partial T(L, t) / \partial x = 0$

Измерив усредненные температуры почвенного слоя $[0, L]$ через каждые *шесть часов в течение суток*, т.е., $T(t_i, b), (i=1, 2, 3, 4)$, следует вычислить $\Omega(b)$ по формуле (9) и, используя график, приведенный на рис. 1, найти значение $b = b^*$, и, далее K по следующей формуле:

$$(10) \quad K^* = \frac{\pi \cdot L^2}{\tau_0 \cdot b_*^2}.$$

Выводы

Предложенный средне-интегральный метод определения параметра K почвенных слоев является простым, надежным и удобным и дает возможность использовать его результаты в прогнозных математических моделях, используемых при оптимизации управления тепловым режимом естественных и искусственных почв.

Список литературы

- [1] Герайзаде А.П. Термо-и влагоперенос в почвенных системах.-Баку.: Элм, 1982.-159с.
- [2] Куртнер Д.А., Решетин О.Л. Об одном решении уравнения теплопроводности в связи с расчетом температуры почвы // Сб. трудов по агрономической физике, вып 26 / Теплообмен в открытом и защищенном грунте.-Л.: Гид.мет.изд., 1970.- С. 38-45.
- [3] Михайлов Ф.Д. Определение параметров модели солепереноса при промывках водонасыщенных поверхностно засоленных почвогрунтов // Почвоведение, № 5, 2007, С. 599-609.
- [4] Михайлов Ф.Д. Некоторые вопросы моделирования температурного режима почвы //Сб. Тр. конференции «16 Зимняя школа по механике сплошных сред», 24-27 Февраля 2009 г., Пермь.
- [5] Рычѐва Т.А. Моделирование температурного режима дерново-подзолистой почвы: определяющая роль условий на поверхности //Почвоведение, № 6, 1999, С. 697- 703.
- [6] Чудновский А.Ф. Теплофизика почв.- М.: Наука, 1976.-352с.
- [7] Шейн Е.В. Теории и методы физики почв. М.: Изд. «Гриф и К», 2005.- 616 с.
- [8] Juri W.A., Gardner W.R., Gardner W.H. Soil Physics.- New York, 1991.-328 p.
- [9] Horton, R.Jr. Determination and use of soil thermal proper ties near the soil surface. New Mexico State University, 1982.-151 p.
- [10] Marinova, T.K. On determining the conductivity coefficient of the basic soils in Bulgaria. Bulgarian Journal of Meteorology & Hydrology, No 2, 1993, pp. 65-69.

УСТОЙЧИВОЕ УПРАВЛЕНИЕ ЗНАНИЯМИ В КОМПАНИИ: ПРОБЛЕМЫ И ЗАДАЧИ

А.М. Прокофьев

Санкт-Петербургский государственный политехнический университет
195220, Санкт-Петербург, Гражданский пр., 28, Россия
alex@abz.ru
тел: +7 (812) 534-10-01 факс: +7 (812) 534-13-65

Ключевые слова: *устойчивость управления, оценка результатов, сложные социо-технические системы, проблемы управления знаниями.*

Abstract

Two decades after its inception, knowledge management shows increasing number of problems with sustainability of results. In this paper we study methods for assessing results of knowledge management projects and argue for their inconclusiveness. Based on recent empirical research, we highlight several current problems in the field. We provide new approaches for addressing these problems and outline tasks that will be done in further research.

Введение

В начале 90-х годов прошлого века, на основе опыта ведущих корпораций, были сформулированы модели управления знаниями. Первоначально большее внимание уделялось технической стороне и применению различных существующих информационных технологий для этой задачи. В 2000-х возросло внимание исследователей к социально-психологическим аспектам управления знаниями. Долгое время количество публикаций, предлагающих новые методы и подходы, существенно превосходило количество эмпирических исследований этих методов. Также, результаты, получаемые непосредственно после внедрения новых технологий, зачастую не удавалось сохранить в течение продолжительного времени. Возникла проблема устойчивости управления знаниями в компании. Современные эмпирические исследования [1, 2], выполненные на основе многолетних наблюдений, показывают важность и актуальность этой проблемы.

1 Обзор метрики и оценок результатов

Предлагается множество метрик для оценки результатов проекта по внедрению технологий управления знаниями. Некоторые из них опираются на финансовые показатели верхнего уровня, например, рост прибыли компании или сокращение расходов на поддержку. Другие, ближе подходят к самому предмету проекта, например количество поданных патентных заявок [3, 4, 5]. Значения таких метрик получаются на основе бухгалтерской и другой отчетности, которая формировалась в организации в течение длительного времени. Это позволяет сравнить изменение ряда показателей до и после проекта. Однако, на метрики верхнего уровня оказывает влияние большое количество внешних факторов. Например, рост прибыли можно объяснить изменением ситуации на рынке, рост числа патентов – завершением исследования, которое началось задолго до внедрения технологии УЗ. Поэтому, они не могут являться основным доказательством эффективности внедренной технологии.

Информационные системы, используемые в проектах УЗ, обычно предоставляют чисто технические метрики, которые дают возможность сделать лишь поверхностный анализ. Примером таких метрик являются: количество созданных документов, объем коммуникации между сотрудниками, количество поисковых запросов. На их основе можно сделать вывод о том,

используется система или нет. Однако, они не предоставляют информации об удовлетворенности сотрудников и не доказывают, что внедренная технология приносит ожидаемые результаты. Например, при переходе от совещаний к общению по электронной почте, можно измерять количество электронных сообщений, отправляемых сотрудниками. Рост этой метрики можно интерпретировать как успех проекта и положительный сдвиг. Но на самом деле могло произойти другое. Из-за отсутствия регулярных совещаний, сотрудники стали по-разному представлять рабочий процесс и задачи, стоящие перед группой. Обсуждение новых предложений стало проходить не всей группой, а парами отдельных участников, что только усугубляло различия в понимании задач коллектива. А в возросшем потоке электронных сообщений, обсуждались разногласия и недопонимания, которые возникли в результате потери общей точки зрения, которая ранее формировалась с помощью регулярных совещаний. Таким образом, и чисто технические метрики не пригодны для оценки технологии УЗ.

В сложных социально-технических системах, часто используется метод анкетирования для контроля результативности действий. В области управления знаниями наблюдается сильная поляризация мнений. Опросы, проводимые среди руководителей высшего звена, говорят о высокой эффективности реализуемых в области управления знаниями инициатив [6, 7, 8]. В то же время, анкетирование рядовых сотрудников показало диаметрально противоположенные результаты [2, 1]. При этом нужно учитывать, что именно рядовые сотрудники являются основными пользователями анализируемых систем и технологий. Субъективность мнений в данном случае дополняется материальной заинтересованностью руководителей и консультантов в успешности проектов. Поэтому в большей степени следует доверять результатам, полученным при анкетировании сотрудников, непосредственно вовлеченных в процессы обмена знаниями.

Предлагается разработать новый метод оценки эффективности обмена знаниями в компании, который будет совмещен с системой мотивации сотрудников и при этом будет реализован техническими, а не организационными средствами. Для того, чтобы поставить задачу на разработку такого метода, дополнительно необходимо проанализировать подходы к вовлечению сотрудников в рассматриваемый процесс.

Мотивация персонала при обмене знаниями является обширной темой для исследований. Ранние работы использовали методы экономической и организационной теории для этой цели. Однако, еще в 1993 году Kohn аргументировано показал неэффективность материальных наград, как способов устойчивой стимуляции деятельности. В ряде случаев, раздача значительных вознаграждений за достижение метрик могла существенно ухудшить реальные результаты работы и снизить инновационный потенциал [9]. Его предположения были подтверждены в эмпирическом исследовании в области обмена знаниями (Vock и Kim). Исследование показало, что для сотрудников мотивирующими факторами являются: ожидаемый вклад в общее дело компании и возможность улучшить отношения с коллегами в процессе обмена знаниями. В тоже время, было опровергнуто предположение, что материальные награды усилят это стремление. Для опытных сотрудников обмен знаниями является естественной необходимостью и способом самореализации. Выплата вознаграждений воспринималась как попытка подкупа и манипуляций, в результате которой обмен знаниями только сокращался [10]. Принуждение сотрудников к обмену знаниями со стороны руководства также показало себя как неэффективный метод стимуляции, который приводит к противоположенным эффектам [11].

2 Определение проблемных областей

Учитывая результаты предыдущих исследований и наблюдений, было выделено несколько препятствий для устойчивого применения процесса управления знаниями с использованием информационных систем.

2.1 Необходимость "лишнего шага"

В исследовании Ralph и Ellis рассматривалось использование системы управления знаниями для оптимизации работы библиотек. Подразумевалось, что библиотекари будут делиться своими находками и вносить их в систему. Это должно было сократить время обработки повторных запросов и сократить дублирование работы. Сотрудники были обучены, и умели пользоваться рассматриваемой системой. Также они представляли, какие выгоды это принесет для их организации. Результаты оказались неожиданными для авторов проекта: 96% библиотекарей системой не пользовались. Среди причин такого результата респонденты указали: необходимость дополнительного шага в работе (68%), значительные временные затраты на работу с системой (48%). При этом только 19% считали инструмент недостаточно удобным [2].

2.2 Неоправданная или несвоевременная формализация

Некоторые теоретики считают, что обмен знаниями должен быть строго регламентирован, знания должны быть кодифицированы и проверены экспертом в данной области. Они предъявляют жесткие требования к качеству знаний. Такой подход хорошо работает в научном сообществе, однако в корпоративной среде он вызывает ряд проблем. Основная из которых – существенное замедление процесса. Далеко не все сотрудники, обладающие знаниями, являются аналитиками высокого уровня или учеными и могут выполнить грамотную кодификацию и рубрикацию. Привлечение к этой задаче профессиональных аналитиков замедляет процесс только сильнее. Исследователи, пропагандирующие этот подход, пытаются обобщить использование метода экспертных систем на всю компанию, игнорируя область его применимости. В современной компании, новые знания появляются и устаревают с высокой скоростью. В результате, за время их тщательного улучшения они могут полностью потерять свою практическую ценность. Эта проблема отмечается многими авторами, однако они не предоставляют методологии, указывающей, каким образом следует выбирать уровень формализации, оставляя это на откуп консультантам [6, 12].

2.3 Игнорирование существующих способов работы с информацией в компании

Каждая организация работает с информацией в своем стиле. Некоторые все еще ориентируются на бумажный документооборот и совещания, другие частично перешли на электронные формы совместной работы. Однако, уровень информатизации в отечественных компаниях остается невысоким и обычно ограничивается базовыми офисными пакетами и системами бухгалтерского учета. Поэтому не оправдываются расчеты, что большинство сотрудников в компании единомоментно примут новую парадигму работы с информацией. Как показало многолетнее исследование Wilensky, они не принимают ее даже после тщательного обучения, и длительных попыток по вовлечению персонал в новый подход к работе с информацией и обмену знаниями с помощью информационной системы [1].

Использование чисто социальных методов управления знаниями также не принесло ожидаемых результатов. Ряд авторов предлагал использовать концепцию сообществ практиков для организации эффективного обмена знаниями [13]. Не смотря на то, что авторы строили свои подходы на основе принципов социологии и психологии, ожидаемые результаты оказались не достигнуты. Рассматриваемые сообщества в итоге занимались не обменом знаний, а рассуждениями о необходимости такого обмена [14].

3 Цель и задачи

Цель исследования: разработка новой технологии устойчивого управления знаниями в компании. Учитывая результаты предыдущих исследований, были сформулированы следующие задачи:

3.1 Динамический выбор оптимального уровня формализации

Необходимо определить востребованные уровни формализации, сформулировать критерии, характеризующих каждый из них. Разработать технологии перехода между различными уровнями, а также предоставить оценку трудоемкости таких переходов. Сформулировать метод выбора корректного уровня формализации представления области знаний (на основе результатов ее использования и заданных ограничений). Предполагается, что уровень формализации может не только повышаться со временем, но и понижаться. Например, при росте изменчивости внешней среды, трудоемкость строгой формализации и структуризации новых знаний может перестать быть оправданной.

3.2 Технология микро-мотиваций

Для решения проблемы устойчивого вовлечения сотрудников в процесс обмена знаниями будет разработана технология микро-мотивации. В ее основе лежит принцип незамедлительного поощрения сотрудника за небольшое, но полезное действие. Для его реализации будет сформулировано большое количество достижений – осязаемых и измеримых вкладов в процесс обмена, накопления и улучшения знаний. Разрабатываемый метод будет отличаться от традиционных подходов к мотивации (например, определение лучшего сотрудника раз в квартал) рядом характеристик. Достижением будет признаваться не только качественное изложение знаний, но и их дополнительная структуризация, оценка, а также своевременность этих действий. Планируется разработать несколько десятков таких достижений, которые будут ориентированы на различные навыки и ценности сотрудников. Реализация такой технологии невозможна чисто организационными средствами, так как она требует постоянного детального мониторинга и анализа всех действий сотрудников. Поэтому, будет разработан метод определения нетривиальных действий пользователя в информационных системах.

3.3 Метрики для изолированной оценки новой технологии

Для оценки эффективности предлагаемой технологии будет использоваться статистический анализ достижений, которые были получены сотрудниками в рамках микро-мотиваций. Эти метрики будут с одной стороны соответствовать воспринимаемой выгоде от использования технологии, а с другой будут достаточно изолированными от воздействия внешних факторов.

3.4 Реализация и способы проверки

В ходе исследования на основе онтолого-семантической модели [15] будет разработана система управления знаниями, которая реализует описанные задачи. С ее помощью будут проверены предлагаемые технологии и методы.

Заключение

В этой статье мы изучили распространенные способы оценки результатов внедрения технологии управления знаниями. Рассмотренные метрики относились к техническим, экономическим и социологическим областям. Была отмечена их недостаточная проработанность и показательность. В частности, рассмотренные метрики были либо подвержены влиянию внешних факторов, либо не предоставляли однозначную оценку полезности технологии. По результатам недавних эмпирических исследований, мы выделили ряд текущих проблем, препятствующих устойчивому управлению знаниями в компании. Среди них: неоправданная и несвоевременная формализация, необходимость "лишнего шага", игнорирование существующих способов работы с информацией в компании, некорректная мотивация. Для разрешения обозначенных проблем мы сформулировали цель и задачи будущих исследований.

Список литературы

- [1] Wilensky Hiroko, Redmiles David и Su Norman Makoto The Dissemination of Knowledge Management[Конференция] // Материалы конференции "ACM 2009 international conference on Supporting group work". - 2009.
- [2] Ralph L. L. и Ellis T. J. An Investigation of a Knowledge Management Solution for the Improvement of Reference Services[Статья] // Journal of Information, Information Technology, and Organizations. - 2009 г.. - Т. 4.
- [3] Лукичева Л. И. Управление интеллектуальным капиталом[Книга]. - Москва : Омега-Л, 2007.
- [4] Liebowitz J. и Suen C. Y. Developing knowledge management metrics for measuring intellectual capital[Статья] // Journal of Intellectual Capital. - 2000 г.. - Т. 1.
- [5] Grossman M. An overview of knowledge management assessment approaches[Статья] // Journal of American Academy of Business. - Cambridge : [б.н.], 2006 г.. - 8.
- [6] Jennex ME и Olfman L A Model of Knowledge Management Success[Раздел книги] // Current Issues in knowledge management. - [б.м.] : Information Science Reference, 2008.
- [7] Прусак Лоренс и Лессер Эрик Как превратить знания в стоимость[Книга]. - Москва : Альпина бизнес-букс, 2006.
- [8] С. Curado Perceptions of knowledge management and intellectual capital in the banking industry.[Статья] // Journal of Knowledge Management. - 2008 г.. - 141 : Т. 12.
- [9] Kohn A. Why incentive plans cannot work[Статья] // Harvard business review, 1993 г..
- [10] Bock G. W. и Kim Y. G. Breaking the myths of rewards: An exploratory study of attitudes about knowledge sharing[Статья] // Information Resources Management Journal. - 2002 г.. - Т. 15.
- [11] Harder M. How do rewards and management styles influence the motivation to share knowledge?[Доклад]. - [б.м.] : Center for Strategic Management and Globalization, 2008.
- [12] Davenport Thomas, Long David и Beers Michael Building Successful Knowledge Management Projects[Статья] // Managing the Knowledge of the Organization. - [б.м.] : Ernst & Young LLP, 1997 г..
- [13] Harvard Business Review Управление знаниями[Книга]. - Москва : Альпина Бизнес Букс, 2006.
- [14] Su NM [и др.] The gospel of knowledge management in and out of a professional community[Конференция] // Proceedings of the 2007 international ACM conference. - 2007.
- [15] Тузовский Анатолий Федорович Онтолого-семантические модели в корпоративных системах управления знаниями // Диссертация на соискание ученой степени доктора технических наук. - Томск : ТПУ, 2007 г..

КОГНИТИВНАЯ ТЕХНОЛОГИЯ УПРАВЛЕНИЯ. ТЕХНОЛОГИЯ СО-УПРАВЛЕНИЯ, МОДЕЛЬ ДЛЯ УПРАВЛЕНИЯ

Д.И. Пунда

ООО «НАУКОМП - Технологии Моделирования»

443010, Самара, ул. Самарская, 146, Россия

poundadi@yandex.ru

тел: +7 (846) - 246-40-30, +7 (916) -571-63-17 (Москва), тел/факс: +7 (846) - 221-19-13

Ключевые слова, основные понятия: технологии управления, технологии информационного обеспечения управления, модель для управления, модель управления для организационных систем, архитектура деятельности организационных систем, сложная организационная система, интеграция знаний, формализация знаний, открытые технологии, открытые модели, закрытые технологии, закрытые модели

Abstract

The effect of both individual and collective management resource on a complex organization systems management is analyzed. It's shown that the objective (natural) restrictions for possibilities of the open-modeling of multi-technological organization systems (or non-additivity of open-modeling of complex organization systems) are one of major reasons for weak manageability and control in complex organization systems. The co-management technology (i.e. a cognitive control-and-management technology) is suggested as an artificial method for the open-modeling possibility increase for multi-technological organization systems, and as a complex organization systems management crisis compensator.

Введение

Современный прогресс количественно можно охарактеризовать тем, что каждые три года в мире удваивается число технологий человеческой деятельности (по оценкам отдельных экспертов такое удвоение происходит каждый год) [1]. Однако изученность, определенность создаваемых технологий зачастую оставляет желать лучшего. Одной из основных причин слабой изученности современной деятельности является то, что в условиях нынешней, слабоуправляемой формы свободного рынка, системный (определенный, изученный) продукт менее приоритетен, чем быстрый финансовый результат.

Такое большое и растущее разнообразие технологий деятельности с вероятной их низкой системностью может порождать серьезные сложности в управлении организационными системами любого уровня – от конкретных организаций до отраслей, экономик на государственном уровне. Это согласуется с тем, что экономические кризисы первой половины прошлого века имели относительно общий характер – они происходили на уровне государств, при достаточно управляемой деятельности отдельных предприятий [2,3], а в последние десятилетия все более вероятными становятся локальные кризисные процессы (как, например, известный кризис управления GM в 90-е годы на фоне относительно спокойного общего рынка).

Косвенным подтверждением растущего усложнения предметной управляемости современной деятельности в обществе является заметное увеличение времени согласований понятий и терминов при регламентированных взаимодействиях в организационных системах, особенно на конференциях, деловых или политизированных встречах и совещаниях с участием сторон из многих отраслей [4].

Кроме многотехнологичности и низкой системности современной деятельности, конечно, есть и другие причины, оказывающие влияние на управление. Например, это политические,

организационные, финансовые, конъюнктурные вопросы и проблемы, проблемы конкуренции и иные общественно-деловые аспекты, которые всегда присутствовали, и будут присутствовать. Но эти вопросы и проблемы принципиально могут регулироваться согласованиями, кооперациями, введением норм и другими технологиями делового, исполнительно-правового и социального взаимодействия организационных систем. Поэтому считать, что они в основном могут быть причинами возникновения современных проблем управления, будет не корректно.

Дальнейшее “эволюционное” развитие данных “технологий регулирования” проблем и отношений, конечно, целесообразно и явно действенно. Но такое развитие не дает уверенности в достижении полного (хотя бы удовлетворительного) контроля над кризисными процессами, в решении проблем управления. Это постоянно подтверждает практика применения различных методов согласования условий и интересов (от внутрикорпоративных до межотраслевых и межгосударственных), практика применения различных способов коопераций и сотрудничества, практика совершенствования нормативных баз и т.п. [4-6].

В последнее время участились публичные утверждения того, что нынешний экономический кризис по своей сути является кризисом управления, что нужны «правильные модели управления, которых пока нет» [4,7]. Появляются и другие предложения повышения действенности коллективного и индивидуального управления. Так в мире постоянно совершенствуются теории экономики и управления, развиваются информационные технологии (ИТ) решения различных задач проектирования, планирования и учета, системы поддержки принятия решения (DSS) и иные *технологии информационного и коммуникационного обеспечения управления* [8]. Они явно улучшают предметную организацию деятельности и увеличивают степень определенности и формализуемости технологий деятельности, дают представление об экономике в целом и частном, объясняют “прошлое” и явно упорядочивают процессы управления на любых его уровнях. Однако они либо относительно скоро теряют свою прогностическую силу, либо постепенно теряют актуальность своего использования в управлении, либо, просто, их применение не приводит к удовлетворительному уменьшению вероятности возникновения общих и локальных кризисов (частота и масштабы которых постоянно растут).

То есть «зарекомендовавшие себя» технологии, направленные на улучшение управления организационными системами пока не дают полной уверенности в их эффективности. Предметом настоящей работы является практический анализ естественных возможностей индивидуального и коллективного управления (на уровне анализа мышления при управлении), определение влияния этих возможностей на проблемы управления организационными системами, построение ментальной (“мыслительной”) структуры процесса управления, и, на основе этого, создание технологии управления, искусственно увеличивающей возможности индивидуально-го управления, качественно повышающей управляемость организационными системами.

1 Сложные и «простые» организационные системы

В частных компаниях владельцы или их представители-управленцы стараются не усложнять свою деятельность по технологической насыщенности и неопределенности используемых технологий, то есть обеспечивать ее управляемость (например, обеспечивать устойчивые возможности руководителя формализовать любые свои решения до уровня, однозначно воспринимаемого и рассчитываемого его заместителями и исполнителями). Но во многих организационных системах, корпоративных структурах, и даже не только государственных, такое «упрощение» деятельности может быть нежелательным или даже невозможным. Коллективное управление и перераспределение управления по “иерархической лестнице” уже далеко не всегда эффективно для управления многотехнологичными организационными системами [6].

Тогда актуальна постановка такого вопроса, - какие причины или механизмы процесса управления, усложняющие управление организационной системой с ростом технологий ее деятельности, можно устранять урегулированием и информационно-коммуникационным

обеспечением только при числе технологий деятельности не больше определенного их количества? И есть ли вообще такие причины, которые с ростом технологий выше этого определенного количества меняют качество управления и делают его принципиально сложным?

С одной стороны, любая организационная система вправе считаться сложной по причине ограниченной предсказуемости принятия решения ее руководителями разного уровня. С другой стороны, известно, что человек может в среднем эффективно управлять не более 6-ти подчиненными, плюс-минус (в зависимости от субъективных возможностей, от уровня используемых интеллектуальных технологий, от степени системности деятельности подчиненных подразделений, от системности используемых в них технологий, от количества технологий, отнесенных к классу полностью определенных, формализованных). В работах Миллера показано практически, что человек в состоянии одновременно воспринимать не более “7 плюс-минус 2” разнородных целостных объектов (субъектов, организационных систем) [5]. Что, даже разбираясь в «ста» технологиях, он может эффективно контролировать не более тех же “7 +/- 2” взаимосвязанных, не полностью формализованных технологий. То есть индивидуальные возможности управления многотехнологичной деятельностью ограничены.

С еще «другой», третьей стороны, если на место директора предприятия “посадить” коллектив из 10 управленцев, то они смогут эффективно управлять не 60-ю подчиненными, а реально практически все теми же 6-ю. То есть возможности коллективного управления тоже ограничены, причем не количеством управленцев. И потенциальная эффективность коллективного управления качественно не много выше, чем при индивидуальном управлении [6].

Поэтому причину неуправляемости сложными организационными системами, не устраняемую согласованиями, информационными и коммуникационными обеспечениями, нужно искать в ограниченных возможностях индивидуального управления.

И мы вправе разделить современные реальные организационные системы на относительно простые в управлении системы и сложные (разделить качественно и относительно количественно). И для эффективного управления в сложно-управляемых организационных системах требуются отдельные специальные технологии - инструменты для руководителя (дополнительные к технологиям информационного обеспечения управления, и увеличивающие индивидуальные возможности руководителя). Такие дополнительные технологии управления могут быть вполне не актуальными для «простых» организационных систем.

На практике нам приходилось иметь предметные контакты с руководителями и специалистами достаточно крупных, но простых с точки зрения управления предприятий (Энергосбытовых, Коммунальных, структур Газпрома, Нефтяных компаний). Номенклатурная, клиентская насыщенность предметных технологий в таких компаниях может быть высокой, что даже самые мощные современные системы сквозного учета (например, IT-системы интегрального учета класса ERP SAP R-3) не всегда справляются с информационной поддержкой баз данных. Но при этом такие компании имеют относительно небольшое количество предметных видов деятельности или разно-технологичных подразделений. И, как показала наша практика, внедрение технологий управления в них носит, в основном, консультационный характер, не требующий создания постоянно используемого инструментария для руководителя.

И имели место прямо противоположные примеры. Даже примеры компаний с меньшими объемами капитализации и оборотами средств (чем в вышеуказанных предприятиях), но в активе которых (компаний) имелось, например, более 10-ти разно-отраслевых производственно-технологических подразделений, несколько проектно-конструкторских и строительных отделов, крупное транспортно-складское хозяйство, технологический ресурсо-добывающий и перерабатывающий комплекс. И все эти подразделения были значимы для бизнеса – их не могли передать в доверительное управление или отнести к классу полностью определенных, формализованных технологий (как например отдел кадров, бухгалтерию или охрану). В таких компаниях, оснащенных самыми передовыми информационными технологиями, демонстрирова-

лась устойчивая заинтересованность руководителя в инструменте для управления, на всех этапах деятельности по созданию технологии управления [7].

2 Неаддитивность открытых моделей и технологий организационных систем

Под открытой моделью организационных систем понимается их представление, включающее все внутренние и внешние понятия (представления объектов и явлений) и связи организационной системы и представления всех используемых технологий. Под открытыми технологиями мы понимаем способы и средства, учитывающие все внутренние и внешние объекты, явления и связи организационных систем и технологий [7].

На протяжении своего развития человечество изучало различные объекты и явления, находило и выделяло закономерности и особенности, строило объяснения и множественные теории их описывающие, и технологии с ними “работающие”. Формализованные теории и технологии появлялись в результате абстрагирования от “невыведенного” в закономерностях окружения. С использованием этих теорий и технологий строились формализованные представления объектов и явлений, как целого в частном. В этом смысле такие формализованные теории, технологии и представления являются закрытыми, отделенными от окружения. И чем они более абстрагированы от внешнего окружения, тем они более закрыты.

Человеческая система мышления в состоянии накапливать, суммировать знания, заключенные в формализованных теориях, технологиях и представлениях целого в частном. В таком понимании эти *формализованные категории аддитивны*. Отдельные “полиглоты” при желании и усилиях осваивают десятки наук, языков, технологических процессов. Коллективы могут использовать формализованные знания в полном располагаемом и суммированном ими объеме, даже не создавая для этого электронные или тексто-графические версии.

В то же время, в процессе развития человечества и даже на протяжении жизни отдельно человека, мы, как правило, имеем дело с небольшим числом одновременно взаимосвязанных субъектов, объектов и явлений. Это объясняет то, что “руководитель способен эффективно управлять только ограниченным числом подчиненных”.

Управление предполагает поиск вариантов решений, и направлено на совершение действий. Поэтому в развитом *подходе познания, как представления целого в частном и конструирования формализованных знаний о частном, для управления в основном актуальны знания о технологиях* (здесь знания – это представления о практике, позволяющие делать суждения [9]). Кроме технологий важными для управления организационными системами являются *растановка приоритетов “частного” в представлениях этих систем и степени определенности, формализации знаний о них (в основном о технологиях)* [7].

Как сказано в предыдущем разделе (п.п. 1), индивидуальные возможности строить представление взаимосвязанных, не полностью формализованных категорий ограничены количественно. Такими слабоформализованными категориями являются открытые технологии и открытые представления-модели организационных систем, с учетом возможной *неустойчивости приоритетов и не полной формализации знаний об этих системах в процессе управления*.

Поэтому при создании открытых моделей для эффективного управления комплексными организационными системами, состоящими из нескольких организационных подсистем, ограничение представления слабоформализованных категорий не позволит построить комплексную модель путем простой состыковки, “суммирования” моделей подсистем. Модель комплексной системы здесь строится с новыми комплексными приоритетами и формализацией. То же справедливо и для представления комплексных открытых технологий. Те подсистемы или технологии, которые уже сложно учесть или/и имеют меньшую приоритетность, руководитель “переводит” в разряд закрытых формализованных технологий и/или передает в доверительное управление замам, а сам уже подробно и постоянно, как правило, не контролирует.

Аналогичная ситуация будет и при создании открытой модели для эффективного управления организационной системой путем суммирования знаний о ней “нескольких руководителей коллективного управления” и специалистов. Ограничение представления слабоформализованных категорий в лучшем случае позволит построить одинаковую для каждого и для всех, оптимальную (предельно наполненную и “отточенную”) открытую модель, но с теми же “приблизительно шестью открытыми подсистемами или технологиями”.

В этом понимании открытые модели организационных систем и представления открытых технологий, используемых в них не аддитивны (не суммируются “неограниченно” при коллективном управлении, в отличие от аддитивности закрытых формализованных моделей и технологий, таких, например, как универсальные ИТ и формализованные DSS).

Поэтому неаддитивность открытых моделей и технологий является причиной сложности контролируемого управления многотехнологичными организационными системами. И компенсировать такую причину урегулированиями (как это принципиально допустимо для причин конъюнктурного, организационного, финансового, криминального и т.п. характера) или системами информационно-коммуникационного обеспечения принципиально невозможно.

3 Управление

Устоявшаяся и общепринятая толковательная интерпретация управления кратко звучит, как «поддержание и урегулирование» [9].

Если управление рассматривать, как деятельность, то, раскрывая его «толковое определение» для практических приложений [7], оно будет совокупностью следующих процессов:

- расстановка приоритетов, ценностей в деятельности организационной системы (выделение главного),
- построение и постановка задач,
- прогнозирование и контроль,
- согласование интересов сторон как внутри организационной системы, так и вне нее.

4 Технология управления

Под технологией управления понимаются способы и средства, увеличивающие индивидуальные возможности управления и предназначенные для прямого функционального использования руководителем [7]. Это совершенно не означает, что техническая работа с технологией управления будет производиться самим руководителем (для этого могут использоваться услуги секретаря-оператора, аналогично услугам офисного секретаря), и не значит, что руководителю необходимы дополнительные технические знания и навыки. Входными данными для технологии управления являются варианты управленческих решений, конструируемые руководителем в принятых для руководителя понятиях и терминах, а выходными – результаты возможной реализации этих решений.

Если технологии разделить на технологии постановки задач и технологии решения задач, то технологии управления относятся к технологиям постановки задач. Поскольку для технологии решения задач можно формализовать исполнительскую регламентированную функцию.

Являясь технологией постановки задач, такие способы и средства не обязательно полностью формализованы. Не полностью формализованным является и используемое в технологии управления моделирование деятельности – модель для управления, которая соответствует представлению руководителя о деятельности организационной системы.

Полная формализация варианта управленческого решения и возможная дополнительная формализация самой технологии управления (и модели для управления) происходят в каждом конкретном случае. Такая формализация может зависеть не только от воли руководителя, но и от принимаемых решений на остальных уровнях управления, от приоритетов и от конкретной формализации технологий деятельности, которые могут быть полностью не формализованы.

Современные, традиционно используемые интегральные и специальные формализованные ИТ-технологии являются технологиями решения задач. Прямыми функциональными пользователями таких инструментариев являются исполнители, решающие поставленные руководителем задачи. И поэтому такие технологии, способные выполнять функции информационного обеспечения управления, не могут выступать в качестве технологий управления. Технологии информационного обеспечения управления и технологии управления не исключают друг друга, а способствуют развитию каждого. Но они отличаются по целям и средствам.

5 Варианты решения современных проблем управления

Интерес к развитию научных и прикладных способов и средств управления организационными системами заметно оживился «благодаря» сегодняшнему кризису. Попытаемся структурировать основные подходы к качественному повышению эффективности управления организационными системами, которые (подходы) “на слуху”.

- 1) Совершенствование создания и внедрения ИТ-технологий и соответствующих подразделений специалистов для решения специальных и комплексных задач.
- 2) Совершенствование информационного и коммуникационного обеспечения управления в форме утвердивших себя систем поддержки принятия решений.
- 3) Развитие управленческого консалтинга.
- 4) Поиск новых моделей управления (функциональных и организационных структур, архитектур управления). От микроэкономических моделей до моделей общеэкономических.
- 5) Упрощение деятельности (разбиение сложной организации на “простые” предприятия).
- 6) Развитие методов коллективного управления.
- 7) Создание искусственного интеллекта.

Какой вариант (или какая их совокупность) наиболее правильный? Возможно, таковой есть, но отличный от этих?

Ясно одно – максимально возможная управляемость сложной организационной системой реализуется в современных условиях при лучших ИТ-технологиях, учитывающих всё, информирующих обо всем необходимом и решающих любые поставленные актуальные задачи. Для оперативного, согласованного и точного информационного и коммуникационного обеспечения управления необходимы лучшие системы поддержки принятия решений. И конечно для эффективного управления нужна лучшая модель - архитектура организации управления. То есть проблему управляемости организационной системы нужно решать системно – взаимосвязано использовать все возможные интеллектуальные технологии. Это все логично и практически привлекательно, и реализуется далеко не в единичных случаях. Но пока, как изложено выше, даже такая логичная интеллектуальная интеграция, не всегда обеспечивает стабильную управляемость организационными системами, отраслями, рынками.

Для предметного анализа и выбора технологий управления предлагается детально представить структуру процесса управления, и выделить в ней те составляющие, которые влияют на возможности индивидуального и коллективного управления.

6 Структурирование управления. Со-управление

Мы вправе формализовать процесс управления в том объеме, который позволяют современные практические знания (понимая, что эти знания пока не позволяют нам полностью формализовать и «считать» процесс человеческого мышления, лежащий в основе управления организационными системами).

Подход к представлению целого, как композиции частного, предполагает структурирование на «три кита» - множества, ограничения-правила и действия [10]. Тогда с учетом изложенного в п.п. 1 и 2 структуру процесса управления можно представить следующим образом [7] (см. рисунок 1):

- Информационно-подготовительная составляющая управления:
 - Интеграция знаний (в основном знаний о технологиях)
 - Расстановка приоритетов — выделение главного (расстановка ценностей явных, неявных, их изменений — неявные те, которые не могут формализовать, определить или не хотят делать явными).
 - Определение степени формализуемости представлений параметров деятельности (например, их понятий и связей).
 - Постановка задач (например, на получение информации)
 - Получение новых знаний (например, генерация знаний на основе полученной информации или/и самостоятельного анализа)
 - В итоге — построение целостной картины, модели деятельности, с учетом не полной ее определенности (построение слабоформализованной модели деятельности).
- Творческая, решающая составляющая мышления:
 - Генерация вариантов управленческих решений (в том числе не полностью формализованных).
 - Проигрывание вариантов решений.
 - Постановка задач (для проигрывания вариантов управленческих решений, для получения новой информации).
 - Получение новых знаний.
 - В итоге — выбор варианта решения.
- Все эти составляющие при управлении могут идти в любом порядке и количестве.

Проанализируем, какие из этих составляющих индивидуального мышления при управлении, можно выделить и «безболезненно передать» помощнику-соуправленцу для самостоятельной деятельности, но не для решения задач информационного обеспечения, поставленных руководителем (*соуправленец, то есть вторая сторона индивидуального управления – это технология управления и возможный коллектив ее реализующий*). Причем выделить и передать именно как часть единого управления, - разделить единый процесс индивидуального управления. Такая передача части управления позволит коллективу из двух неравнозначных сторон – руководителю и соуправленцу, увеличить возможности управления большим количеством подчиненных технологий, взаимосвязанных в единую систему. Такое разделение управления сможет реализовать управляемое взаимопонимание для коллектива из двух сторон.

То, что точно и заведомо нельзя выполнять за руководителя, это генерацию вариантов управленческих решений. Постановку задач. Выбор варианта решения (принятие решения). Неявные ценности, неявную формализуемость, и неявные знания (см. рисунок 2).

Построение целостной взаимосвязанной картины деятельности путем интеграции знаний – технологий и их связей, путем расстановки явных (точнее, в основном явных) приоритетов и их ситуационных изменений, определение степени формализуемости параметров деятельности – это то, что может быть выполнено и отслеживаться отдельно от руководителя, но соответствовать его представлению. И с построением считаемого сложного калькулятора (программной версии модели для управления), соуправленец, выполняющий эти составляющие управления (причем достаточно информационно насыщенные), может эффективно проводить:

- Работу по формализации формулируемого директором варианта управленческого решения, по согласованию в конкретной ситуации со всеми службами и замами, с внешними факторами и т.п.
- Проигрывание вариантов управленческих решений с получением результатов их предполагаемой реализации
- Корректировку решений, планов использования ресурсов, финансовых и других планов, и их согласование.

- Составление приоритетов деятельности, и, по необходимости, их согласование, в том числе с директором, но уже в подготовленном, укрупненном виде.
- Системное отслеживание изменений приоритетов и технологий в деятельности организации и внесение изменений в модель.



Рисунок 1 – Составляющие процесса управления



Рисунок 2 – Составляющие процесса управления. То, что можно передать соуправленцу «безболезненно для индивидуального управления» (составляющие без слова “НЕТ”)

7 Слабоформализованная модель деятельности. Модель для управления

Согласно выше изложенному, для технологии управления необходимо построение интегральной модели деятельности, соответствующей представлению руководителя о его деятель-

ности. Такая модель не является полностью формализованной (и может формализоваться в каждом конкретном случае). Она интегрирует все технологии деятельности. Учитывает расстановку приоритетов-ценностей, выделенного главного, их изменения. Соответствует практической степени формализации понятий и связей деятельности. Включает в себя все внешние взаимодействия с учетом их изменчивости, что является обеспечением создания открытой технологии управления. В отличие от универсальных ИТ-систем учета и систем поддержки принятия решений, такая модель для управления оперирует не с подробными данными о деятельности, а с укрупненными, опять же соответствующими представлениям руководителя.

8 О технологиях со-управления

Построению модели для управления, методологически, предшествует построение интегральной модели экономики в целом (модели “мира” с учетом макроэкономических особенностей и тенденций и с учетом микроэкономических параметров, актуальных для деятельности моделируемой организационной системы). Затем следует моделирование обеспеченности деятельности организационной системы в представлении модели экономики в целом (т.е. моделирование основных противоречий и их развития; это позволяет структурировать деятельность и определить качество работы подразделений и организации в целом, и в текущее время, и стратегически, в тенденциях). Дополнительно детализируется структура управления (п.п. 6) для конкретной организационной системы (внутренние технологии деятельности подразделений организации и технологии внешнего взаимодействия делятся на те, которые можно считать полностью определенными, формализованными и на слабоформализованные, расставляются интерфейсы принятия решений на всех уровнях и т.п.). При построении слабоформализованной модели для управления, параллельно моделируется обеспеченность этой модели (степень адекватности формализации модели для управления). Поскольку интегрально и открыто представлять деятельность организационной системы “не должен никто лучше, чем руководитель”, то в процессе построения технологии со-управления Разработчик технологии становится неформальным руководителем моделируемого предприятия (Заказчика технологии со-управления). Затем эти “полномочия” передаются Заказчику и Актом приема-передачи, и по предметной сути технологии управления. Передаются его конкретному формальному подразделению со-управления, которое и будет работать со слабоформализованной моделью для управления, модернизировать ее при необходимости, моделировать обеспеченность деятельности и модели для управления, учитывать корректировки модели экономики в целом.

Несложно видеть, что при создании и использовании технологии со-управления реализуются основные этапы “моделирования” (представления деятельности), которые, как правило, на ментальном уровне проходит собственник организации или его представитель-управленец в процессе управленческой деятельности. И что интеграция слабоформализованных технологий соуправленцем, как руководителем, в рамках использования технологии со-управления, обеспечивает аддитивность этой интеграции (См. п.п. 2,5,6,7). То есть *снимает индивидуальные ограничения моделирования многотехнологичных открытых организационных систем.*

9 “Традиционные” инструменты руководителя

В нашей практике приходилось встречаться с мнениями руководителей крупных компаний, что “Изъятие компьютера со стола руководителя увеличивает эффективность его управленческих решений значительно” [7]. Или что наиболее эффективными и действенными для руководителя ИТ-инструментами при анализе и принятии управленческих решений “В достаточно крупных, даже корпоративных организациях, как ни кажется это архаичным, оказываются простые связанные таблицы в Excel”. А по нашему опыту и мнению основными инструментами руководителя, как прямого функционального пользователя, остаются *лист бумаги, ручка, телефон* (зачастую через секретаря), *калькулятор и отчетно-справочный материал.*

Это не означает, что компьютер и IT-технологии противопоказаны руководителю организационной системы, наоборот, знание любых современных технологий только повышает индивидуальные возможности. Мы имеем в виду то, что если требовать обязательное освоение формализованных технологий решения задач, как правило, специализированных (закрытых технологий), от человека, который ставит задачи и реализует интеграцию многих связанных внутренних и внешних знаний (реализует открытые технологии), то можно перенапрягать его индивидуальные возможности и ущемлять “обязательную для управления открытую интеграцию знаний”. Не говоря даже об ущемлении “обязательного выполнения” других составляющих процесса управления.

10 Коллективное управление, плюсы и минусы

Анализу коллективной модели управления уделялось и уделяется заметное внимание [6], даже с выводами об оптимальном количестве “руководителей на одном директорском месте”. Такой анализ позволяет структурировать особенности коллективного (но неразделенного по предметным областям) управления, выделить положительные и негативные механизмы и тенденции, в зависимости от состояния и развития. Здесь приведены основные, на наш взгляд, “плюсы” и “минусы” коллективного (неразделенного предметно) управления:

- Интеграция знаний. Быстро упорядочивается и проясняется интегральная деятельность системы, знания участников суммируются, накапливаются (когда “одна голова хорошо, а две лучше”) – плюс.
- Новые знания. Имеют место случаи полного взаимопонимания и согласия руководителей, эффективный коллективный анализ и новые решения, особенно на стадии становления-изучения, формализуемости и установления адекватности представлений (тоже “две головы лучше”) – плюс.
- Амбиции, соперничество руководителей и т.п. (“вторая голова” мешает) – минус.
- Наличие неявных ценностей у руководителей (“вторая голова” даже вредит, тормозит) – минус.
- На начальных этапах минусы слабее “включаются”, а плюсы активнее – маленький плюс.
- На развитой относительно стадии минусы все же включаются, а плюсы тают, взаимопонимание в коллективе возможно, но не устойчиво, не управляемо – большой минус.

Можно говорить о такой модели коллективного управления, когда на разных стадиях становления корректируется руководство организационной системы. Или когда нужно прозрачно регламентировать обязанности и ответственности участников-руководителей введением жестких внутренних норм-договоренностей. Либо использовать фискальные технологии (которые, на самом деле, дорогостоящие и не гарантирующие устойчивой долгосрочной стабильности и профессионализма критериев и оценок, и их развития). Вопрос о возможности применения модели коллективного управления зависит от прав собственности, от формы собственности, от конкретной организационной системы и даже от конкретных претендентов. И в каждом конкретном случае, наверное, его нужно решать специально. С другой стороны можно придерживаться достаточно обобщенных выводов о неэффективности коллективного управления [6], и того, что пока кризисных процессов такое управление заметно не уменьшает.

11 О моделях управления

Устоявшиеся укрупненные направления управленческой деятельности практически любой организационной системы - это, в первую очередь, конструирование, регулирование и поддержание стратегии и идеологии (определяющие перспективу и устойчивость системы). Затем это предметное управление текущей деятельностью системы. И в третьих это общее управление материальным и нормативным обеспечением деятельности. Реально, особенно в частных компаниях, все эти направления могут быть сконцентрированы в “руках” первого руководителя. Или они могут быть поделены между директором и председателем совета, например акционе-

ров. Все интеллектуальное, техническое и материальное обеспечение управления (информационное, коммуникационное), являющееся технологиями решения задач, находится в ведении помощников и замов, как прямых функциональных пользователей этих технологий (см. рисунок 3). В случае использования технологии со-управления, к традиционной модели добавляется еще одно направление управленческой деятельности – интеграция знаний (в основном технологий), как уже своеобразное прямое обеспечение управления, или “инструмент для руководителя”, который одновременно “управляет” деятельностью (см. рисунок 4).

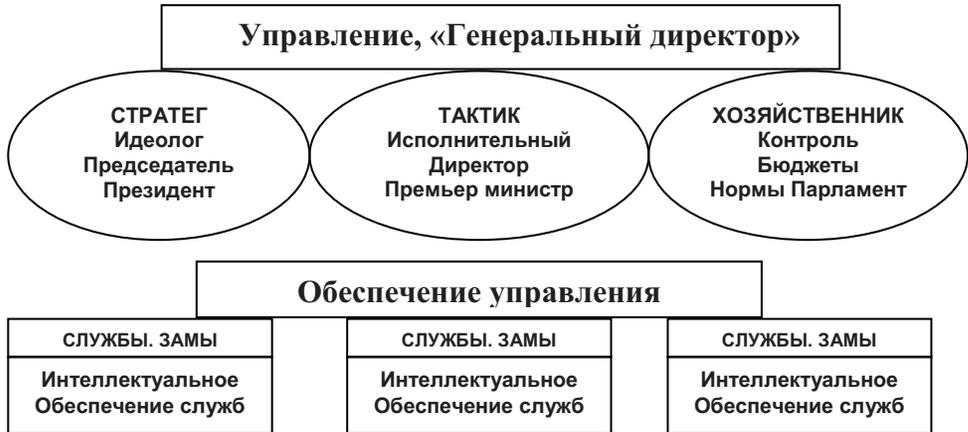


Рисунок 3 – Традиционная модель управления (Государственное управление как «top» реализация, устоявшаяся отработанная модель)



Рисунок 4 – Модель управления с обеспечением управления технологией со-управления

Введение в модель управления технологии со-управления может потенциально способствовать возможности развития модели управления “государство-предприниматель”. И потенциально дать дополнительные возможности подконтрольно и эффективно управлять общественно причастными деятельностями (например, ресурсными, массовыми инфраструктурными и монопольными). И в этом смысле способствовать развитию управляемого свободного рынка.

Что касается такой привлекательной перспективы, как создание искусственного интеллекта (см. п.п. 5), к которой обязательно нужно стремиться, то здесь мы пока не научились “считать процесс мышления”, или, хотя бы, формулировать системную и прозрачную поста-

новку задачи, - что и как считать или, просто, формализовать в мышлении. Поэтому рассматривать такую перспективу как ближайшее решение проблем управления, и даже не очень ближайшее, по нашему мнению, не реально. Хотелось бы надеяться, что Билл Гейтс ошибался, когда говорил в своей презентационной речи при выходе из управления Microsoft, что если и есть перспектива создания искусственного интеллекта, то она необозримо отдаленная.

Заключение

Проанализировано влияние индивидуальных и коллективных возможностей управления на проблемы управляемости организационными системами. Показано, что основной принципиальной причиной неуправляемости в организационных системах может являться существование естественного ограничения возможностей построения открытых моделей многотехнологичных организационных систем (неаддитивность моделирования слабоформализованных технологий). В качестве искусственного увеличения возможностей открытого моделирования организационных систем предложена технология со-управления, - когнитивная технология управления, построенная на основе ментального анализа индивидуального управления.

Создание технологии со-управления на базе разделения единого управления на две выделенные управляющие стороны (деятельность руководителя и со-управление), и на основе открытого интегрального моделирования деятельности, соответствующего представлению руководителя о ней, позволяет *искусственно снять ограничения индивидуальных возможностей открытого моделирования многотехнологичных организационных систем* (снять неаддитивность моделирования слабоформализованных технологий). Это качественно увеличивает возможности индивидуального управления и управляемость многотехнологичными организационными системами. И это может потенциально способствовать развитию *модели управления "государство-предприниматель"* и созданию более *управляемого свободного рынка*.

Проблемы управления в каждой организационной системе индивидуальны. И их решения в каждом конкретном случае уникальны, субъективны, конъюнктурны. Но общее в эффективном решении проблем управления – это комплексные меры с построением модели (архитектуры) управления – на макро и микро уровнях, с построением систем информационного и коммуникационного обеспечения управленческой и исполнительской деятельности. И это, также, построение открытых интегрированных моделей для управления и технологий со-управления, наиболее актуальное в наиболее сложных, многотехнологичных организационных системах.

Список литературы

- [1] Тоффлер, Э. Революционное богатство: как оно будет создано и как оно изменит нашу жизнь. – М., 2008. – 569 с.
- [2] Ambrosius, G. and W. Hibbard, A Social and Economic History of Twentieth-Century Europe (1989).
- [3] Feinstein. Charles H. The European economy between the wars (1997).
- [4] Злобин Н.В. Второй новый миропорядок. Эксмо, 2009. -352.
- [5] Георг Миллер (George Miller), 1956 год. «Волшебное правило» "семь плюс минус два (7 ± 2)"; Восприятие предметов, образов – субъектов и объектов.
- [6] Elinor Ostrom, Governing the Commons: The Evolution of Institutions for Collective Action, 1990 год. Монография 1991 года, Нобелевская Премия по экономике 2009г.
- [7] Пунда Д.И. АКМ-метод анализа слабоформализуемых знаний при моделировании деятельности организационных систем. – Проблемы управления и моделирования в сложных системах: Труды XI Международной конференции. – Самара, Самарский научный центр РАН, 2009, с.520-531.
- [8] Виттих В.А. «Парадигма ограниченной рациональности принятия решений», Препринт Института проблем управления сложными системами, РАН, 28 с., Самара, 2009.
- [9] Всемирная энциклопедия. Философия.- М.: АСТ; Мн.: Харвест, Современный литератор, 2001.
- [10] Колмогоров А. Н., Драгалин А. Г.. Математическая логика. Изд. УРСС. 2003. 240с.

ОНТОЛОГИЧЕСКИЙ ПОДХОД К ПОСТРОЕНИЮ СИСТЕМ СБАЛАНСИРОВАННЫХ ПОКАЗАТЕЛЕЙ

И.Д. Виноградов

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
vid@svptus.ru
тел: +7 (846) 332-39-27, факс: +7 (846) 333-27-70

Ключевые слова: система сбалансированных показателей, онтология ССП, объектная модель ССП

Abstract

This clause presents the ontology approach to creation the Balanced Scorecard system.

Введение

Система сбалансированных показателей (ССП) представляет собой современную концепцию управления предприятием на принципах количественного измерения его хозяйственной деятельности. Построение ССП – это сложная многоплановая задача, охватывающая стратегический, тактический и оперативный уровни управления предприятием. Решение этой задачи затруднено из-за отсутствия методологии по интеграции ключевых показателей эффективности работы предприятия, относящихся к различным уровням и сферам производства с целью обеспечения их согласованности и сбалансированности. Существующие подходы к построению ССП обладают малыми возможностями автоматизации извлечения и структуризации знаний, необходимых для наилучшей реализации информационной системы. Онтологический подход обеспечивает не только автоматизированное построение ССП на основе компьютерного представления обобщенных знаний в форме онтологий ССП, но и имеет целый арсенал методов и средств получения и интеграции этих знаний [1-3]. В данной работе рассматривается постановка задачи построения онтологий ССП с помощью применения анализа формальных понятий из наборов показателей деятельности предприятия (финансовых, экономических, маркетинговых, производственных и т.д.), представленных в форме таблиц «объекты-свойства», а также операций объединения и связывания онтологий.

1 Построение онтологии ССП

Рассмотрим применение анализа формальных понятий (АФП) [4, 5] для автоматизации построения онтологий ССП. Пусть имеются некоторые наборы показателей, отражающие различные аспекты производственной деятельности. Для простоты понимания ограничимся тремя наборами показателей: маркетинговыми, финансовыми и производственными. Каждый показатель в терминологии онтологического подхода представлен в виде характерных для него признаков (атрибутов), т.е. по сути, является описанием класса. Для возможности применения АФП объединим наборы показателей и представим результат объединения в форме формального контекста (см. таблицу 1). Как правило, формальный контекст формируется экспертным путем, но может быть извлечен автоматически из таблиц «объекты-свойства». Первые три признака формального контекста: *стратегический*, *тактический* и *оперативный* указывают на принадлежность показателя к уровню управления предприятием. Остальная группа признаков характеризует отношения между показателями: «главный-подчиненный», «агрегированный-детализированный», «независимый-зависимый». Следует отметить, что один и тот же показатель на разных уровнях управления может выступать как первым ар-

гументом, так и вторым аргументом в перечисленных выше отношениях. Например, он одновременно может обладать признаками агрегированности и детализированности в зависимости от аспекта его рассмотрения. В качестве аспекта рассмотрения показателя выступает объектная модель ССП, созданная на базе онтологии.

Далее из формального контекста автоматически извлекается решетка понятий, представляющая собой онтологию ССП (см. рисунок 1). Онтология включает в себя обобщенные знания в виде абстрактных обобщенных понятий (показателей), необходимые для автоматизации построения объектных моделей ССП.

Таблица 1 – Формальный контекст ССП

Показатели/Атрибуты	Стратегический	Тактический	Оперативный	Агрегированный	Детализированный	Главный	Подчиненный	Зависимый	Независимый
Маркетинговые:									
Рентабельность продаж	+	+		+	+	+	+	+	+
Кол-во сделок			+		+	+	+	+	+
Темп прироста продаж		+		+	+		+	+	+
Финансовые:									
Объем прибыли	+	+	+	+		+	+	+	
Себестоимость продукции		+	+		+		+	+	+
Рентабельность капитала	+			+		+		+	
Производственные:									
Объем продукции	+	+	+	+	+	+	+	+	+
Темп роста производительности труда			+		+		+	+	+
Рентабельность производства	+	+		+		+		+	

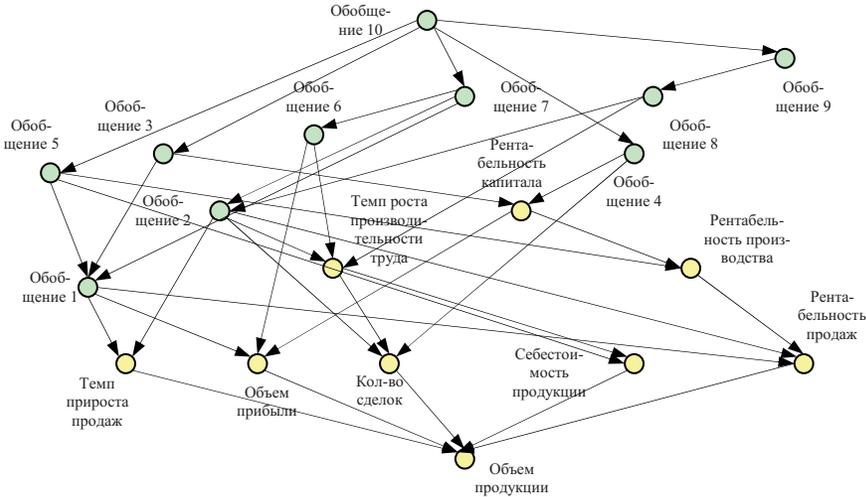


Рисунок 1 – Онтология ССП извлеченная из формального контекста в форме диаграммы Хассе

Таким образом, АФП позволяет извлекать обобщенные показатели для построения ССП, но их анализ и интерпретация по-прежнему остается за человеком. Например, полученное в

примере обобщение 8 интерпретируется как детализированный показатель, обладающий свойствами зависимости от других показателей или независимости от них. В данном примере наиболее общим понятием является показатель, обладающий свойством зависеть от других показателей. Извлеченная решетка понятий может содержать неопределенность и противоречия из-за множественного наследования между классами [2]. Для их устранения применяется операция нормализации онтологии [6]. В ходе ее проведения «разрываются» петли множественного наследования путем введения дополнительных связей наследования между обобщениями и листовыми понятиями. Плата за нормализацию онтологии – избыточность связей между понятиями.

Поиск наиболее предпочтительной онтологии достигается путем многократного построения решеток понятий по различным наборам признаков показателей (количество построений исследователь определяет исходя из сложности предметной области). В результате проведения онтологического анализа по обобщению показателей для реализации ССП получена обобщенная онтология, представленная на рисунке 2.

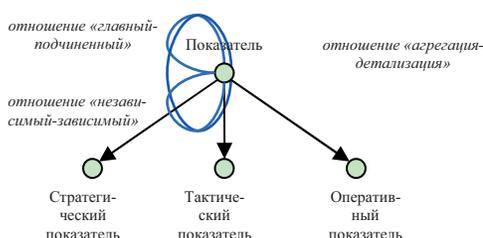


Рисунок 2 – Обобщенная онтология ССП

Далее обобщенная онтология ССП интегрируется с общей онтологией информационно-аналитической системы. Фрагмент такой интеграции показан на рисунке 3. Интеграция онтологий производится с помощью операций объединения и связывания [7-9].

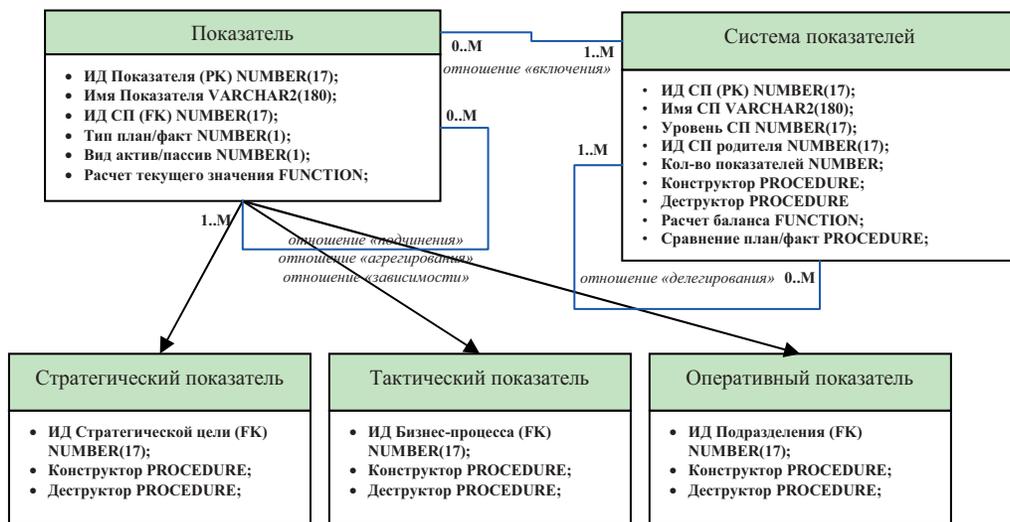


Рисунок 3 – Фрагмент интегрированной онтологии ССП

Класс «Система показателей» выступает в качестве контейнера для хранения показателей. С его помощью можно реализовать различные по-своему наполнению ССП, имеющие иерархическую структуру.

При построении онтологии важно обеспечить согласованность объектных моделей ССП. Объектные модели предназначены для решения различных прикладных задач (выбора, оптимизации, прогнозирования и т.д.) с использованием соответствующих методологий. Для компьютерной реализации методологий строятся соответствующие им методо-ориентированные онтологии [1]. Например, при помощи онтологии метода анализа иерархий Т. Саати (МАИ) строятся компьютерные объектные модели многокритериального выбора из множества альтернатив, а на базе онтологии системной динамики Дж. Форрестера реализованы пакеты прикладных программ имитационного моделирования непрерывных систем, такие как PowerSim, IThink, AnyLogic и др. [11-12]. Согласованность объектных моделей ССП осуществляется при помощи операции связывания онтологий. На рисунке 4 показан фрагмент связывания предметной онтологии ССП с методо-ориентированными онтологиями МАИ и системной динамики.

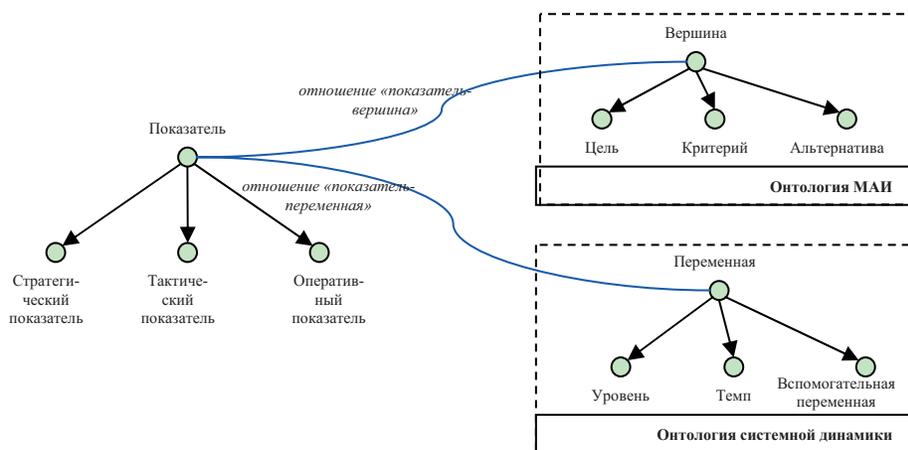


Рисунок 4 – Фрагмент связывания онтологий

2 Построение объектных моделей ССП

Объектные (денотативные) модели ССП состоят из объектов (экземпляров) классов, описанных на концептуальном онтологическом уровне. Например, объектами понятия «Стратегический показатель» являются показатели: валовая прибыль, производство чистой продукции, рентабельность и др. На базе онтологической спецификации с помощью CASE-средств автоматически создается база данных для объектного моделирования. Наполнение объектной модели начинается с конструирования объектов и программной реализации методов их использования. Построение объектной модели ССП ведется на общих принципах разработки ключевых показателей эффективности [13]. В соответствии со стратегическими целями предприятия и спецификой его деятельности строятся иерархические многокритериальные модели для отбора ключевых показателей эффективности для каждого уровня управления:

- kri(c) – стратегические ключевые показатели эффективности предприятия;
- kri(bp) – ключевые показатели эффективности бизнес-процессов;
- kri(d) – ключевые показатели эффективности структурных подразделений;
- kri(p) – ключевые показатели эффективности исполнения должностей.

Пример такой модели показан на рисунке 5.

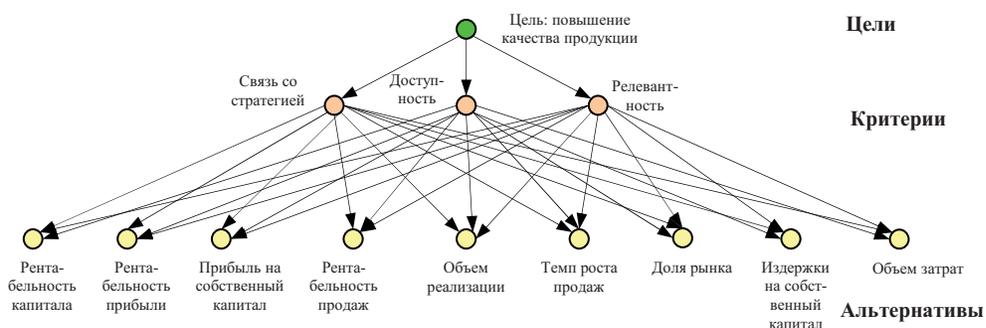


Рисунок 5 – Модель МАИ для оценки показателей эффективности

После отбора ключевых показателей для каждого уровня управления строятся объектные модели, в которых описываются причинно-следственные связи между показателями. Один из фрагментов каузальной сети показан на рисунке 6. Каузальные объектные модели описывают связи между показателями как одного уровня, так между показателями разных уровней управления, образуя единую систему показателей для предприятия.

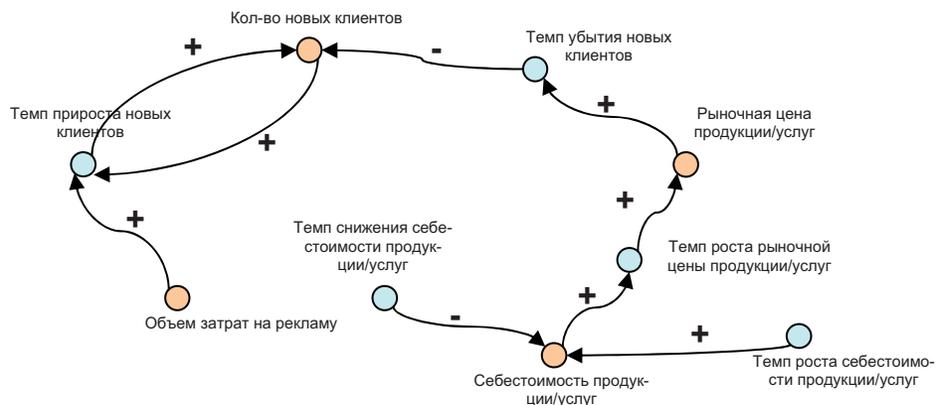


Рисунок 6 – Фрагмент объектной модели каузальной сети ССП

При построении ССП важно учитывать динамику производственных процессов, т.е. необходимость обеспечения актуальности ССП в условиях динамически развивающихся бизнес-процессов. В этом большое значение уделяется имитационным моделям, в частности моделям системной динамики Дж. Форрестера [10]. Модели системной динамики используют переменную времени в форме записей разностных дифференциальных уравнений 1-го порядка. Такие модели имеют графическое «потокосное» представление близкое к нотации построения диаграмм причинно-следственных связей (см. рисунок 7). Сбалансированность показателей эффективности достигается посредством уравнений балансов.

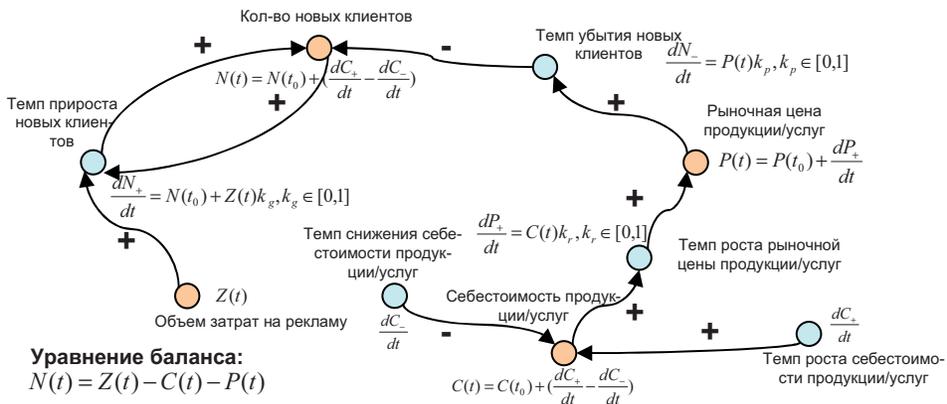


Рисунок 7 – Фрагмент объектной имитационной модели ССП

Заключение

Построение систем сбалансированных показателей на основе онтологического подхода обладает следующими достоинствами:

- автоматизация извлечения знаний из наборов данных;
- унифицированное представление знаний для их компьютерной обработки;
- автоматизация интеграции разнородных знаний путем их композиции и связывания;
- автоматизированное построение объектных моделей ССП и обеспечение их согласованности;
- обеспечение актуальности ССП в условиях динамически меняющихся бизнес-процессов предприятия за счет автоматизации процедур приобретения, накопления и использования знаний и данных;
- доступность реализации ССП специалистами, не обладающими знаниями и навыками программирования.

Данный подход может быть использован при создании информационно-аналитических систем (OLAP-систем) для решения задач автоматизации принятия управленческих решений.

Список литературы

- [1] Смирнов С.В. Онтологический анализ предметных областей моделирования // Известия Самарского научного центра РАН. 2001. Т.3. №1. - С. 62-70.
- [2] Гаврилова Т.А., Муромцев Д.И. Интеллектуальные технологии в менеджменте: инструменты и системы: Учеб. Пособие. 2-е изд. - СПб.: Изд-во «Высшая школа менеджмента»; Издат. Дом С.-Петерб. гос. ун-та, 2008. – 488 с.
- [3] Виттих В.А. Интеграция знаний при исследованиях сложных систем // Известия академии наук. Теория и системы управления. 1998. №5. - С. 132-139.
- [4] Ganter B., Wille R. Formal Concept Analysis: Mathematical Foundations. - Berlin; Heidelberg: Springer-Verlag, 1999.
- [5] Кузнецов С.О., Обьедков С.А. Алгоритмы построения множества всех понятий формального контекста и его диаграммы Хассе // Известия Академии Наук, теория и системы управления. 2001. №1. – С.120-129.
- [6] Виноградов И.Д., Смирнов С.В. Композиция концептуальных схем сложных систем // Проблемы управления и моделирования в сложных системах: Труды международной конференции (15-17 июня, 1999 г., Самара, Россия) – Самара: СамНЦ РАН, 1999. - С. 184-190.

- [7] Виноградов И.Д., Смирнов С.В. Алгоритм объединения концептуальных схем на основе реконструкции их формального контекста // Проблемы управления и моделирования в сложных системах: Труды III международной конференции (4-9 сентября, 2001 г., Самара, Россия) – Самара: СамНЦ РАН, 2001. - С. 213-220.
- [8] Natalya Fridman Noy and mark A. Musen. SMART: Automated Support for Ontology Merging and Alignment // Stanford Medical Informatics Stanford University, CA. – С.1-20.
- [9] Клещев А.С., Артемьева И.Л. Отношения между онтологиями предметных областей. Часть 1. Онтологии, представляющие одну и ту же концептуализацию, упрощение онтологий // Научно-техническая информация, Сер.2, 2002, №1. - С. 4-16.
- [10] Форрестер Дж. Основы кибернетики предприятий (индустриальная динамика). – М.: Прогресс, 1971.
- [11] Карпов Ю. Имитационное моделирование систем. Введение в моделирование с AnyLogic 5. – СПб.: БХВ-Петербург, 2006. – 400 с.
- [12] Сидоренко В.Н. Системно-динамическое моделирование в среде POWERSIM: Справочник по интерфейсу и функциям. - М.: Макс-Пресс, 2001.
- [13] Данилин О. Принципы разработки ключевых показателей эффективности (КПЭ) для промышленных предприятий и практика их применения // Управление компанией. 2003. №2 (21).

СИНТЕЗ И АНАЛИЗ ЭНЕРГОЭФФЕКТИВНЫХ СИСТЕМ УПРАВЛЕНИЯ НА ОСНОВЕ НАБЛЮДАТЕЛЕЙ СОСТОЯНИЯ

А.Н. Дилигенская

Самарский государственный технический университет
443100, Самара, ул. Молодогвардейская, 244, Россия
ADiligenskaya@mail.ru
тел: +7 (846) 337-07-00

Ключевые слова: метод пространства состояний, аналитический регулятор состояния, наблюдатель состояния, энергоэффективность

Abstract

The problem of the synthesis and study of energy-efficient buildings thermal regime control system on the basis of state observer in the presence of non-measurable indoors temperature and uncontrolled disturbances.

Введение

Проблема повышения энергоэффективности является актуальной при решении задач совершенствования и управления производством во всех отраслях экономики. Отсутствие целенаправленной политики в этой сфере привело к тому, что на сегодняшний день удельные затраты энергетических ресурсов в стране превышают аналогичные затраты в передовых экономически развитых странах в 2.5-3.5 раза.

Значительный потенциал энергосбережения сосредоточен в обеспечении жилых и производственных помещений тепловой энергией, и его величина составляет до 25% от общей энергоемкости. Традиционный способ теплоснабжения помещений заключается в регулировании в соответствии с действующими нормативными правилами температуры теплоносителя в помещении. При этом реальная температура воздуха в помещении является неизвестной в условиях действия неконтролируемых возмущений – внешних (наружные воздействия), внутренних (бытовые тепловыделения – люди, электроустановки), параметрических (открытые окна, двери и т.п.). Качество управления по отношению к температуре воздуха в помещении (статическая и динамические ошибки) становится во многих случаях неудовлетворительным. Кроме того, не оптимизируется расход энергии на отопление.

В работе предлагается новый подход к решению таких задач, основанный на синтезе аналитического регулятора и наблюдателя состояния.

1 Описание объекта управления

Рассматривается задача синтеза системы оптимального управления автономной отопительной установкой, предназначенной для поддержания заданной температуры воздуха в комплексе помещений в условиях действия неконтролируемых возмущений разнообразной природы - внешних, внутренних, параметрических. Источником тепла является теплоагрегат (нагреватель), локальная система управления которого обеспечивает регулирование температуры теплоносителя в функциональной зависимости от температуры наружного воздуха. Регулирование температуры внутри помещений обеспечивается за счет управления мощностью теплоисточника. В качестве контролируемой величины используется температура теплоносителя на выходе нагревателя. Выбор такой координаты обусловлен простотой реализации системы. В этом случае величина, подлежащая налагаемым требованиям на поведение, - температура воз-

душной массы помещения – является недоступной для непосредственного контроля и соответствует цели управления со значительной долей отклонений. Предлагается восстанавливать неконтролируемую температуру воздуха с помощью наблюдателя состояния и использовать полученную оценку для оптимизации системы управления теплообеспечением.

2 Математическая модель объекта управления

Модель объекта управления конструируется на основе базового уравнения сохранения тепловой энергии, устанавливающего баланс между выделением энергии в теплоисточнике, с одной стороны, и изменением теплосодержания теплоносителя и теплоотдачей в окружающую воздушную массу помещения, с другой стороны, и уравнения теплового баланса в помещении, описывающего расход тепла от теплоисточника на повышение теплосодержания воздуха и на потери из помещения в окружающую среду [1].

Описанию объекта управления, состоящего из совокупности дифференциальных уравнений для динамики температуры теплоносителя и нагреваемой воздушной массы, соответствует структура, состоящая из двух апериодических звеньев первого порядка, связанных положительной обратной связью [2].

Задача синтеза системы оптимального управления решалась на основе метода пространства состояний, при этом использовалась модель рассматриваемого объекта в универсальной векторно-матричной форме в базисе координат состояния со скалярным входным управляющим воздействием $u(t) = Q_V(t)$, соответствующим объемной мощности тепловыделения в нагревателе, имеющая вид:

$$(1) \quad \begin{aligned} \frac{dx}{dt} &= Ax(t) + Bu(t); \\ y(t) &= Cx(t). \end{aligned}$$

В (1) порядок объекта $n = 2$, $x = [x_1 \ x_2]^T$ – вектор состояния системы; $[...]^T$ – знак транспонирования; y – вектор наблюдений. В качестве компонент вектора состояния объекта рассматриваются величины, характеризующие процесс нагрева: $x_1(t) = \theta_H^0(t)$ соответствует температуре нагревателя, $x_2(t) = \theta_e^0(t)$ – температуре воздушной массы помещения. Температура окружающей среды $\theta_c^0(t)$ обладает значительно большей инерционностью по сравнению с мощностью нагревателя $Q_V(t)$ и рассматривается как возмущающее воздействие. Матрица состояния объекта A , матрица управления B и матрица наблюдений C имеют следующий вид:

$$(2) \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}; \quad B = [b_1 \ 0]^T; \quad C = [1 \ 0].$$

Коэффициенты матриц считаются известными, зависящими от технологических, тепловых и геометрических характеристик нагревательного устройства и помещения [2].

3 Постановка задачи оптимального управления

Далее рассматривается задача в базисе относительных переменных $\xi(t) = [\xi_1(t) \ \xi_2(t)]$, соответствующих отклонениям температур теплоносителя нагревателя и воздуха относительно требуемых значений фазовых координат $x_{\text{зад}} = [x_1^{\text{уст}} \ x_2^{\text{уст}}]^T$, определяемых по соотношениям $\xi_1(t) = x_1(t) - x_1^{\text{уст}}$; $\xi_2(t) = x_2(t) - x_2^{\text{уст}}$, где $x_1^{\text{уст}}$ и $x_2^{\text{уст}} = \theta_e^{\text{зад}}$ – значения координат со-

стояния в установившемся режиме поддержания заданной температуры воздуха помещения $\theta_6^{\text{зад}}$. В качестве управляющего воздействия рассматривается величина относительной мощности теплоисточника $\varphi(t) = u(t) - u_{\text{уст}}(t)$, где $u_{\text{уст}}$ - стационарная мощность нагревателя, обеспечивающая поддержание температуры помещения на уровне $\theta_6^{\text{зад}}$. Значения $x_1^{\text{уст}}$ и $u_{\text{уст}}$ находятся из решения стационарной задачи, отвечающей (1) при $x_2 = \theta_6^{\text{зад}}$.

Сформулируем на основе детерминированной модели объекта управления (1), (2) задачу оптимального управления в стохастической постановке при учете случайных неконтролируемых возмущений. Система теплообеспечения рассматривается в виде линейного стационарного объекта, описываемого вероятностными уравнениями

$$(3) \quad \begin{aligned} \frac{d\xi}{dt} &= A\xi(t) + B\varphi(t) + v(t); \\ z(t) &= C\xi(t) + \eta(t), \end{aligned}$$

где $v(t)$ - помеха, действующая на входе объекта и $\eta(t)$ - погрешности измерения, полагаемые некоррелированными белыми гауссовскими случайными процессами с нулевыми средними и статистическими характеристиками $M\{v(t)v(\tau)^T\} = V\delta(t - \tau)$ и $M\{\eta(t)\eta(\tau)^T\} = H\delta(t - \tau)$, где $M\{\dots\}$ - знак операции математического ожидания, $\delta(t)$ - дельта-функция.

Начальное состояние объекта определяется величинами

$$(4) \quad \begin{aligned} \xi_1(0) &= \theta_n^{\text{нач}} - x_1^{\text{уст}}; \\ \xi_2(0) &= \theta_6^{\text{нач}} - x_2^{\text{уст}}, \end{aligned}$$

где $\theta_n^{\text{нач}}$ и $\theta_6^{\text{нач}}$ - температуры нагревателя и воздуха помещения в начальный момент времени, полагающиеся равными $\theta_n^{\text{нач}} = \theta_6^{\text{нач}} = \theta^{\text{нач}}$.

Объект (2), (3), (4) полагается управляемым и наблюдаемым, т.к. ранги его матриц управляемости и наблюдаемости равны размерности вектора состояния $n = 2$ [2].

Рассматривается задача стабилизации состояния объекта относительно положения равновесия $\xi = [0 \ 0]^T$ из заданного начального состояния $\xi(0) = [\xi_1(0) \ \xi_2(0)]^T$.

Конкретные требования к переходному процессу задаются в виде критерия качества, определяющего алгоритм управления тепловым режимом. В соответствии с классическими работами Летова А.М., Калмана Р., Красовского А.А [3], [4], [5] постановка задачи аналитического конструирования на некотором интервале $[0, t_k]$ оптимального регулятора формулируется при минимизации функционала, отвечающего взвешенной сумме интегральных квадратичных значений отклонений вектора состояния $\xi(t)$ от требуемых конечных значений, и управляющего воздействия φ

$$(5) \quad J(\xi, \varphi) = \frac{1}{2} \int_0^{t_k} (\xi^T(t)Q\xi(t) + \varphi^T(t)R\varphi(t))dt.$$

В (5) Q и R - симметричные вещественные положительно определенные матрицы соответствующих размерностей. Матрица Q задает меру (веса значимости) отклонений фазовых координат от заданных значений и выбирается на основе технологических и экономических соображений. Скалярное значение R характеризует затраты на управляющее воздействие и с физической точки зрения определяет расход энергии на обогрев помещений.

4 Синтез системы управления

Для стационарных матриц системы A, B, C решение задачи оптимизации реализуется с помощью обратных связей с пропорциональным регулятором отклонений фазовых переменных в виде [3]

$$(6) \quad \varphi = -K\xi,$$

где K – матрица коэффициентов обратных связей.

В этом случае для синтеза системы управления с регулятором состояния необходимо измерение всех компонент вектора состояния объекта ξ . В условиях действия неконтролируемых возмущений $\nu(t)$ и $\eta(t)$ одна из компонент вектора состояния объекта ξ_1 , необходимая для синтеза управления, измеряется с существенными случайными ошибками, а подлежащая налагаемым требованиям величина ξ_2 не измеряется.

При недоступной непосредственному контролю компоненте ξ_2 , ее возможно восстановить с помощью наблюдателя состояния, формирующего на выходе некоторую оценку $\hat{\xi}$ частично ненаблюдаемого и зашумленного вектора состояния. После этого для синтеза управляющего воздействия $\varphi(t)$ (6) в цепи обратной связи неконтролируемый вектор ξ заменяется полученной оценкой $\hat{\xi}$, и управляющее воздействие формируется в виде $\varphi = -K\hat{\xi}$.

Согласно принципам разделимости и дуальности [6], характеристические полиномы замкнутой системы управления и наблюдателя независимы друг от друга, и их собственные числа, в дальнейшем входящие без изменений в замкнутую систему управления, выбираются отдельно друг от друга. Таким образом, задача синтеза системы управления сводится к синтезу наблюдателя состояния и синтезу регулятора состояния, решаемым независимо друг от друга.

Для восстановления значения ненаблюдаемой величины ξ_2 применяется наблюдатель состояния с минимальной дисперсией – фильтр Калмана, формирующий на выходе оптимальную в смысле минимума среднеквадратичного отклонения

$$J_1 = \int_0^{t_k} (\xi(t) - \hat{\xi}(t))^2 dt = \|\xi(t) - \hat{\xi}(t)\|^2 \rightarrow \min \text{ оценку } \hat{\xi} \text{ по измеряемым сигналам мощности теп-$$

ловыделения и температуре теплоносителя нагревателя. При конструировании фильтра Калмана в качестве идентификатора состояния принимается математическая модель системы, и для уменьшения дисперсии оценок в алгоритм идентификатора вводится корректирующая обратная связь по выходу системы $z(t)$.

Оптимальный идентификатор осуществляет процедуру оценивания по алгоритму:

$$\frac{d\hat{\xi}}{dt} = A\hat{\xi}(t) + B\varphi(t) + L \left[z(t) - C\hat{\xi}(t) \right];$$

$$\hat{\xi}(0) = \xi_0,$$

где $L = GC^T H^{-1}$ – матрица коэффициентов усиления фильтра, формирующая динамические свойства идентификатора и зависящая от ковариационной матрицы ошибок оценивания G , определяемой в случае стационарности процессов единственным неотрицательно определенным решением алгебраического матричного уравнения Риккати $AG + GA^T - GC^T H^{-1} CG + V = 0$ относительно матриц V и H интенсивностей случайных воздействий $\nu(t)$ и $\eta(t)$ [7].

При синтезе системы управления, в случае ее линейности и стационарности, минимизация квадратичного функционала (5) может быть реализована с помощью линейного алгоритма

управления, в замкнутой форме имеющего вид (6), где матрица коэффициентов обратных связей $K = R^{-1}B^T P$ определяется единственным неотрицательно определенным решением P алгебраического матричного уравнения Риккати относительно весовых матриц функционала качества $A^T P + PA - PBR^{-1}B^T P = -Q$.

Объединяя решение задачи синтеза наблюдателя состояния и регулятора состояния, структурную схему системы оптимального управления можно представить в следующем виде:

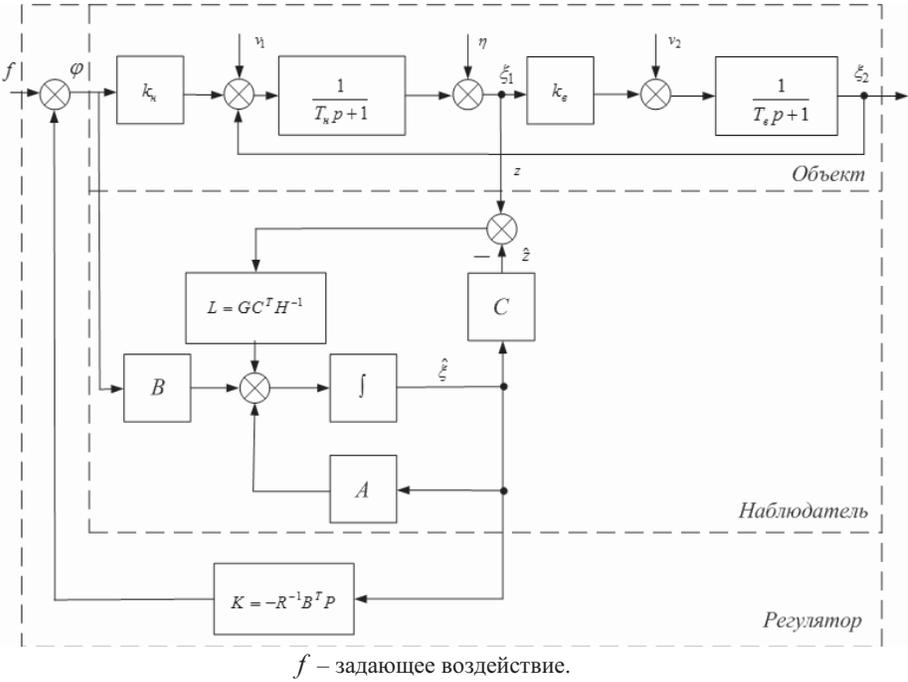


Рисунок 1 – Структурная схема системы оптимального управления

Неконтролируемые возмущающие воздействия на объект и погрешности измерения включены в аддитивные помехи входных и выходных процессов и промоделированы случайными гауссовскими сигналами.

Вычислительные эксперименты, выполненные при реализации контуров обратных связей по управлению с использованием наблюдателя, построенного по принципу фильтра Калмана при учете аддитивных случайных воздействий типа белого шума, показывают работоспособность рассмотренных алгоритмов наблюдателя и регулятора.

На основе представленной структурной схемы оптимального управления (рисунок 1) был проведен сравнительный анализ качества регулирования процессов обогрева помещений в условиях действия различного рода помех для случаев использования традиционного П- и аналитического регуляторов. Конкретные значения параметра П-регулятора и весовых коэффициентов функционала качества выбраны таким образом, что при отсутствии случайных возмущений траектории переходных процессов температур, управляющих воздействий и интегральных критериев практически совпадают между собой.

При учете случайных воздействий в случае использования классического П- регулятора появляется некомпенсируемая погрешность по отношению к применению аналитического регулятора с наблюдателем состояния (рисунки 2, 3).

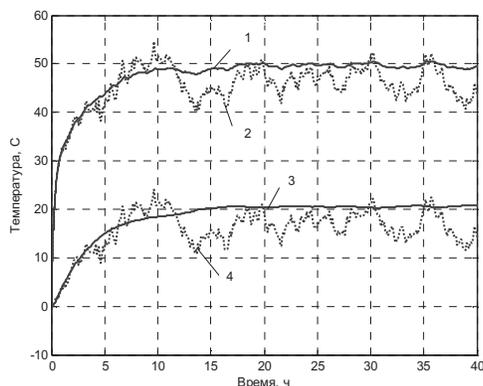


Рисунок 2 – Переходные процессы при наличии помех при аналитическом (кривые 1, 3 для температуры нагревателя и воздуха) и классическом П-регулировании (кривые 2, 4 для тех же температур)

Из результатов расчета (рисунки 2, 3) видно, что применение аналитического регулирования с наблюдателем состояния в установившемся режиме обеспечивает практически точное поддержание требуемой температуры воздуха в помещении, в то время, как использование П-регулятора по температуре приводит к значительным колебаниям температуры воздуха. При применении П-регулирования виден перерасход энергии и ухудшение качества управления по интегральному квадратичному функционалу оптимизации.

Заключение

Решение задачи регулирования теплоснабжением помещений на основе аналитического конструирования с использованием наблюдателя состояния, восстанавливающего подлежащую регулированию неизмеряемую величину – температуру воздуха - с минимальной дисперсией, по сравнению с традиционным подходом имеет существенные преимущества по показателям качества интегрального квадратичного критерия: точности нагрева и затратам энергии.

Список литературы

- [1] Табунщиков Ю.А., Бродач М.М. Математическое моделирование и оптимизация тепловой эффективности зданий. - М.: АВОК-ПРЕСС, 2002, 194с.
- (1) Дилигенская А.Н., Щетинин В.Г. Синтез структуры системы обогрева помещений в условиях неполноты измерений // Вестник СамГТУ. Серия «Технические науки». 2009. №3 (25). С. 206-214.
- [2] Летов А.М. Аналитическое конструирование регуляторов, I-IV // Автоматика и телемеханика, 1960. № 4. С. 436–441; 1960. № 5. С. 561–568; 1960. № 6. С. 661–665; 1961. № 4. С. 425–435.
- [3] Калман Р., Фалб П., Арбиб М. Очерки по математической теории систем. М.: Мир, 1971.
- [4] Справочник по теории автоматического управления / Под ред. А.А. Красовского. – М.: Наука. Гл. ред. физ.-мат. лит., 1987. – 712 с.
- [5] Андреев Ю.Н. Управление конечномерными линейными объектами. - М.: Наука, 1976, 424 с.
- [6] Эйкхофф П. Основы идентификации систем управления. Пер. с англ. - М.: Мир, 1975. - 683 с.

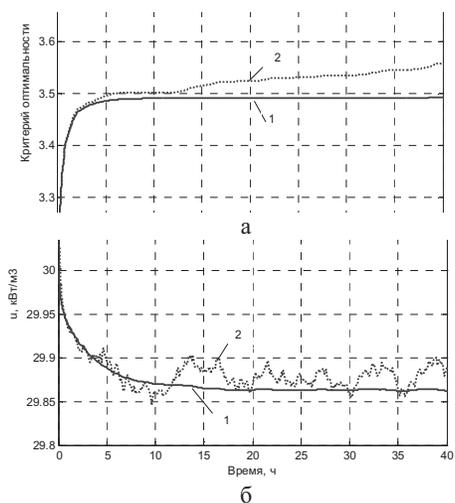


Рисунок 3 – Кривые функционалов оптимизации (а) и управляющих воздействий (б) при использовании аналитического (1) и П-регулятора (2) при учете случайных возмущений

ПРИМЕНЕНИЕ МЕТОДОЛОГИИ СИСТЕМНОГО АНАЛИЗА ПРИ РАЗРАБОТКЕ КОГНИТИВНЫХ МОДЕЛЕЙ КРЕДИТНОГО РИСКА

М.А. Чефранова

Дополнительный офис Ростовского филиала ОАО «Промсвязьбанк»
347900, Ростовская область, Таганрог, ул.Октябрьская, 9, Россия
тел/факс: +7 (8634) 38-38-32

Ключевые слова: системный анализ, факторные модели, когнитивные модели, случайные величины, статистические методы

Abstract

The article considers the application of factor analysis on the cognitive maps for model building process of lending. Article includes features of systems analysis and adaptation of its methods in the semi-systems. Also in the article proposed use of existing methods and descriptions in the database of the cognitive model.

Введение

При моделировании кредитного процесса в коммерческом банке для прогнозирования возможности реализации кредитного риска используется множество методик. В основном они основаны на анализе существующего финансового положения заемщика и существующих прочих факторов, влияющих на размер риска. Однако в последнее время все острее становится проблема прогнозирования возможности реализации кредитного риска на основе статистических данных с использованием методологии системного анализа.

В статье рассмотрены вопросы построения факторных моделей, в том числе через инструментарий когнитивного моделирования, рассматривая процесс банковского кредитования как слабоструктурированную систему.

Система – это сущность, которая в результате взаимодействия ее частей может поддерживать свое существование и функционировать как единое целое.

Такое понимание системы акцентирует внимание на трех моментах:

- на целостности системы, которая самостоятельно поддерживает свое существование и функционирование;
- на источнике или первопричине этой целостности как результате взаимодействия ее частей, косвенно предполагающем интерпретацию составного характера системы, отрицающего редукционизм, то есть представление о целом как о сумме составляющих его частей;
- на связях между частями, обеспечивающих устойчивость существования и функционирования системы.

Для поддержания целостности системы в условиях изменяющейся внешней среды и внутренних трансформаций требуется особая организация системы, обеспечивающая ее устойчивость. Устойчивость проявляется через способность системы парировать действие возмущений, а значит в конечном итоге через сохранение целостности.

В экономических системах для обозначения устойчивого состояния используется понятие «равновесие». Это ограниченная характеристика устойчивости, полученная посредством наблюдения за критическими параметрами и удержания их значений в допустимых пределах.

Наличие обратных связей и использование их для саморегулирования являются еще одним характерным свойством системы, отличающим ее от простого нагромождения взаимосвязанных элементов.

Наличие вышеперечисленных характеристик необходимо и достаточно для того, чтобы причислить исследуемое образование к классу систем и применить к нему методологический арсенал системного анализа.

При моделировании экономических систем необходимо помнить, что в них всегда присутствует задержка, и более того, следствие (результат) может проявиться совсем не в том месте, где его ожидали. Таким образом, имея дело с экономическими системами, нужно быть готовым к тому, что последствия могут отстоять от вызвавшей их причины во времени и пространстве.

Наибольшее практическое применение среди динамических моделей нашли структурные динамические модели сетевого типа.

Таким образом, условие управляемого целенаправленного развития системы выглядит так:

$$U(t) \in U(t) \cap \bar{U}(t), t = \bar{1}, \bar{T}$$

Таким образом, построение динамической структурной модели системы заключается в формализованном описании траектории ее развития путем задания промежуточных состояний системы и управляющих воздействий, последовательно переводящих систему из начального состояния в конечное, соответствующее цели ее развития.

Существуют выходные и входные параметры объекта. При рассмотрении процесса кредитования:

- $\bar{y}(y_1, \dots, y_j)$ – выходные параметры, характеризующие экономические результаты действия объекта: возврат либо невозврат основного долга, процентов за пользование кредитом; доходность операции кредитования с учетом всех издержек; процент проблемных кредитов в кредитном портфеле банка и т.п.

На эти параметры наложены ограничения:

$$y_{i \min} \leq y_i \leq y_{i \max}$$

- $\bar{x}(x_1, \dots, x_n)$ – группа управляющих и возмущающих воздействий. Это всевозможные параметры, от которых зависит протекание процесса (например, обеспеченность кредита, деловая репутация заемщика, динамика коэффициентов ликвидности заемщика и т.д.). На параметры этой группы также наложены ограничения:

$$x_{i \min} \leq x_i \leq x_{i \max}$$

В реальных условиях все или часть этих параметров – случайные величины. Управляющие параметры также характеризуются случайными величинами – случайные действия персонала и т.п. Поэтому и выходные параметры \bar{y} также величины случайные.

Часть переменных групп \bar{x}, \bar{y} может быть контролируема, часть – не поддаваться контролю по тем или иным причинам.

Для того чтобы найти оптимальный (с заданных позиций) закон управления, необходимо с той или иной форме описать зависимость между всеми переменными \bar{x}, \bar{y} . Далеко не всегда существуют детерминированные связи между ними, поскольку большинство процессов вероятно по своей природе и описать их можно только статистически.

Метод «оптимума номинала» позволяет на основании статистических данных о действующем объекте составить алгоритм наилучшего управления им с точки зрения достижения максимальной экономической эффективности

В большинстве реальных процессов, и в процессе кредитования в том числе, имеет место несимметричное распределение отклонений от номинала и несимметричное распределение результата процесса.

Метод оптимума номинала позволяет для любого закона распределения отклонений и заданного распределения результата реализации процесса определить такое смещение номинала от середины поля допуска, которое дает наибольшее значение математического ожидания результата – наибольшую эффективность действия изучаемого процесса.

Функция эффективности определяется выражением

$$\varphi\left(\bar{\mu}_1, \dots, \bar{\mu}_k\right) = \iint_R b\left(\bar{x}\right) \times f\left(\bar{x}, \bar{\mu}_1, \dots, \bar{\mu}_k\right) d\bar{x} - B\left(\bar{\mu}_1, \dots, \bar{\mu}_k\right)$$

где $\bar{\mu}_1, \dots, \bar{\mu}_k$ - значения 1-го, ..., k-го моментов распределения входных параметров, $k = 1, 2, \dots$,

$f\left(\bar{x}, \bar{\mu}_1, \dots, \bar{\mu}_k\right)$ - плотность вероятности значений входных параметров \bar{x} ,

R – область допустимых входных параметров,

$b\left(\bar{x}\right)$ - результат реализации процесса, соответствующий входному параметру \bar{x} .

Величина $B\left(\bar{\mu}_1, \dots, \bar{\mu}_k\right)$ характеризует стоимость установления и поддержания рассматриваемого процесса с зафиксированной плотностью вероятности значений входных параметров.

Задача оптимизации заключается в установлении таких значений моментов распределения $\bar{\mu}_{10}, \dots, \bar{\mu}_{k0}$, которые обеспечивали бы максимум функции $\varphi\left(\bar{\mu}_1, \dots, \bar{\mu}_k\right)$.

При зависимости процесса от n параметров, эффективность его можно оценить функцией

$$\varphi = \iint_R \dots \int b\left(\bar{x}\right) f\left(\bar{x}\right) d\bar{x}$$

R – объем области n -мерного пространства управляющих параметров,

$b\left(\bar{x}\right)$ - скалярная функция цены.

В рамках данного исследования следует рассматривать вероятностные задачи принятия решений, задачи с риском, задачи в условиях неопределенности первого рода, «игры с природой», которые порождаются слабоструктурированным и проблемами сложных систем.

Остановимся на трех основных моментах, имеющих значение для объединения задач принятия решений и когнитивного анализа.

Во-первых, для построения модели задач принятия решений, в том числе, обобщенной функции эффективности оптимума номинала, необходимо провести исследования для определения всех ее составляющих и параметров: $A, U, X, Y, C(Y), M_h, f(Y, M_h), M_h = \eta(X)$.

Во-вторых, необходимо сформулировать задачи принятия решений в общем виде для каждого этапа когнитивного моделирования.

В-третьих, необходимо сформулировать задачу для выбора наилучшего сценария развития ситуаций на когнитивной карте изучаемого объекта и расчета управляющих воздействий, реализующих этот сценарий.

Технология объединения моделей вероятностных задач принятия решений и когнитивных моделей сводится к реализации следующих основных этапов:

- 1) Построение когнитивных карт.
- 2) Моделирование возможных импульсных процессов в вершинах V построенной карты при внесении возмущающих воздействиях и вариациях $f(x_i, x_j, e_{ij})$. Для чего возможным является использование следующей известной формулой импульсного процесса на графе

$$x_{v_i}(n+1) = x_{v_i}(n) + \sum_{v_j: e_{ij} \in E}^{k-1} f(x_i, x_j, e_{ij})P_j(n) + Q_i(n+1).$$

- 3) Выделение интервалов значений параметров в вершинах когнитивной карты, которые соответствуют разным интервалам значений полезности C_i . Данные интервалы $[y_{in}, y_{iv}]$ рисков, угроз или допустимых значений показателей-индикаторов безопасности и устойчивого развития системы. Необходимо иметь данные о соответствующих этим интервалам полезностям C_i .

Использование когнитивных моделей для представления сложных систем и возможного их поведения путем импульсного моделирования на когнитивных моделях позволяет получить достаточное количество реализаций импульсных процессов. Наличие таких реализаций позволяет поставить и решить задачу оптимума номинала по выбору «лучшего» импульсного процесса, который в дальнейшем может быть принят в качестве желаемой стратегии развития исследуемого объекта.

Для случая k факторов многофакторная модель может быть записана в следующем виде:

$$r_{it} = a_i + b_{i1}F_{1t} + b_{i2}F_{2t} + \dots + b_{ik}F_{kt} + e_{it},$$

где каждый кредитный риск в портфеле характеризуется k чувствительностями, по одной на каждый из k факторов.

Для оценки факторных моделей используется много методов, которые могут быть отнесены к трем группам:

- 1) Методы временных рядов.
- 2) Методы пространственной выборки.
- 3) Методы факторного анализа.

В рамках факторно-аналитического метода построения факторной модели неизвестны ни значения факторов, ни чувствительности рисковости кредитов к этим факторам. Для определения числа факторов и чувствительностей риска по кредиту в прошлом просто применяется статистический метод, называемый факторным анализом. При использовании этого метода объем невозвратов (то есть реализации кредитного риска в полном объеме) некоторой выборки рассматриваются за большое число временных рядов в целях установления одного или нескольких статистически значимых факторов, которые могли бы привести к ковариации рисков, наблюдаемых в этой выборке. По сути, в этом подходе данные по невозвратам сами указывают на структуру факторной модели. Однако, факторная модель, которая хороша для одного временного периода, не всегда будет хорошей моделью для последующего периода. Риски, связанные с разными факторами, а также чувствительности этих рисков к факторам могут меняться с течением времени. Методы статистических оценок должны быть дополнены здравым смыслом, познанием текущей ситуации, для того, чтобы построенная модель учитывала динамическую природу рыночной среды.

Во многих случаях, когда мы сталкиваемся с нечеткими, размытыми параметрами системы и взаимосвязями между ними, целесообразно применять относительно молодой и перспективный метод построения модели системы – когнитивное моделирование.

Заключение

При исследовании кредитного риска как случайного процесса наиболее целесообразным автору представляется совмещение факторно-аналитических, статистических методов на базе когнитивного моделирования. Это позволяет совместить как количественные, числовые характеристики процесса с характеристиками качественными, описательными и уменьшить ошибку при прогнозировании реализации кредитного риска.

Список литературы

- [1] Бокс Дж., Джекинс Г. Анализ временных рядов. Прогноз и управление. Выпуск 1. – М.: Мир, 1974. – 274 с.
- [2] Горелова Г.В., Верба В.А., Захарова Е.Н. Принятие решений на когнитивных моделях сложных систем //Интеллектуальные и многопроцессорные системы-2005. Материалы Международной научно-технической конференции. – Таганрог: Изд-во ТРТУ, 2002. Т.2.
- [3] Корноушенко Е.К. Управление процессами в слабоформализованных средах при стабилизации графовых моделей среды/ Корноушенко Е.К., Максимов В.И. – М: Труды ИПУ, вып.2 – 1998. – 157 с.
- [4] Свечарник Д.В. Задача об оптимуме номинала при вероятностных расчетах /В тр. Института машиноведения АН СССР, семинар по точности в машиностроении и приборостроении, Вып. 10. – М.: Изд-во АН СССР, 1957.

МОДЕЛИРОВАНИЕ НЕРАВНОВЕСНЫХ СОСТОЯНИЙ РЫНКА ТРУДА И ЕГО ВЛИЯНИЯ НА ДИНАМИКУ КРУГООБОРОТА ФИНАНСОВЫХ ПОТОКОВ

Б.Г. Ильясов, И.В. Дегтярева, Е.А. Макарова, А.Н. Павлова

Уфимский государственный авиационный технический университет
450000, Уфа, ул. К.Маркса, 12, Россия
ea-makarova@mail.ru, pavlova.ugatu@gmail.com
тел: +7 (347) 273-78-35

***Ключевые слова:** рынок труда, спрос, предложение, ставка заработной платы, макроэкономический кругооборот, рынок благ, потенциальный выпуск, динамическая модель*

Abstract

The cognitive model of the macroeconomic system functioning regarding goods and labor markets is presented. This article shows the characteristics of the dynamics of interaction of macroeconomic goods and labor market. The dynamic model of labor market functioning is included into the dynamic model of macroeconomic system functioning. The results of experimental research of the controlled and uncontrolled scenarios considering behaviour of the macroeconomic system are discussed.

Введение

В условиях обострения проблем занятости и безработицы особую актуальность приобретают исследования в области системного моделирования неравновесных процессов функционирования макроэкономического рынка труда [1,2]. Цель проводимых исследований заключается в моделировании неравновесных процессов функционирования рынка труда и влияния их на динамику формирования доходов и расходов в системе макроэкономического кругооборота финансовых потоков.

В статье рассматривается решение задач разработки когнитивной и динамических моделей функционирования макроэкономической системы (МЭС) с учетом неравновесных условий взаимодействия рынков благ и труда, а также в проведении экспериментальных исследований динамики взаимосвязанного изменения уровней цен и заработной платы в системе макроэкономического кругооборота финансовых потоков.

Решение перечисленных задач ведется на основе разработанных ранее концепции системного моделирования и динамических моделей воспроизводственного процесса МЭС, формирующего макроэкономический кругооборот финансовых потоков.

1 Когнитивная модель функционирования МЭС с учетом рынков благ и труда

Разработанная ранее когнитивная модель макроэкономического кругооборота МЭС без учета макроэкономических рынков включает в себя четыре концепта – макроэкономических агента (МЭА), в роли которых выступают: реальный сектор (фирмы), сектор домохозяйств (население), секторы финансовых и государственных учреждений. Связями являются финансовые потоки (сплошные линии).

Взаимодействие МЭА во времени сопровождается формированием разнотемповых и разнонаправленных потоков финансовых ресурсов, образующих доходы и расходы МЭА. Помимо потоковых связей в модели выделены информационные связи по запасам (штриховые линии). Каждый МЭА имеет запасы финансовых ресурсов (заштрихованная область

концепта); деятельность МЭА является управляемой на основе информации о накопленных запасах [3].

С концептуальных позиций включение в модель новых элементов – рынков, во-первых, не изменяет сущности построения потоковых взаимосвязей между МЭА, заключающейся в обеспечении кругооборота финансовых ресурсов; во-вторых, не затрагивает локальных механизмов информационной корректировки темпов формирования расходов финансовых ресурсов каждого МЭА на основе информации о запасах; и, в-третьих, дополняет существующие информационные механизмы корректировки потоков по запасам качественно новыми информационными механизмами – механизмами регулирования финансовых потоков, которые основаны на анализе состояния всей МЭС в целом в виде информации о совокупном спросе $\dot{A}d$ и совокупном предложении $\dot{A}s$, а также на анализе состояния рынка труда [4].

Концептуальные положения системного моделирования и исследования рынка труда состоят в следующем.

1) Моделирование и анализ закономерностей функционирования рынка труда ведется на основе кейнсианской концепции занятости с учетом современных теорий рынка труда [5].

2) Процессы функционирования рынка труда рассматриваются в динамике; взаимодействие спроса, предложения и цены труда (ставки заработной платы) рассматриваются во времени, при этом исследуются неравновесные состояния рынка труда. Разрабатываемые когнитивная и динамическая модели функционирования рынка труда включаются в соответствующие модели функционирования МЭС с учетом рынка благ [4].

3) Рынок труда – макроэкономический рынок, построенный путем агрегирования микроэкономических рынков труда с целью выделения наиболее важных закономерностей взаимодействия спроса на труд Nd , предложения труда Ns и формирования ставки заработной платы w .

4) Спрос на труд Nd – это количество трудовых ресурсов (работников), которые могут быть наняты для производства валового продукта при каждой возможной цене труда (ставки заработной платы) при прочих равных условиях [6]. Спрос Nd представлен функцией от реальной ставки заработной платы w , неценовых детерминант ΔNd_{mw} и времени: $Nd = f(w, \Delta Nd_{mw}, t)$. Неценовыми детерминантами спроса на труд являются изменения совокупного спроса, производительности труда, цен других ресурсов. Спрос на труд Nd по своей природе является запасом. Отметим, что спрос на труд определяется стоимостью предельного продукта труда и согласно кейнсианской концепции спрос является производным от совокупного спроса на товары и услуги [6].

На предлагаемой когнитивной модели функционирования МЭС с учетом рынков благ и труда спрос на труд Nd представлен в виде информационной связи (рисунок 1).

5) Совокупное предложение Ns – это количество трудовых ресурсов (работников), которые желают участвовать в производстве валового продукта при каждой возможной цене труда при прочих равных условиях. Предложение труда Ns является функцией от реальной ставки заработной платы w , множества неценовых детерминант ΔNs_{mw} и времени: $Ns = f(w, \Delta Ns_{mw}, t)$. К неценовым детерминантам предложения труда относятся уровень цен, денежные иллюзии, связанные с оценкой влияния роста уровня цен на реальную ставку заработной платы [5].

Поскольку согласно принятым ранее допущениям моделируется движение только финансовых потоков, а движение потоков благ, равных по значению и противоположных по направлению движения, не рассматривается, то на когнитивной модели предложение трудовых ресурсов Ns , так же, как и совокупное предложение $\dot{A}s$ (материальный поток) целесообразно представить в виде специальной информационной связи (штрихпунктирная линия). Необходимо отметить только, что предложение труда, как и спрос на труд, является по

природе своей запасом – запасом трудновосполнимых ресурсов в виде человеческого капитала.

б) Макроэкономические рынки могут находиться как в равновесных, так и в неравновесных состояниях. Равновесное состояние на рынке имеет место, когда спрос и предложение равны. Эти состояния являются идеализированными, их достижение в реальности не представляется возможным, но понятие равновесного состояния используется при исследовании тенденций движения на множестве неравновесных состояний. Такие макроэкономические рынки, как рынок благ и рынок денег, обладают механизмом самовыравнивания спроса и предложения. Например, при возникновении неравновесных состояний рынок благ, обладая механизмом самовыравнивания темпов материального потока благ (совокупного предложения) и финансового потока (совокупного спроса) за счет регулирующего воздействия уровня цен, стремится к восстановлению равновесного состояния.

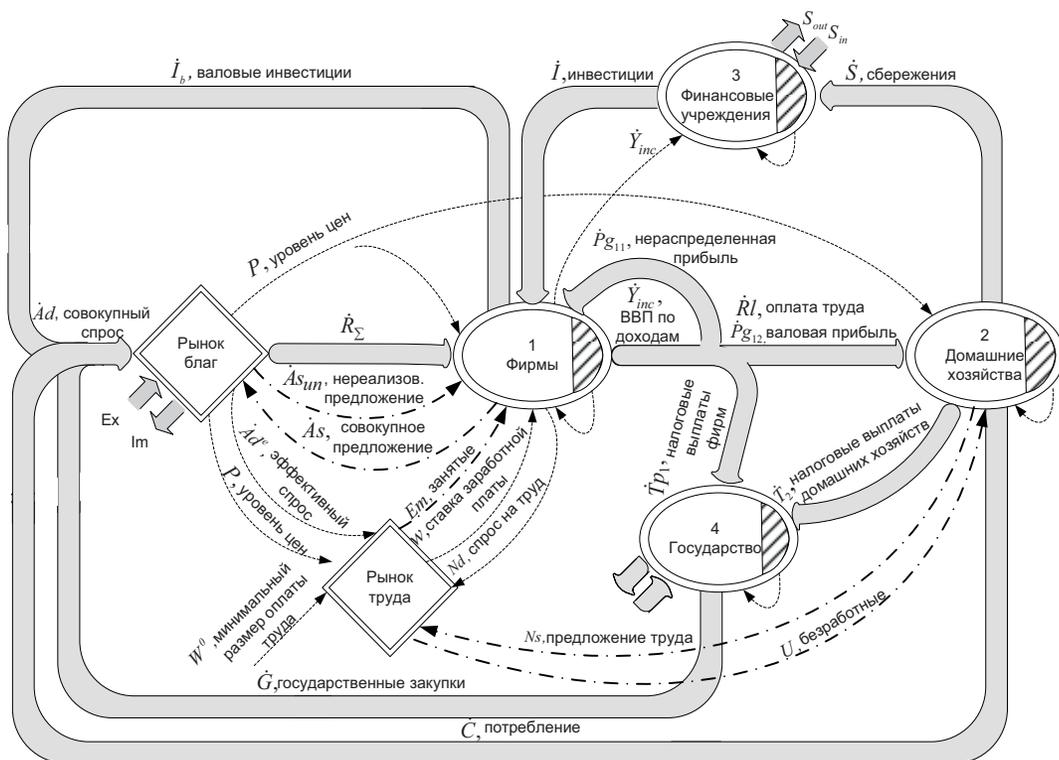


Рисунок 1 – Когнитивная модель функционирования МЭС с учетом рынков благ и труда

В отличие от рынка благ рынок труда является исключением и не обладает свойством саморегулирования согласно Кейнсу [5-8]. На рынке труда отсутствуют силы, которые бы обеспечили выравнивание спроса на труд и предложения труда. Поэтому на рынке труда устанавливается состояние неравновесия, которое приобретает устойчивый характер. В реальности, как правило, имеют место ситуации превышения предложения труда над спросом: $Ns > Nd$, где Ns – численность экономически активного населения, а Nd – численность занятых. Разность между указанными величинами, выраженная в процентах по отношению к численности экономически активного населения, определяет уровень безработицы U :

$U = \frac{Ns - Nd}{Ns} * 100\%$ [7]. Таким образом, неравновесные состояния рынка труда, характеризующиеся наличием безработицы, являются стабильными и соответствуют состоянию «квазиравновесия», которое не мешает движению к равновесию и восстановлению равновесия на оставшихся рынках.

7) Среди множества неравновесных состояний предполагается существование состояния полной занятости, при котором безработица U находится на уровне естественной безработицы U_n : $U = U_n$.

Естественная безработица включает в себя: фрикционную безработицу, связанную с инерционностью объективных процессов взаимного поиска работников требуемых специальностей со стороны фирм и подходящей работы со стороны населения; и структурную безработицу, обусловленную несоответствием структуры спроса на труд и предложения труда. Уровень естественной безработицы U_n для разных стран колеблется в пределах 3%-11% [8]. Естественной норме безработицы придается особое значение с позиций макроэкономического регулирования, поскольку считается, что принимать решения по регулированию рынка труда в условиях естественной безработицы с целью повышения занятости, не следует. Это объясняется тем, что естественная безработица рассматривается как резерв, запас трудовых ресурсов, который необходим экономике для обеспечения гибкости реагирования при решении различных проблем.

В состоянии полной занятости (при естественном уровне безработицы) производство валового внутреннего продукта (ВВП) осуществляется на уровне потенциального \dot{Y}^* . Отметим, что потенциальный ВВП \dot{Y}^* – это количество продукции, который возможно произвести за определенный период времени (темп) при полном использовании имеющихся ресурсов (капитальных, трудовых). Если безработица ниже естественной нормы, то имеет место избыточная занятость, и это свидетельствует о возможном перегреве экономики. Если безработица превышает естественную норму, то имеет место и циклическая безработица, или неполная занятость, вызванная циклическими колебаниями производства. В этих ситуациях ВВП ниже потенциального; целесообразны решения, направленные на повышение темпа производства ВВП и снижение безработицы.

8) Из множества возможных неравновесных состояний в области неполной занятости выбрано одно состояние, которое принято за базовое равновесное состояние. Считается, что на базовом равновесном режиме МЭС функционирует в условиях равновесия на рынке благ при наличии квазиравновесия на рынке труда. Используется понятие эффективного спроса $\dot{A}d^e$, введенное Кейнсом для определения совокупного спроса $\dot{A}d$ в ситуациях его равенства совокупному предложению $\dot{A}s$: $\dot{A}d = \dot{A}s = \dot{A}d^e$ [5].

9) Особенности динамики взаимодействия рынка труда и рынка благ заключаются в следующем.

Во-первых, рынок благ оказывает определяющее влияние на состояние рынка труда и является причиной его изменения. Это связано с формированием спроса на труд Nd , который зависит от эффективного спроса $\dot{A}d^e$ на рынке благ. Кроме того, влияние оказывает сформированный на рынке благ уровень цен P , который несет информацию о снижении реальной ставки заработной платы w в условиях негибкой номинальной ставки заработной платы W : $w = W/P$.

Особенности реагирования рынка труда на изменения состояния рынка благ заключаются в наличии запаздывания и инерционности в изменении спроса на труд при изменении совокупного спроса. Теоретические положения кейнсианской теории гласят о том, что рынок труда формирует «отклик» на изменение состояния рынка благ только при приближении рынка благ к новому состоянию равновесия, то есть к новому значению эффективного спроса

$\dot{A}d^e$. В начальный же период действия возмущений на рынке благ (в виде шоков совокупного спроса или совокупного предложения), когда МЭС выведена из равновесия и находится еще далеко от нового состояния равновесия (рассогласование между спросом и предложением велико), рынок труда еще не может сформировать реакцию. Наблюдается своеобразный эффект гистерезиса, обусловленный невозможностью для производителей определить новые потребности в трудовых ресурсах Nd , соответствующие новому будущему, пока неизвестному, равновесному спросу $\dot{A}d^e$.

По результатам взаимодействия сформировавшегося спроса на труд Nd , и предложения труда Ns определяется количества занятых Em и безработных Un . На когнитивной модели перечисленные связи изображены штриховыми (соответствующими информационным связям) и штрихпунктирными (соответствующими материальным потокам) линиями (рисунок 1).

Во-вторых, рынок труда оказывает опосредованное влияние на состояние рынка благ через влияние на реальный сектор и сектор домохозяйств.

Влияние рынка труда на поведение реального сектора носит информационный характер и направлено на корректировку планов производителей по выпуску ВВП на основе информации об изменении ставки заработной платы, характеризующей будущие издержки. Например, повышение ставки заработной платы приводит к росту издержек и способствует снижению совокупного предложения. В результате снижение совокупного предложения приводит к дополнительному воздействию на рынок благ и уже не может не сказаться на финансовых потоках доходов и расходов всех макроэкономических агентов, формирующих кругооборот финансовых потоков.

Влияние рынка труда на поведение сектора домохозяйств состоит в изменении располагаемого дохода и, следовательно, запасов домохозяйств под воздействием ставки заработной платы. Иными словами, домохозяйства реагируют на изменение состояния рынка труда изменением финансовых потоков доходов, запасов финансовых ресурсов и последующим за этим изменением финансовых потоков расходов на потребление, что не может не сказаться на совокупном спросе на рынке благ.

2 Функциональные схемы динамических моделей функционирования МЭС в неравновесных условиях рынков благ и труда

На обобщенной функциональной схеме (рисунок 2) динамической модели функционирования МЭС с учетом рынков благ и труда выделены шесть моделей, которые взаимосвязаны между собой потоковыми (сплошные линии) и информационными (пунктирные линии) связями.

Первые четыре модели предназначены для описания функционирования всех МЭА: реального сектора, сектора домохозяйств, секторов финансовых и государственных учреждений. Подробное описание функционального назначения каждого сектора и взаимосвязей между ними представлено в работе [3]. Модель А5 описывает функционирование рынка благ и представлена двумя блоками, предназначенными для формирования динамического уровня цен $P(t)$ на основе информации о совокупном спросе $\dot{A}d(t)$ и совокупном предложении $\dot{A}s(t)$ [4].

Модель А6 функционирования рынка труда предназначена для формирования спроса на труд Nd на основе знания предельной производительности труда $\dot{Y}'_N = \frac{d\dot{Y}}{dN}$; предложения труда Ns на основе количества экономически активного населения N_a и номинальной ставки заработной платы W^0 , соответствующей базовому равновесному режиму. Естественная норма безработицы Un применяется для определения превышения текущего уровня безработицы U над естественным. Для корректировки финансового потока оплаты труда, формируемого при

распределении ВВП, в зависимости от изменяемой ставки заработной платы w рассчитывается коэффициент k_{rlc} корректировки темпа формирования оплаты труда.

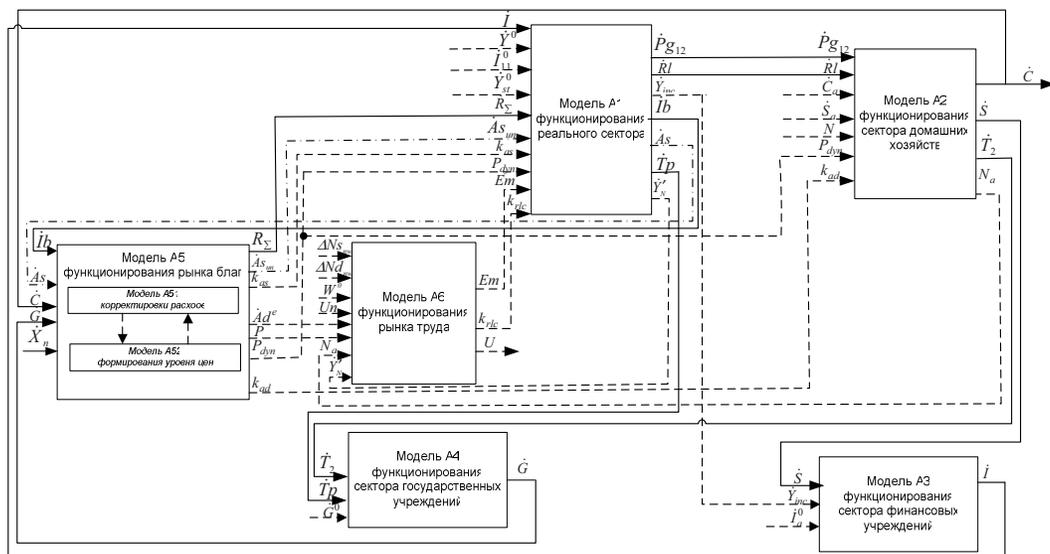


Рисунок 2 – Обобщенная функциональная схема динамической модели функционирования МЭС в неравновесных условиях рынков благ и труда

Функциональная схема динамической модели А6 функционирования рынка труда включает шесть динамических подмоделей, взаимосвязанных информационными связями (рисунок 3). Для наглядности все связи представлены сплошными линиями.

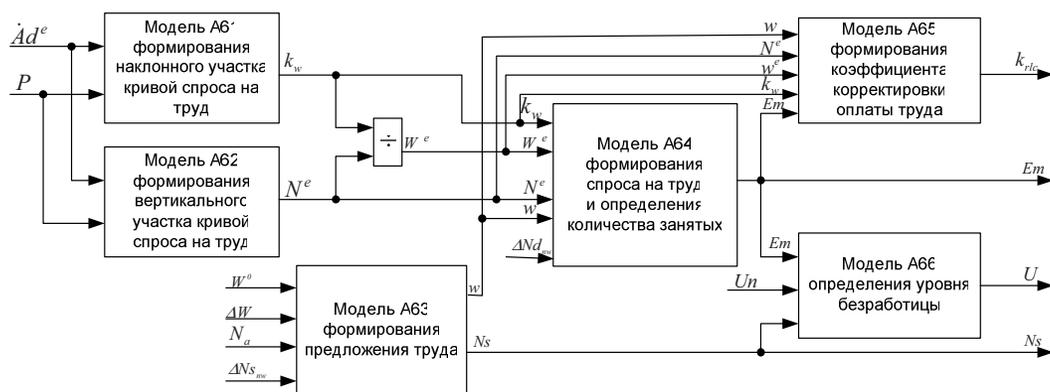


Рисунок 3 – Функциональная схема динамической модели функционирования рынка труда

Модель А61 предназначена для формирования наклонного участка кривой спроса на труд. Модель А62 предназначена для определения эффективной занятости N^e , необходимой для формирования ограничения по спросу на труд (вертикального участка кривой спроса на труд). Модель А63 формирует предложение труда N_s и текущую ставку заработной платы. На основе построенных отрезков кривой спроса на труд в модели А64 формируется единая зависимость спроса на труд $Nd = f(w, \Delta Nd_{nw}, t)$, а также с учетом сформированного предложения труда N_s , текущей ставки заработной платы W определяется уровень занятости Nd . Модель А65 позволяет рассчитать коэффициент k_{rlc} корректировки темпа формирования

оплаты труда. коэффициент k_{rlc} корректировки темпа формирования оплаты труда. В модели А66 с учетом естественной нормы безработицы Un рассчитывается текущий уровень безработицы U .

3 Особенности построения динамической модели неравновесных процессов функционирования рынка труда

Особенности моделирования функциональной зависимости *спроса на труд* $Nd = f(w, \Delta Nd_{nw}, t)$ в моделях А61, А62 состоят в следующем.

Во-первых, определяется значение эффективного спроса \dot{Y}^e и эффективной процентной ставки i^e , соответствующие совместному равновесию на рынке благ и рынке денег в виде точки пересечения кривых IS и LM [8].

Во-вторых, на основе производственной функции $\dot{Y} = F(N) = AK^{\alpha_1} N^{\alpha_2}$, где K – капитал, N – живой труд, α_1, α_2 – коэффициенты эластичности по труду и капиталу, определяется количество труда N^e , необходимое для производства благ с темпом, равным эффективному спросу \dot{Y}^e .

В-третьих, вычисляется предельный продукт труда $\frac{d\dot{Y}}{dN}$, проверяется условие эффективности дополнительного найма, а также определяется предельная ставка номинальной заработной платы W^e . Предельный продукт труда $\frac{d\dot{Y}}{dN}$ вычисляется путем

дифференцирования производственной функции: $\frac{d\dot{Y}}{dN} = \alpha_2 AK^{\alpha_1} N^{\alpha_2-1} = A\alpha_2 \frac{\dot{Y}}{N}$.

Известно, что наемные работники как важнейший фактор производства являются для работодателей источником получения прибыли, однако необходимость несения определенных издержек на рабочую силу ограничивает стремление предпринимателей к наращиванию спроса на труд.

В соответствии с теорией предельной производительности труда заработная плата работников равна стоимости продукта, которая была бы потеряна при уменьшении занятости на одну единицу [8]. Иными словами, ставка реальной заработной платы $\frac{W}{P}$ равна

предельному продукту труда $\frac{d\dot{Y}}{dN}$. Известно, что предпринимателям выгодно использовать

труд работника до тех пор, пока предельный продукт труда больше или равен ставки реальной зарплаты $\frac{W}{P}$: $\frac{d\dot{Y}}{dN} \geq \frac{W}{P}$. В этом случае предпринимателям выгодно увеличить количество

наемных работников, поскольку с каждой дополнительной единицей труда они будут получать дополнительную прибыль, равную $(P \frac{d\dot{Y}}{dN} - W)$. Когда ставка зарплаты оказывается

значительно ниже стоимости предельного продукта труда, это свидетельствует о том, что предприниматели эксплуатируют своих наемных работников. Если же $\frac{d\dot{Y}}{dN} < \frac{W}{P}$, то это

означает, что все предприятия платят наемным работникам заработную плату выше стоимости предельного продукта труда и, следовательно, несут убытки, что является недопустимым. Эти

ситуации должны быть исключены при моделировании. Если же $\frac{d\dot{Y}}{dN} = \frac{W}{P}$, то имеем предельный случай, которому соответствует предельная (максимальная) ставка реальной заработной платы $\frac{W^e}{P}$ (или предельная ставка номинальной заработной платы W^e):

$\frac{W^e}{P} = \frac{d\dot{Y}}{dN}$. При условии, что $\dot{Y}(t) = \dot{Y}^e$, то есть в ситуации, когда темп выпуска ВВП $\dot{Y}(t)$

соответствует эффективному спросу $\dot{A}d^e$, получим $\frac{W^e}{P} = A\alpha_2 \frac{\dot{Y}^e}{N^e}$ или $W^e = PA\alpha_2 \frac{\dot{Y}^e}{N^e}$.

В-четвертых, формируется график зависимости $N_d = f(w)$ в виде ломаной, состоящей из двух участков. Первый участок моделируется в блоке А62 и представляет собой прямую вертикальную линию, проведенную уровне занятости N^e , который соответствует величине эффективного спроса \dot{Y}^e и предельной ставке W^e номинальной заработной платы. Второй участок формируется в блоке А61 и представлен кривой, которая соответствует убывающей зависимости вида: $N_d = f(w) = \frac{k_w}{W}$, $k_w = A\alpha_2 P\dot{Y}$, полученной из равенства

$$\frac{W}{P} = \frac{d\dot{Y}}{dN} = A\alpha_2 \frac{\dot{Y}}{N}.$$

Отметим, что при превышении текущей ставки W предельной ставки номинальной заработной платы W^e ($W > W^e$) предприниматели будут сокращать свой спрос на труд (зависимость убывающая). Иначе, если текущая ставка W ниже предельной ставки номинальной заработной платы W^e ($W < W^e$) спрос сохраняется на прежнем уровне N^e .

Поскольку в кейнсианской концепции спроса на труд регулятором рынка труда является не цена труда (зарботная плата), а эффективный спрос, определяемый конъюнктурой рынка благ и рынка денег, влиять на состояние рынка труда (спрос на труд и ставку реальной заработной платы) может только государство. Как известно, эффективный спрос определяется пересечением линий IS и LM , то есть на него можно влиять путем увеличения или уменьшения составляющих совокупного спроса (например, снижение налогов стимулирует рост потребительского спроса, что в свою очередь, приводит к росту производства и занятости) для сдвига кривой IS или количества денег для сдвига кривой LM .

Особенности формирования в блоке А63 функциональной зависимости *предложения труда* $N_s = f(w, \Delta N_{s_{nw}}, t)$ состоят в следующем.

Кривая предложения труда так же, как и кривая спроса, представляет собой ломаную, состоящую из двух участков. Первый участок кривой представляет собой функцию предложения труда в ситуации неполной занятости. В этой ситуации функция предложения труда $N_s = f(W)$ стабильно возрастает при росте ставки номинальной заработной платы W .

При достижении полной занятости N_s^* , соответствующей потенциальному выпуску, ситуация меняется и предложение труда становится близким к совершенно эластичной зависимости, что и представляет собой второй участок кривой предложения труда.

Таким образом, снижение уровня безработицы возможно путем изменения ситуации на рынке благ и денег при проведении мер государственного регулирования, направленных на рост компонентов эффективного спроса \dot{Y}^e , который и будет определять темп выпуска ВВП.

Динамическая модель функционирования рынка труда представлена системой нелинейных дифференциальных уравнений первого порядка.

4 Экспериментальные исследования динамики макроэкономического кругооборота финансовых потоков в неравновесных условиях рынка труда

Предложенные модели положены в основу экспериментальных исследований влияния неравновесных условий рынка труда на динамику макроэкономического кругооборота финансовых потоков в рамках макроэкономической системы. Результаты экспериментальных исследований представлены на рисунке 4.

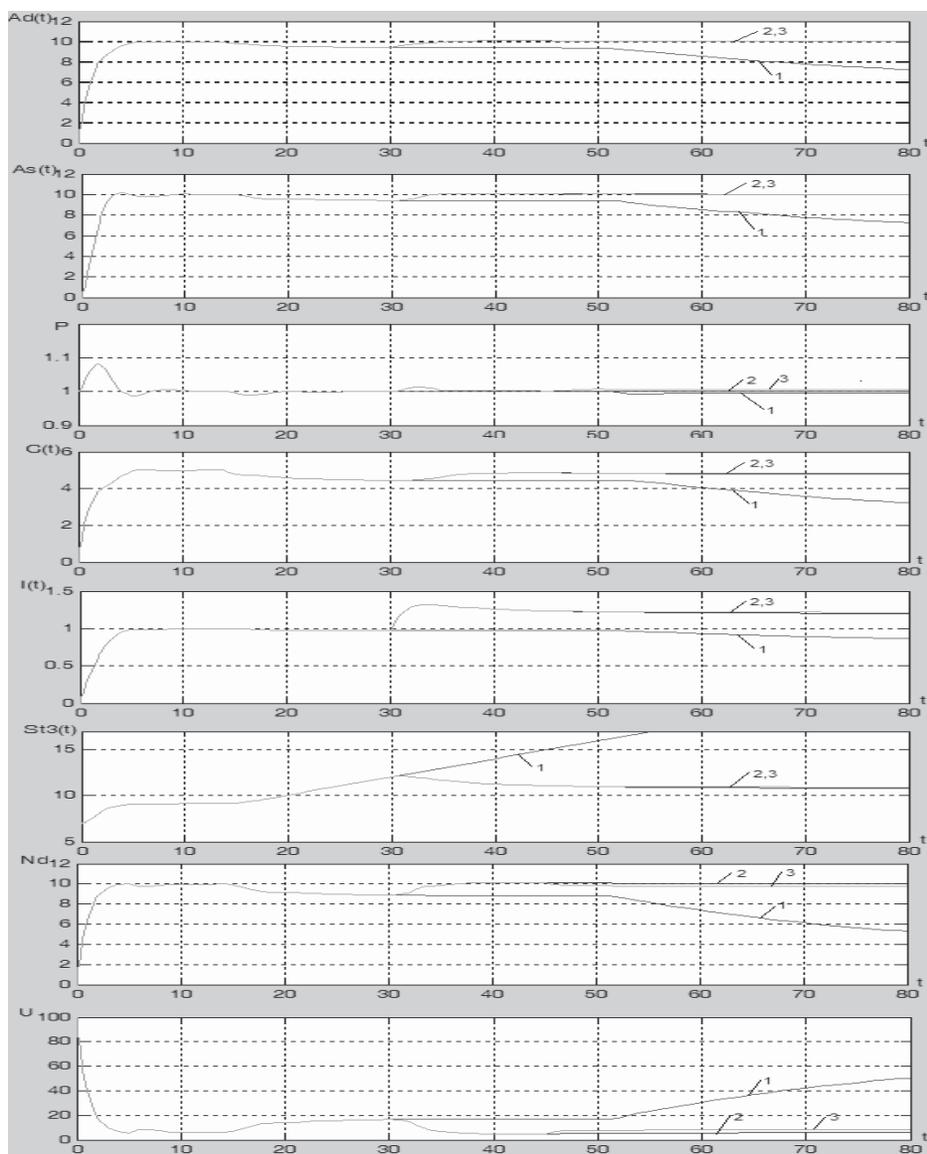


Рисунок 4 – Графики переходных процессов функционирования реального сектора, секторов домохозяйств и финансовых учреждений, рынка труда

В качестве единицы модельного времени выбран месяц. Моделирование проводится для среднесрочного периода. Время моделирования равно 80 месяцам. Установлены следующие начальные значения плановых темпов расхода ресурсов, для формирования предложения

$\dot{A}s^0(t) = 10$; формирования автономного потребления $\dot{C}^0_a(t) = 2$; для автономного сбережения $\dot{S}_a(t) = 1,5$; для автономных инвестиций $\dot{I}^0_a(t) = 0,5$; для формирования госзакупок $\dot{G}^0(t) = 3$. При распределении ВВП приняты следующие значения коэффициентов формирования: оплаты труда $k_{rl} = 0,4$, валовой прибыли $k_{pg} = 0,4$; налогов $k_t = 0,2$.

Сценарий 1 является неуправляемым. В момент времени $t=14$ предполагается подача возмущения в секторе домохозяйств в виде снижения темпов формирования автономного потребления $\Delta \dot{C}_a(t) = -1$. Темп формирования потребления как составляющая темпа формирования совокупного спроса снижает последний. Следствием этого, в первую очередь, станет реакция рынка благ, где наблюдается неравновесная ситуация, когда совокупное предложение превышает совокупный спрос ($\dot{A}s > \dot{A}d$). В такой ситуации в реальном секторе образуются излишки продукции, которые приобретают форму товарно-материальных запасов. Представленная неравновесная ситуация влечет снижение цены на блага в $t=14,5$, что приводит к сокращению объемов производства, то есть совокупного предложения $\dot{A}s$ в $t=15$, а также приводит к сокращению налоговых поступлений от реального сектора. Последним на изменение темпа формирования совокупного спроса $\dot{A}d$ отреагирует рынок труда. Темп формирования совокупного предложения сократился и, следовательно, уменьшилось количество рабочих мест, что эквивалентно резкому падению спроса на рынке труда Nd в момент времени $t=16$ и одновременному росту безработицы U . Отметим, что значение спроса на труд оказывает влияние на планируемый темп производства благ реальным сектором. К моменту времени $t=32$ равновесие на рынке благ восстанавливается, хотя и на более низком уровне. Установившееся равновесное значение совокупного спроса и совокупного предложения является величиной эффективного спроса \dot{Y}^e . Более низкая величина эффективного спроса \dot{Y}^e приведет к сокращению количества занятых, требуемое для его удовлетворения N^e . Взаимное изменение темпа формирования совокупного спроса Nd на рынке труда и количества занятых, требуемого для удовлетворения эффективного спроса N^e , приведет к снижению минимальной ставки реальной заработной платы. Сектор домохозяйств в этом случае получает меньше доходов, что приводит к сокращению налоговых поступлений в бюджет государства. Существенное сокращение налоговых поступлений от реального сектора и сектора домохозяйств привели к тому, что государственный сектор постоянно расходует свои запасы и, исчерпав их, в момент времени $t=60$ вынужден сократить величину госзакупок. Такая ситуация спровоцировала значительный «обвал» (снижение запасов) всех секторов за исключением сектора финансовых учреждений.

Сценарий 2 является управляемым. Управление заключается в предоставлении дополнительных инвестиций реальному сектору со стороны сектора финансовых учреждений в момент времени $t=40$. Увеличение инвестиций со стороны сектора финансовых учреждений привело к росту валовых инвестиций, которые являются составляющей совокупного спроса. Следовательно, темп формирования совокупного спроса $\dot{A}d$ увеличился. На увеличение темпа формирования совокупного спроса отреагировал рынок благ повышением цены, так как возникла ситуация дефицита ($\dot{A}d > \dot{A}s$). Реальный сектор увеличил производство. Для реализации этого потребовались новые рабочие места, таким образом, увеличился спрос на рынке труда Nd и снизился уровень безработицы. Увеличение спроса привело к росту доходов населения в виде получаемой заработной платы, следовательно, увеличились выплачиваемые государству налоги. Величина налоговых поступлений в казну государства увеличилась также и со стороны реального сектора. Рост вызван ростом темпов производства. За счет поступлений в бюджет государство сохраняет величину госзакупок на требуемом

уровне. Однако рост цен на рынке благ привел к уменьшению реальных доходов населения, вследствие чего сектор домохозяйств вынужден сократить потребление. Эта ситуация привела к взаимосвязанному падению совокупного спроса, совокупного предложения, спроса на труд и росту безработицы. В результате взаимосвязанного изменения поведения секторов МЭС наблюдается установление общего макроэкономического равновесия на первоначальном уровне. Принятие решения в виде выделения дополнительного финансирования сектором финансовых учреждений реальному сектору благоприятно сказалось на поведении всей МЭС в целом, позволив избежать падения уровня ВВП.

Сценарий 3 также является управляемым. В момент времени $t = 45$ подано возмущение на рынок труда в виде увеличения ставки заработной платы W . В этом случае реальный сектор, вынужденный увеличить свои расходы, сократит спрос на рабочую силу Nd , что повлечет за собой рост безработицы. Однако при этом наблюдается положительный эффект в виде роста доходов занятых в экономике.

Заключение

Предложена когнитивная модель функционирования макроэкономической системы с учетом рынков благ и труда, отражающая влияние рынков на формирование финансовых потоков макроэкономического кругооборота. Роль рынка труда в процессах функционирования МЭС состоит как в корректировке плановых темпов производства ВВП на основе информации об изменении ставки заработной платы, так и в корректировке фактических темпов формирования оплаты труда при распределении произведенного ВВП. Показано, что неравновесные ситуации на рынке труда определяются взаимосвязанными изменениями совокупного спроса и совокупного предложения на рынке благ. Предложена динамическая модель функционирования рынка труда, позволяющая отразить характер влияния неравновесных ситуаций на рынке труда на динамику макроэкономического кругооборота финансовых потоков. Проведенные экспериментальные исследования продемонстрировали взаимное влияние динамики доходов и расходов секторов МЭС, формирующих кругооборот финансовых потоков, на неравновесные состояния рынка труда; а управляемые сценарии показали возможность снижения уровня безработицы и повышения доходов населения за счет принятия решений по увеличению составляющих совокупного спроса под влиянием роста инвестиционных вложений.

Список литературы

- [1] Петров А.А., Поспелов И.Г. Математические модели экономики // Вестник РАН, 2009. Том 9, №6. С 492-506.
- [2] Прогноз и моделирование кризисов и мировой динамики / Отв. ред. А.А. Акаев, А.В. Коротаев, Г.Г. Малинецкий. – М.: Издательство ЛКИ, 2010. – 352 с.
- [3] Ильясов Б.Г., Дегтярева И.В., Макарова Е.А., Габдуллина Э.Р. Моделирование динамики кругооборота финансовых потоков с учетом накопления финансовых ресурсов // Вестник компьютерных и информационных технологий. – 2009. - №1 (61). - С. 28-38.
- [4] Ильясов Б.Г., Макарова Е.А., Габдуллина Э.Р., Павлова А.Н. Динамическая модель формирования уровня цен в неравновесных условиях рынка благ // Проблемы управления и моделирования в сложных системах: Труды XI Международной конференции – Самара: Самарский научный центр РАН, 2009. – С. 203-211
- [5] Маневич В.Е. Кейнсианская теория и российская экономика. Изд.2-е, перераб. и доп. М. КомКнига, 2010. – 224 с.
- [6] Селищев А.С. Макроэкономика – СПб.: «Издательство «Питер», 2005. – 448с.
- [7] Макроэкономика: научные школы, концепции, экономическая политика: учеб. пособие / А.А. Никифоров, О.А. Антипина, Н.А. Миклашевская. М.: Издательство «Дело и сервис», 2008. 534 с.
- [8] Макроэкономика: учебник / Л.С. Тарасевич, П.И.Гребенников, А.И. Леусский. – 7-е изд. испр. и доп. – М.: Высшее образование, Юрайт – Издат., 2009. – 654 с.

АЛГОРИТМ СИНТЕЗА РЕГУЛЯТОРА АДАПТИВНОЙ МНОГОСВЯЗНОЙ СИСТЕМЫ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ НА ОСНОВЕ АНАЛИЗА ХАРАКТЕРИСТИЧЕСКИХ ФУНКЦИЙ

Б.Г. Ильясов¹, А.Ш. Назаров¹, Г.А. Саитова²

¹ГОУ ВПО Уфимский государственный авиационный технический университет
450000, Уфа, ул. К. Маркса, 12, Россия

²ИМех УфНЦ РАН
Уфа, Проспект Октября 71, Россия
anv.naz@mail.ru
тел: +7 (347) 273-78-35

Ключевые слова: *Управление в адаптивных системах, многосвязная система автоматического управления, алгоритм синтеза, многосвязное регулирование*

Abstract

The given article is dedicated to problems of designing an adaptive multivariable automatic control systems. Such us synthesis of connections between separate subsystems of multivariable automatic control systems at adaptive multivariable control of multivariable multiplemode object.

Введение

Высокая эффективность работы адаптивной МСАУ сложным многорежимным динамическим объектом (СДО) может быть достигнута также за счет введения искусственных связей между сепаратными подсистемами, носящих, например, коалиционный характер, с целью приспособления к изменяющимся режимам функционирования [6]. В этом случае говорят о связанном управлении сепаратными подсистемами МСАУ через многомерный регулятор. Поэтому в целях улучшения динамических свойств системы, обеспечения структурной устойчивости МСАУ на всех режимах ее функционирования, повышения запасов устойчивости, стабилизации на критических режимах функционирования, при возможных отказах в сепаратных подсистемах актуальным является проектирование систем управления СДО с использованием дополнительных корректирующих и стабилизирующих связей между отдельными ее каналами, образующих многомерный регулятор.

Задача адаптации МСАУ к различным режимам ее функционирования путем введения в нее дополнительных связей между сепаратными подсистемами также относится к классу задач адаптации МСАУ. Для решения данной задачи необходимо, во-первых, определить условия функционирования МСАУ, а также ее характеристики, при которых целесообразно осуществлять данный способ адаптации, во-вторых, предложить формализованный алгоритм структурного синтеза ее многосвязного регулятора.

Для решения данной задачи предлагается использовать частотный критерий устойчивости, основанный на системном описании МСАУ на уровне сепаратных подсистем и многомерных элементов связи между ними [5]. Суть метода заключается в нахождении корней μ_i ($i=1,2,\dots,n$, где n – размерность МСАУ) характеристического уравнения связи МСАУ и установлении факта неохвата их годографом АФХ $\Phi(j\omega)$ сепаратных подсистем при $\omega \in (0, +\infty)$ [3].

Алгоритм синтеза регулятора адаптивной МСАУ

Рассмотрим режим работы МСАУ, на котором происходит некоторое отклонение ее динамических характеристик от желаемых, при которых система становится либо неустойчивой, либо находится на границе устойчивости, либо не удовлетворяет заданным требованиям к качеству управления. При этом, если преимущественно изменяются характеристики связи отдельных подсистем, определяемые расположением корней μ_i относительно осей комплексной плоскости, а форма и положение АФХ подсистем $\Phi(j\omega)$ существенно не меняются: $\Phi(j\omega) \approx \Phi^*(j\omega)$; $\mu_i \neq \mu_i^*$, тогда требуемые преобразования структуры МСАУ необходимо проводить путем введения дополнительных связей R_{ij} в структуру многомерного регулятора, так как основное влияние на устойчивость и качество МСАУ в данном случае оказывают передаточные функции (ПФ) связей между локальными подсистемами. Сформулированное выше условие и будем считать условием, при котором целесообразно осуществлять предложенный метод адаптивного управления МСАУ.

Задачу структурного синтеза связей в многомерном регуляторе адаптивной МСАУ, структурная схема которой представлена на рисунке 1, из условий обеспечения желаемых динамических характеристик, предлагается решить следующим образом.

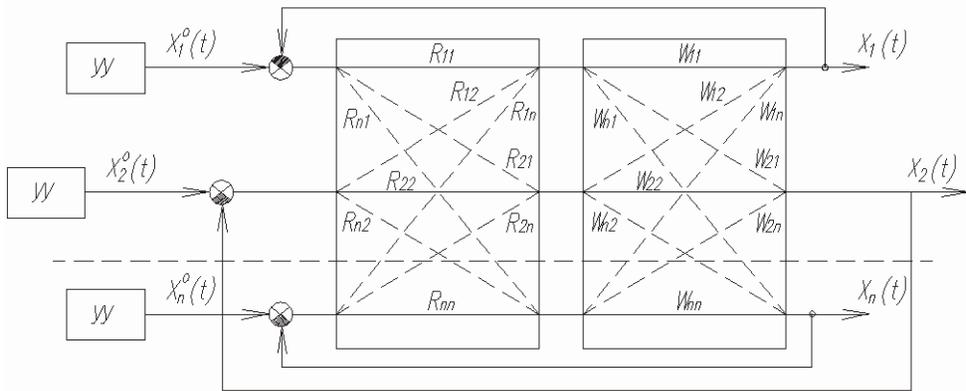


Рисунок 1 – Структурная схема n -мерной МСАУ с n -мерным регулятором

Пусть для исследуемой МСАУ, передаточная функция n -мерного объекта управления и n -мерного регулятора описываются в форме матричных ПФ (МПФ) размерности $n \times n$:

$$W(s) = \|W_{ij}(s)\|_{n \times n} = \begin{bmatrix} W_{11} & \dots & W_{1i} & \dots & W_{1n} \\ \vdots & & \vdots & & \vdots \\ W_{i1} & \dots & W_{ii} & \dots & W_{in} \\ \vdots & & \vdots & & \vdots \\ W_{n1} & \dots & W_{ni} & \dots & W_{nn} \end{bmatrix}, \quad R(s) = \|R_{ij}(s)\|_{n \times n} = \begin{bmatrix} R_{11} & \dots & R_{1i} & \dots & R_{1n} \\ \vdots & & \vdots & & \vdots \\ R_{i1} & \dots & R_{ii} & \dots & R_{in} \\ \vdots & & \vdots & & \vdots \\ R_{n1} & \dots & R_{ni} & \dots & R_{nn} \end{bmatrix}.$$

Пусть заданными являются:

- структура неизменяемой части системы: объекта управления, регуляторов отдельных подсистем, определяемых как диагональные элементы МПФ $R(s)$: $diag(R(s)) = \{R_{11}(s), R_{22}(s), \dots, R_{nn}(s)\}$, удовлетворяющей условию полной управляемости;
- список корректирующих звеньев;
- ограничения на способы соединения указанных структур друг с другом.

При этом предполагается, что заданная часть системы и корректирующие звенья описываются системой линейных дифференциальных уравнений и могут быть представлены соот-

ветствующими ПФ. В общем случае задача заключается в том, что бы найти некоторый алгоритм формирования или изменения связей R_{ij} (при $i \neq j$) в многомерном регуляторе $R(s)$ – алгоритм реструктуризации, с тем, чтобы удовлетворить заданным требованиям к обеспечению желаемых динамических свойств системы, а также к запасам ее устойчивости.

Будем полагать, что исследуемая система относится к классу МСАУ с однотипными характеристиками подсистем, а требуемые динамические свойства в системе обеспечиваются заданным распределением на комплексной плоскости корней μ_i характеристического уравнения связи подсистем, которое можно представить в виде некоторого характеристического полинома (функции) $D(x)$, т.е. выполнением условия:

$$(1) \quad \begin{cases} \operatorname{Re}(\mu_i) \rightarrow \operatorname{Re}(\mu_i^*) \\ \operatorname{Im}(\mu_i) \rightarrow \operatorname{Im}(\mu_i^*) \end{cases},$$

где $i \in \overline{1, n}$, где n – размерность исследуемой МСАУ. Очевидно, что условие (1) выполняется если выполняется условие:

$$(2) \quad D^*(x) = D^*(x), \text{ где } D(x) = 1 + a_2 x^2 + \dots + a_n x^n; \quad D^*(x) = 1 + a_2^* x^2 + \dots + a_n^* x^n;$$

$D^*(x)$ – желаемая характеристическая функция с заданным распределением корней μ_i^* , при которых МСАУ на данном режиме статически и динамически устойчива и отвечает желаемым требованиям к качеству управления.

Известно, что при решении задачи синтеза такого рода в первую очередь выделяют ряд принципиальных требований, которым должна удовлетворять допустимая структура управляющей части [2]. В соответствии с [2] синтезированная структура управляющей части системы считается допустимой, а алгоритм ее синтеза корректным, если обеспечивается выполнение следующих условий синтеза:

- 1) математическая корректность задачи синтеза;
- 2) структурная устойчивость системы;
- 3) астатизм заданного порядка;
- 4) физическая реализуемость управляющей части.

Поэтому дальнейший синтез требуемой многосвязной структуры регулятора $R(s)$ целесообразно вести с учетом перечисленных требований.

Исходя из вышесказанного, формализуем алгоритм синтеза корректирующего устройства на основе анализа характеристических функций $D(x)$ и докажем его корректность.

На первом этапе запишем характеристическое уравнение МСАУ, в которой кроме естественных связей через многомерный объект управления (МОУ) имеются искусственные связи через многомерный регулятор, согласно выражению 2.5 оно будет иметь следующий вид:

$$(3) \quad D(h_\Sigma, \Phi) = 1 + \sum_{i,j=1}^n h_{\Sigma ij} \Phi^2 + \sum_{i,j,k=1}^n h_{\Sigma ijk} \Phi^3 + \dots + h_{\Sigma 1n} \Phi^n = 0,$$

$$\text{где } n - \text{порядок МСАУ, } h_{\Sigma 1,2,3\dots k} = \frac{\det \left\| \left\| W(s)R(s) \right\|_{ij} \gamma_{ij} \right\|_{kxk}}{\det \left\| \left\| W(s)R(s) \right\|_{ij} \delta_{ij} \right\|_{kxk}} - \text{характеристика связи между под-}$$

системами, учитывающая влияние, как естественных связей через МОУ, так и искусственных связей через многомерный регулятор. Положим, что каждому элементу a_i характеристической функции $D(x)$ ставится в соответствие свое значение $h_{\Sigma 1,2,3\dots k}$, то есть $a_i = h_{\Sigma 1,2,3\dots k}$. Тогда, искомое решение будем искать из условия наилучшего приближения (2), то есть равенства коэффициентов $a_i = a_i^*$ при одинаковых степенях x характеристических функций $D(x)$ и $D^*(x)$.

На втором этапе формируется система нелинейных алгебраических уравнений вида:

$$(4) \quad a_i = f(R_{ij}), \text{ где } i \in \overline{1, n}.$$

То есть целью введения искусственных перекрестных связей R_{ij} в МСАУ является поддержание на постоянном уровне коэффициентов характеристической функции $D(x)$ МСАУ. В общем случае условие (4) для наглядности можно записать через характеристики связи $h_{\Sigma 1,2,3\dots k}$ между подсистемами в виде системы уравнений:

$$(5) \quad \begin{cases} h_{\Sigma 1,2}(W_{11}, \dots, W_{1n}, W_{21}, \dots, W_{2n}, R_{11}, \dots, R_{1n}, R_{12}, \dots, R_{n2}) = h^*_{\Sigma 1,2}(W_{11}, W_{12}, W_{21}, W_{22}); \\ \dots \\ h_{\Sigma n-1,n}(W_{n-1,1}, \dots, W_{n-1,n}, W_{n1}, \dots, W_{nn}, R_{1n-1}, \dots, R_{nn-1}, R_{1n}, \dots, R_{nn}) = h^*_{\Sigma 1,n}(W_{n-1n-1}, W_{n-1n}, W_{nn-1}, W_{nn}); \\ \dots \\ h_{\Sigma 1,2,3\dots n}(W_{11}, \dots, W_{nn}, R_{11}, \dots, R_{1n}) = h^*_{\Sigma 1,2,3\dots n}(W_{11}, \dots, W_{nn}). \end{cases}$$

В результате решения системы (5) на этапе параметрического синтеза из условия физической реализуемости определяются искомые ПФ перекрестных связей R_{ij} корректирующего устройства.

Далее для обеспечения математической корректности процедуры синтеза и структурной устойчивости системы в целом введем некоторые ограничения на способы соединения отдельных подсистем в многомерном регуляторе $R(s)$ друг с другом, а также на вид их ПФ. Известно, что одним из критериев структурного синтеза является разрешимость уравнений (5) в базе выбранной структуры [2]. Это означает, что количество r синтезируемых перекрестных связей не должно превышать значения n , т.е. $r < n$. Необходимо отметить, что данное условие вполне выполнимо, если в процессе анализа МСАУ на исследуемом режиме выяснить, какие связи (между парами $h_{\Sigma ij}$, тройками $h_{\Sigma jik}$ ит.д. подсистем) оказывают отрицательное влияние на запасы устойчивости и качество системы в целом. Дальнейший синтез структуры корректирующего устройства между подсистемами вести только для данных подсистем. Скажем, что подобный анализ весьма не трудно провести, используя предложенную в [7] методику анализа МСАУ. Это же замечание справедливо с точки зрения еще одного условия корректности задачи синтеза, которое также предполагает, что синтезируемая структура должна соответствовать критерию минимальной сложности управляющей части системы [2].

Одним из основных критериев структурной устойчивости МСАУ согласно [7] является отсутствие в системе стимулирующих контуров с положительной обратной связью, поэтому еще одним условием введения в систему корректирующих связей является положительный знак их ПФ.

Также стоит отметить, что при подобном способе адаптации путем введения в систему дополнительных корректирующих связей не нарушается заданный порядок астатизма подсистем, который в данном случае обеспечивается структурой регуляторов отдельных подсистем, которая не изменяется в процессе синтеза связей.

Пример. Рассмотрим систему управления трехсвязным динамическим нестационарным объектом. Модель МСАУ исследуемым объектом, также представлена на рисунке 1, при $n=3$.

Рассмотрим два основных режима работы исследуемой МСАУ.

Базовый режим работы объекта управления – режим 1. Пусть матричная передаточная функция многомерного объекта на заданном режиме равна:

$$W(s) = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{13} & W_{32} & W_{33} \end{bmatrix} = \frac{1}{0.6s + 1} \begin{bmatrix} 0.5 & 0.6 & -0.15 \\ 0.25 & 0.6 & -0.1 \\ 0.8 & -0.6 & 0.15 \end{bmatrix}.$$

Пусть желаемые индивидуальные характеристики $\Phi_i(s)$ подсистем описываются дробно-рациональной функцией: $\Phi_i(s) = 1/(\tau^2 s^2 + 2\xi\tau s + 1)$, где $\tau = 0.5$ с, $\xi_1 = 0.8$. Тогда синтезированные регуляторы отдельных подсистем, обеспечивающие желаемые АФХ, имеют следующий вид: $R_{ii}(s) = K_{ii}(T_0 s + 1)/s$, где $K_{11} = 2$; $K_{22} = 1.7$; $K_{33} = 7.7$; $T_0 = 0.6$.

На рисунке 2 приведен результат моделирования исследуемой МСАУ по координате X_1

Произведем в выражении для $\Phi_i(s)$ замену $s=j\omega$, построим годограф $\Phi(j\omega)$ отдельных подсистем МСАУ.

Затем запишем характеристическое уравнение (3) исследуемой МСАУ:

$$1 + h_2 \Phi^2(s) + h_3 \Phi^3(s) = 0, \text{ где } h_2 = \sum_{i,j=1}^n \frac{\det \|W_{ij}(s) \cdot \gamma_{ij}\|_{2 \times 2}}{\det \|W_{ij}(s) \cdot \delta_{ij}\|_{2 \times 2}} = 0.433; \quad h_3 = \frac{\det \|W_{ij}(s) \cdot \gamma_{ij}\|_{3 \times 3}}{\det \|W_{ij}(s) \cdot \delta_{ij}\|_{3 \times 3}} = -0.567$$

интегральные характеристики связи отдельных подсистем,

$$\gamma_{ij} = \begin{cases} 1, & i \neq j \\ 0, & i = j \end{cases} \quad k = 2,3; \quad \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad k = 2,3, \quad \Phi(s) - \text{передаточная функция отдельных}$$

подсистем. Произведем в данном уравнении замену $x = \Phi(s)$, определим корни μ_i , решив его относительно переменной x .

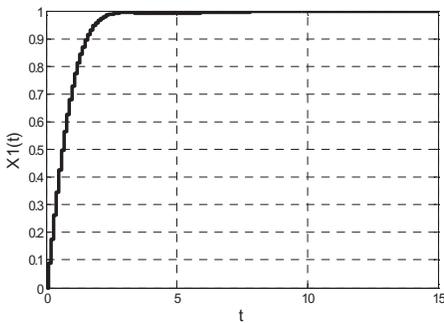


Рисунок 2 – Результаты моделирования переходных процессов координаты X_1 исследуемой МСАУ на расчетном режиме 1

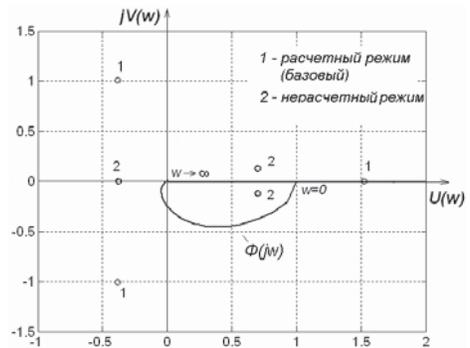


Рисунок 3 – Взаимное расположение годографа отдельных подсистем МСАУ $\Phi(j\omega)$ и корней μ_i на исследуемых режимах

На рисунке 3 приведено взаимное расположение годографа отдельных подсистем МСАУ $\Phi(j\omega)$ и корней μ_i характеристического уравнения связи системы. В соответствии с рисунком 3 можно сделать вывод, что на расчетном режиме – 1 исследуемая МСАУ устойчива, так как годограф $\Phi(j\omega)$ не охватывает ни один из корней характеристического уравнения системы [3].

Далее рассмотрим нерасчетный (для данного режима регуляторы локальных подсистем не синтезировались) режим работы объекта управления.

Запишем передаточную функцию многомерного объекта для данного режима – 2:

$$(6) \quad W(s) = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{13} & W_{32} & W_{33} \end{bmatrix} = \frac{1}{0.6s + 1} \begin{bmatrix} 0.5 & 0.6 & -0.15 \\ 3.25 & 0.6 & -0.1 \\ 0.8 & -0.6 & 0.15 \end{bmatrix}.$$

Решим задачу адаптации исследуемой МСАУ к ее новому режиму работы – режим 2, при условии, что индивидуальные характеристики $\Phi(j\omega)$ подсистем не изменяются.

На первом этапе решения задачи проведем анализ устойчивости МСАУ аналогично проведенному выше для базового режима.

Анализ взаимного расположения годографа сепаратных подсистем МСАУ $\Phi(j\omega)$ и корней μ_i характеристического уравнения связи системы для режима работы 2, представленного на рисунке 3, позволяет сделать вывод, что исследуемая система неустойчива, так как годограф $\Phi(j\omega)$ охватывает корни характеристического уравнения связи системы.

Из выражения $D(h(0), \Phi(0)) = 1 + \sum_{k=2}^n h_k(0) = 0.85$ можно сделать вывод о статической устойчивости МСАУ на данном режиме [1].

Третий вывод заключается в том, что сделанное нами допущение о незначительном изменении характеристики $\Phi(j\omega)$, делает достаточным адаптацию МСАУ за счет изменения связей между подсистемами.

Желаемые изменения расположения корней характеристического уравнения связи между локальными подсистемами будем осуществлять при помощи введения перекрестных связей между локальными регуляторами из условия (5) для данного режима. Анализ влияния связей между подсистемами на устойчивость МСАУ показал, что отрицательное влияние на запасы устойчивости системы в целом оказывает характеристика связи h_{12} , поэтому дальнейший синтез стабилизирующих ХС между подсистемами будем вести только для 1,2 подсистем. С учетом вышесказанного синтезируем искусственные стабилизирующие связи R_{12} и R_{21} между соответствующими локальными подсистемами исследуемой МСАУ. Запишем выражение для характеристики связи $h_{\Sigma 1,2,3}$ между подсистемами исследуемой МСАУ.

$$h_{\Sigma 1,2,3} = \frac{\det \left\| \left\| W(s)R(s) \right\|_{ij} \gamma_{ij} \right\|_{3 \times 3}}{\det \left\| \left\| W(s)R(s) \right\|_{ij} \delta_{ij} \right\|_{3 \times 3}} = \frac{\det \begin{vmatrix} 0 & W_{11}R_{12} + W_{12}R_{22} & W_{13}R_{33} \\ W_{21}R_{11} + W_{22}R_{21} & 0 & W_{23}R_{33} \\ W_{31}R_{11} + W_{32}R_{21} & W_{31}R_{12} + W_{32}R_{22} & 0 \end{vmatrix}}{\det \begin{vmatrix} W_{11}R_{11} + W_{12}R_{21} & 0 & 0 \\ 0 & W_{21}R_{12} + W_{22}R_{22} & 0 \\ 0 & 0 & W_{33}R_{33} \end{vmatrix}}_{3 \times 3},$$

при этом МПФ многомерного регулятора $R(s)$ имеет следующий вид:

$$R(s) = \begin{bmatrix} R_{11} & R_{12} & 0 \\ R_{21} & R_{22} & 0 \\ 0 & 0 & R_{33} \end{bmatrix}.$$

Тогда система уравнений (5) для исследуемой МСАУ имеет следующий вид:

$$\begin{cases} \frac{-(W_{11}R_{12} + W_{12})(W_{21} + W_{22}R_{21})}{(W_{11} + W_{12}R_{21})(W_{21}R_{12} + W_{22})} = -0.5 \\ \frac{-W_{13}(W_{31} + W_{32}R_{21})}{(W_{11} + W_{12}R_{21})W_{33}} = 1.6 \\ \frac{-W_{23}(W_{31}R_{12} + W_{32})}{(W_{21}R_{12} + W_{22})W_{33}} = -0.667 \\ \frac{[(W_{13}W_{21}W_{31}R_{12} + W_{21}W_{13}W_{32} + W_{13}W_{22}R_{21}W_{31}R_{12} + W_{13}W_{22}R_{21}W_{32} + W_{23}W_{31}W_{11}R_{12} + W_{31}W_{12}W_{23} + W_{23}W_{32}R_{21}W_{11}R_{12} + W_{23}W_{32}R_{21}W_{12})]}{[(W_{11} + W_{12}R_{21})(W_{21}R_{12} + W_{22})W_{33}]} = -0.567 \end{cases}$$

Уравнения полученной системы также упрощаются, если организовывать стабилизирующий многомерный регулятор не между локальными регуляторами подсистем, а перед ними, при этом коэффициенты усиления прямых регуляторов R_{ii} многомерного регулятора R выбирать единичными $K_{Rii} = 1$. С учетом вышесказанного можно принять: $R_{11} = R_{22} = R_{33} = 1$, тогда полученная система уравнений с учетом выражения (6), будет иметь следующий вид:

$$\left\{ \begin{array}{l} \frac{-(0.5R_{12} + 0.6)(3.25 + 0.6R_{21})}{(0.5 + 0.6R_{21})(3.25R_{12} + 0.6)} = -0.5 \\ \frac{(0.12 - 0.09R_{21})}{(0.075 + 0.09R_{21})} = 1.6 \\ \frac{(-0.06 + 0.08R_{12})}{(0.488R_{12} + 0.09)} = -0.667 \\ \frac{(0.244 - 0.43R_{12} + 0.09R_{21} - 0.42R_{21}R_{12})}{(0.5 + 0.6R_{21})(0.488R_{12} + 0.09)} = -0.567 \end{array} \right. \Leftrightarrow$$

$$\left\{ \begin{array}{l} \frac{-60580R_{12}^2 + 70311R_{12} + 94200R_{12}^2R_{21} + 64206R_{12}R_{21} + 15030 + 8640R_{21}}{(5 + 6R_{21})(65R_{12} + 12)(244R_{12} + 45)} = -0.433 \\ \frac{0.244 - 0.43R_{12} + 0.09R_{21} - 0.042R_{21}R_{12}}{(0.5 + 0.6R_{21})(0.488R_{12} + 0.09)} = -0.567 \end{array} \right.$$

Решая полученную систему, получаем следующие значения стабилизирующих связей: $R_{12} = -0.2, 0.56; R_{21} = -3.42, -0.56$.

Оценим полученные связи с учетом указанных выше ограничений. Как показывает анализ получившихся контуров управления, представленных на рисунке 4, введение в систему связей $R_{12} = -0.2; R_{21} = -3.42; -0.56$ приводит к возникновению контуров, образованных каналами $R_{21}W_{12}$ и $R_{12}W_{21}$, с сильной положительно обратной связью, которые приводят к структурной неустойчивости всей МСАУ. Поэтому задачу обеспечения устойчивости исследуемой МСАУ на режиме 2 будем решать путем введения в нее стабилизирующей связи $R_{12} = 0.56$.

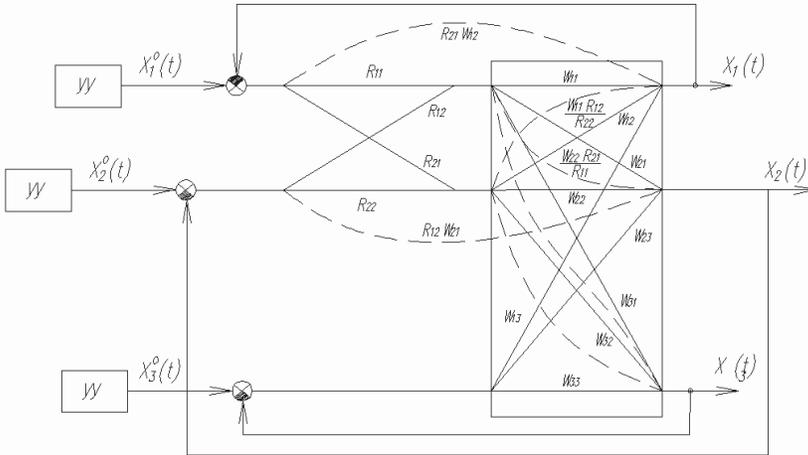


Рисунок 4 – Исследуемая трехмерная МСАУ с дополнительными контурами управления, возникшими в результате введения в нее стабилизирующих связей R_{12} и R_{21}

Результаты моделирования переходных процессов МСАУ при введении в нее синтезированной стабилизирующей связи $R_{12} = 0.56$, представленные на рисунке 5, показали эффектив-

ность предложенного способа адаптации МСАУ к нерасчетным режимам работы за счет структурных преобразований регуляторов ее подсистем.

Дальнейшее повышение качества управления МСАУ может осуществляться при помощи коррекции регуляторов сепаратных подсистем, либо за счет изменения в устройстве управления программы управления координатами сепаратных подсистем.

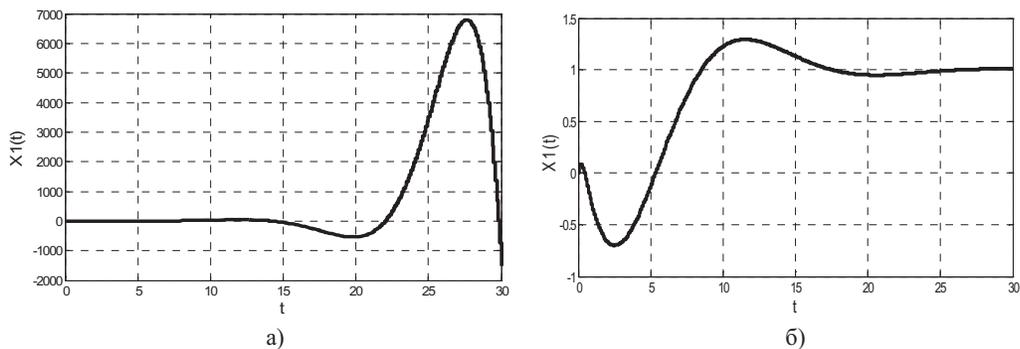


Рисунок 5 – Результаты моделирования переходных процессов координаты X_1 исследуемой МСАУ на нерасчетном режиме 2 до (а) и после (б) введения в нее стабилизирующей связи $R_{12} = 0.56$

Заключение

Предложенный алгоритм формирования структуры многосвязного регулятора МСАУ на основе анализа характеристических функций позволяет сформировать такие связи между подсистемами, которые обеспечивают требуемые запасы устойчивости на всех режимах функционирования МСАУ.

Данный метод по сравнению с известными методами [4], во-первых, имеет небольшую вычислительную сложность при большом числе варьируемых параметров системы, во-вторых, позволяет строить гарантированно устойчивые системы на всех режимах их функционирования, к тому же, метод весьма нагляден, прост для понимания и освоения с точки зрения прикладного инженерного применения.

Список литературы

- [1] Бабак С.Ф., Васильев В.И., Ильясов Б.Г. и др. Основы теории многосвязных систем автоматического управления летательными аппаратами: Учеб. пособие / Под ред. Красильщикова М.Н. - М.: Изд-во МАИ, 1995. - 288 с.
- [2] Васильев С.Н., Бадамшин Р.А., С.С. Валеев и др. Интеллектуальные системы управления и контроля газотурбинных двигателей / Под ред. Васильева С.Н. - М.: Машиностроение, 2008. - 550 с.
- [3] Ильясов Б.Г., Кабальнов Ю.С. Исследование устойчивости однотипных многосвязных систем автоматического управления с голономными связями между подсистемами // Автоматика и телемеханика. - 1993. - №8. - С. 81-90.
- [4] Морозовский В.Т. Многосвязные системы автоматического регулирования.. - М.: Энергия, 1970. - 288 с.
- [5] Петров Б.Н., Черкасов Б.А., Ильясов Б.Г., Куликов Г.Г. Частотный метод анализа и синтеза многомерных систем автоматического регулирования // Докл. АН СССР. - 1979. - Т. 247, № 2. - С. 304-307.
- [6] Рутковский В.Ю., Ильясов Б.Г., Кабальнов Ю.С. и др. Адаптивные системы управления газотурбинными двигателями летательных аппаратов / Под ред. Красильщикова М.Н. - М.: Изд-во МАИ, 1994. - 224 с.
- [7] Шевяков А.А., Мартянова Т.С., Рутковский В.Ю. и др. Оптимизация многомерных систем управления газотурбинных двигателей летательных аппаратов. / Под общей ред. Шевякова А.А., Мартяновой Т.С. - М.: Машиностроение, 1989. - 256 с.

КОМПЛЕКСНЫЙ АНАЛИЗ СОСТОЯНИЯ ТЕРРИТОРИАЛЬНЫХ СИСТЕМ НА ОСНОВЕ РЕШЕНИЯ ЗАДАЧИ ТИПОЛОГИЧЕСКОЙ КЛАССИФИКАЦИИ

В. Е. Гвоздев, Д. В. Блинова

Уфимский государственный авиационный технический университет
450000, Уфа, ул. К. Маркса, 12, Россия
kaf_apris@mail.ru
тел: +7 (347) 272-74-65

***Ключевые слова:** территориальная система, типологическая классификация, классификационная шкала, закон распределения непрерывной случайной величины, порядковая статистика, статистический индекс*

Abstract

This paper is devoted to complex estimation of territorial system condition based on statistical indexes. Calculation of indexes' values is carried out on a base of typological classifications territories by polytypic signs of condition. Procedure of determination intersecting classes bounds based on analysis of the serial statistics distribution laws is described.

Введение

Одной из задач информационной поддержки управления состоянием территориальных систем (ТС) на стратегическом и тактическом уровнях является комплексный анализ территориально-временной изменчивости ТС.

Территориальная система представляет собой совокупность природных и технических объектов, находящихся в тесной взаимосвязи. Управление ТС должно осуществляться с учетом единства природной, техногенной, социальной, экономической компонент. Эффективное управление состоянием ТС на разных уровнях невозможно без комплексного анализа данных, характеризующих состояние объекта управления.

Учитывая сложность и взаимосвязанность протекающих в территориальной системе природных и техногенных процессов, недостаточную изученность их механизмов, среди методов комплексного анализа территориально-временной изменчивости состояния ТС широкое распространение получили методы математической статистики в сочетании с геоинформационными технологиями.

В рамках комплексного анализа состояния ТС в качестве самостоятельной задачи следует выделить решение задачи типологической классификации с использованием методов математической статистики. Это обусловлено тем, что решение задачи классификации с использованием таких методов позволяет получить информационную основу для решения таких управленческих задач, как оценка эффективности ранее принятых мер по улучшению состояния территорий, комплексная оценка состояния территории, анализ динамики изменчивости состояния ТС, и др.

Исходя из вышеизложенного, целью работы является разработка методики комплексного анализа состояния территориальных систем на основе методов математической статистики.

Задачами исследования являются:

- разработка метода построения классификационных шкал, позволяющего усовершенствовать методику типологической классификации участков ТС;
- разработка методики анализа территориально-временной изменчивости состояния ТС и методики анализа тенденций её изменения, в том числе по комплексу признаков;

1 Типологическая классификация территорий как основа комплексного анализа состояния

Целью классификации является разработка процедуры отнесения участков ТС к тому или иному классу состояния в контексте анализируемых характеристик. В силу того, что территориальный объект является сложной системой, для оценки его состояния по какому-либо признаку, в зависимости от целей исследования, могут быть построены различные классификационные шкалы.

Существуют различные принципиальные подходы к построению классификационных шкал, в которых границы классов могут определяться:

- исходя из свойств статистических данных,
- на основе мнения эксперта,
- по результатам сплошного обследования территории,
- на основе космических снимков.

Наиболее доступным, экономически выгодным и объективным является подход построения классификационных шкал на основе статистических данных.

Из литературы известны различные методы построения классификационных шкал, некоторые из которых получили широкое распространение и были встроены в мощные программные пакеты. Одним из таких пакетов является ArcGIS X.X, в рамках которого возможно использование шести методов классификации, четыре из которых позволяют заранее назначать число классов.

Для сравнения результатов оценки состояния территориальной системы, получаемых при использовании одного и того же известного аппарата статистических индексов и различных методов определения границ классов, были использованы одни и те же данные о числе пожаров на территории Республики Башкортостан за 2001 - 2007 года (на 10000 населения), результаты приведены на рисунке 1. В первом случае (рисунок 1, а) был применен метод равных интервалов; во втором случае (рисунок 1, б) – метод квантилей.

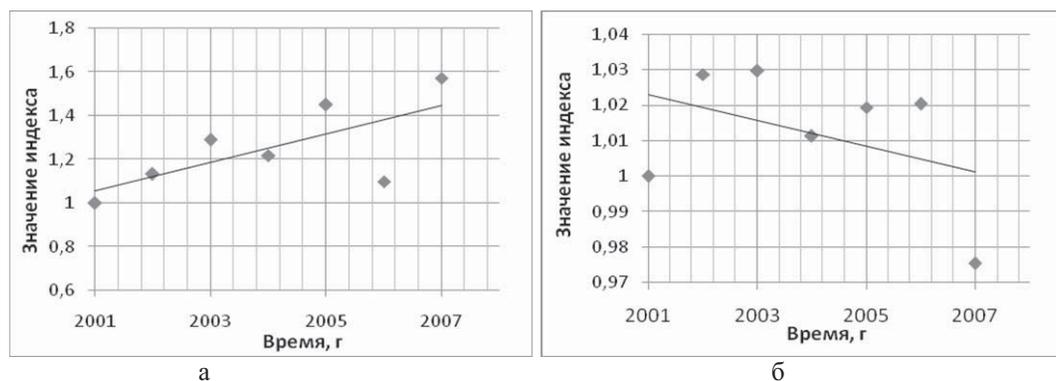


Рисунок 1 – Влияние подходов к формированию классификационных шкал на оценку состояния ТС:
а – классификация с использованием метода равных интервалов;
б – классификация с использованием метода квантилей

Из рисунка видно, что на основе одних и тех же данных, в зависимости от выбора классификационной шкалы, можно получить совершенно разные, взаимоисключающие оценки состояния ТС по рассматриваемому признаку. Таким образом, задача построения классификационных шкал является важной для комплексной оценки состояния ТС и может быть выделена в отдельную задачу.

2 Закон распределения случайной величины как основа построения классификационной шкалы

Одно из центральных мест в статистическом анализе территориально-временной изменчивости занимают задачи, связанные с оцениванием законов распределения случайных величин по выборочным данным. [1]. Это связано с тем, что вид закона распределения показателя состояния тесно связан с характером протекающих в территориальных системах процессов. Исследование вида закона распределения показателя состояния является основой изучения пространственных структур, исследования взаимосвязей между оказываемым воздействием и поведением системы.

Известно, что для случайной величины, измеряемой количественно, знание закона распределения позволяет определять все остальные статистические величины. Именно поэтому при построении классификационной шкалы, ориентированной на эффективное использование выборочных данных, в качестве основы целесообразно использовать закон распределения случайной величины.

В ранее проведенных исследованиях были разработаны унифицированные модели уни-модальных законов распределения случайных величин с левосторонним ограничением. [2, 3, 4] Построение этих моделей позволило формализовать процедуру оценивания закона распределения непрерывной случайной величины по выборочным данным. Так же было установлено, что учет границ интервала возможных значений случайной величины позволяет строить оценки законов распределения на основе лишь двух первых начальных моментов.

В ходе ранее проведенных исследований установлено, что учет левой границы интервала возможных значений случайной величины приводит к возникновению функциональной зависимости между моментами высших порядков и первым начальным моментом (при фиксированном значении среднеквадратического отклонения). Зависимости, иллюстрирующие полученные результаты, приведены на рисунке 2.

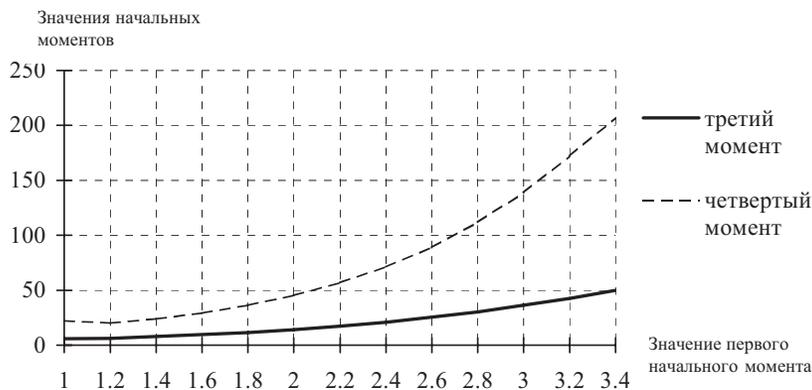


Рисунок 2 – Взаимосвязь значений начальных моментов

Полученные результаты, по мнению авторов, являются уточнением известного положения теории математической статистики, заключающегося в том, что закон распределения случайной величины можно описать ограниченным числом начальных моментов.

Возможность построения оценок законов распределения случайной величины по значениям двух первых моментов является важным с точки зрения обработки малых по объему и низких по точности выборочных данных. Это обусловлено тем, что погрешность оценивания значений моментов по выборочным данным возрастает с увеличением порядка момента.

То обстоятельство, что для описания унимодальных законов распределения достаточно учитывать лишь два первых начальных момента, позволило установить монотонный характер изменения формы интегральных функций распределения в зависимости от коэффициента вариации (рисунок 3).

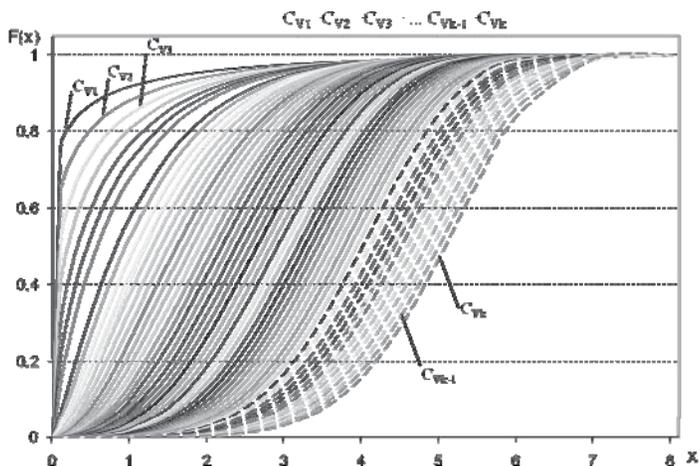


Рисунок 3 – Характер изменения формы интегральных функций распределения при фиксированном значении среднеквадратического отклонения

В ходе проведенных исследований были разработаны решающие правила, определяющие предпочтительность использования определенных унифицированных моделей [5].

3 Построения классификационной шкалы с пересекающимися классами состояния

Использование пересекающихся классов состояния позволяет избежать ситуации отнесения объектов с близкими значениями характеристического признака в разные классы состояния и получить промежуточные интервалы группирования, в которых объекты не обладают ярко выраженными свойствами принадлежности к какому-либо классу.

Предложенный подход к оцениванию законов распределения непрерывных случайных величин позволил разработать схему построения классификационных шкал с использованием аппарата порядковых статистик и с учетом доверительной вероятности попадания объекта в определенный класс состояния β . Графическая иллюстрация схемы приведена на рисунке 4.

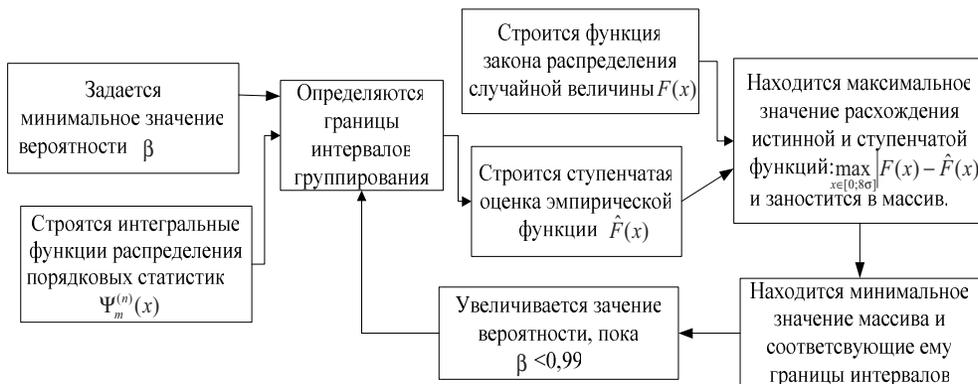
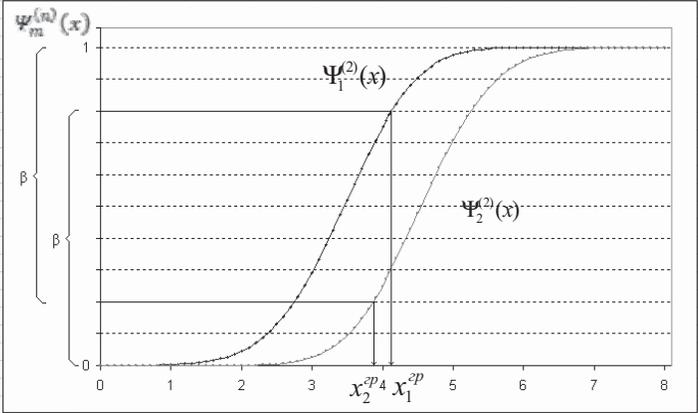


Рисунок 4 – Схема определения границ интервалов группирования

На рисунке 5, в качестве примера, проиллюстрировано определение границ интервалов группирования в зависимости от значения вероятности β для случая двух пересекающихся классов, которые в этом случае выглядят следующим образом: $[0 \div x_2^{zp}]; (x_2^{zp}; x_1^{zp}]; (x_1^{zp}; 8\sigma]$, где x_1^{zp} , x_2^{zp} – границы первого и второго класса соответственно, σ – среднеквадратическое отклонение, рассчитанное по выборочным данным. Объекты, попавшие в данном примере во второй интервал группирования, не обладают выраженными свойствами принадлежности к одному из классов состояний.



$\Psi_m^{(n)}(x)$ – интегральные функции распределения порядковых статистик

Правила формирования границ классов:

$$\Psi_1^{(2)}(x_1^{gp}) - \Psi_1^{(2)}(0) = \beta$$

$$\Psi_2^{(2)}(8\sigma) - \Psi_2^{(2)}(x_2^{gp}) = \beta$$

Рисунок 5 – Формирование классификационной шкалы для случая двух пересекающихся классов состояний

Законы распределения порядковых статистик рассматриваются в данной работе как законы распределения случайных величин внутри классов состояний. Оценивание интегральных законов распределения порядковых статистик осуществляется на основе известного соотношения:

$$(1) \quad \varphi_m^{(n)}(x) = n \cdot C_{n-1}^{m-1} [F(x)]^{m-1} [1 - F(x)]^{n-m} f(x),$$

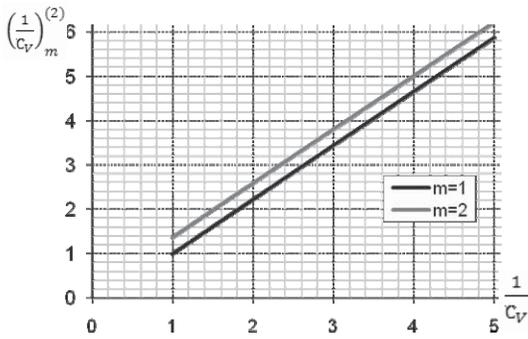
где $f(x)$ – плотность распределения (дифференциальная функция распределения) случайной величины, соответствующая $F(x)$;

n – число классов группирования;

m – рассматриваемый класс ($m = \overline{1, n}$).

Проведенные исследования позволили разработать решающее правило выбора вероятности β в зависимости от числа классов состояний и объема выборочных данных.

Для упрощения процедуры построения оценок законов распределения порядковых статистик в работе были выявлены зависимости между характеристиками формы и масштаба закона распределения случайной величины $F(x)$ и характеристиками формы и масштаба законов распределения внутри классов состояний $\varphi_m^{(n)}(x)$. На рисунках 6 и 7, в качестве примера, приведена часть полученных результатов, соответствующих случаю двух классов состояний.



C_V – коэффициент вариации;

$\left(\frac{1}{C_V}\right)_m^{(n)}$ – параметр, характеризующий форму закона распределения случайной величины внутри m -го класса состояния при общем числе классов n .

Рисунок 6 – Зависимость характеристик формы законов распределения случайных величин внутри двух классов от характеристики формы распределения случайной величины

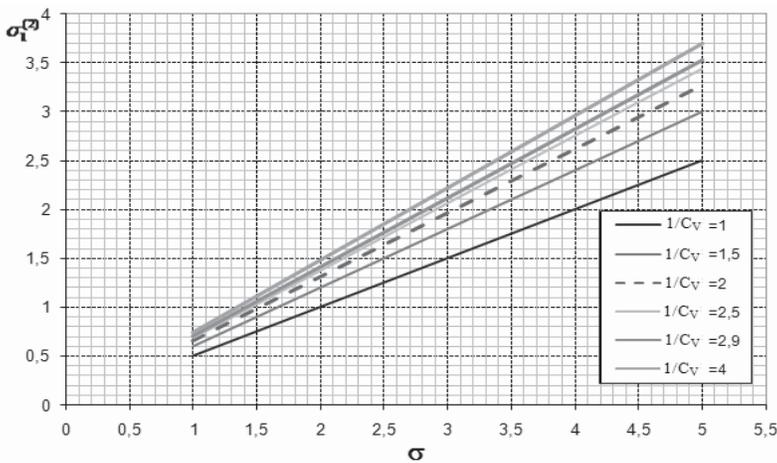


Рисунок 7 – Зависимость характеристики масштаба закона распределения случайной величины внутри первого класса от характеристики масштаба распределения случайной величины для различных характеристик формы закона распределения случайной величины

Полученные результаты позволяют разработать достаточно простую схему построения законов распределения случайных величин внутри классов (рисунок 8), что позволяет уменьшить количество сложных математических вычислений, тем самым уменьшив погрешность при определении границ классов.

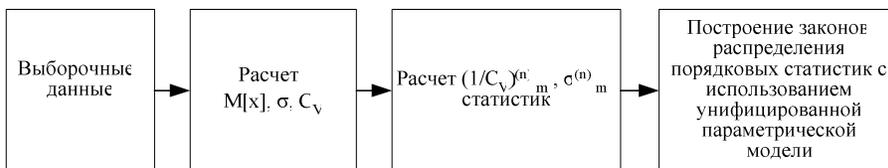


Рисунок 8 – Схема построения законов распределения порядковых статистик на основе унифицированной параметрической модели

В результате исследований разработана простая схема, позволяющая полностью формализовать оценивание законов распределения случайных величин внутри классов состояний.

Разработанная схема позволяет упростить получение результата без потери точности за счет исключения операции численного интегрирования.

4 Комплексный анализ состояния территориальных систем на основе решения задачи типологической классификации

В настоящей работе для комплексной оценки состояния ТС в качестве подхода предлагается использование описанных в литературе статистических индексов, конструируемых по результатам решения задачи типологической классификации.

Анализ состояния ТС должен содержать в себе ответы на два основных вопроса: стабильно ли состояние рассматриваемой ТС и каким образом изменяется состояние ТС. Для ответа на эти вопросы в данной работе предлагается использование двух видов индексов.

В основе конструирования индексов, характеризующих стабильность состояния ТС, лежат результаты типологической классификации ТС:

$$(2) \quad Q_k^{(g)} = \sum_{i=1}^{n_k} \frac{S_i^{(g,k)}}{S} \ln \frac{S_i^{(g,k)}}{S},$$

где $S_i^{(g,k)}$ – суммарная площадь территорий, отнесенных к i -му классу состояния в g -м временном срезе;

S – общая площадь исследуемой территории;

n_k – число классов состояния k -го признака.

На рисунке 9 приведена часть результатов классификации территории Республики Башкортостан по числу возникновения пожаров на 10 тысяч населения за 2001 – 2007 гг.

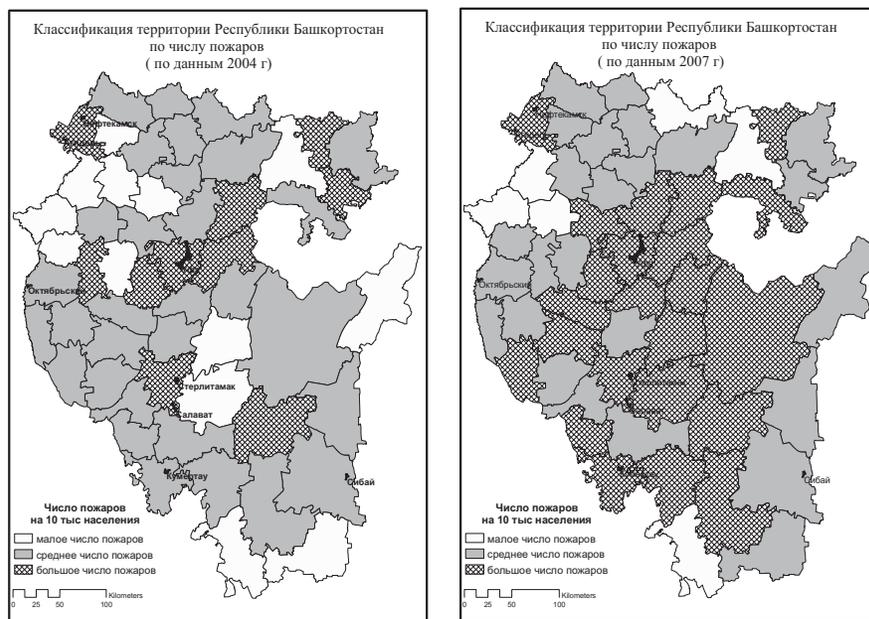


Рисунок 9 – Классификация муниципальных образований Республики Башкортостан по числу возникновения пожаров на 10 тысяч населения

Для оценки направления изменения состояния ТС предлагается использовать индекс, характеризующий среднее значение рассматриваемого признака, индекс рассчитывается на основе соотношения:

$$(3) \quad I_k^{(g)} = \frac{\sum_{i=1}^N \frac{\Pi_i^{(g,k)}}{S_i^{(g,k)}}}{N},$$

где $S_i^{(g,k)}$ – площадь i -го участка территории, входящего в ТС в g -м временном срезе;

$\Pi_i^{(g,k)}$ – значение рассматриваемого k -го признака i -го участка территории в g -м временном срезе;

N – общее число участков территорий, входящих в ТС.

Одновременное использование совокупности индексов позволяет охарактеризовать стабильность изменчивости ТС и определить направление изменения состояния (улучшение или ухудшение). По результатам классификации, представленным выше, были рассчитаны значения статистических индексов (4), результаты приведены на рисунке 10.

$$(4) \quad Q_k^{(g,l)} = \frac{Q_k^{(g)}}{Q_k^{(l)}}; \quad I_k^{(g,l)} = \frac{I_k^{(g)}}{I_k^{(l)}}$$

На рисунке 11, в качестве примера, приведен график изменения значений индексов для случая, когда в качестве характеристического признака рассматривалось число погибших на 10 тысяч населения.

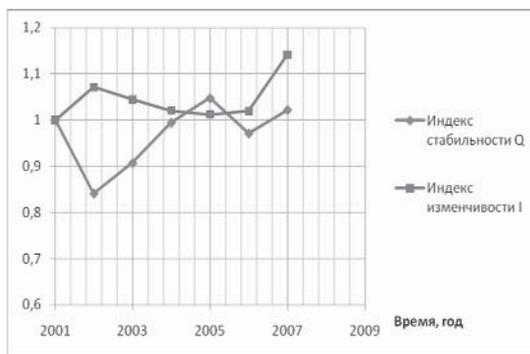


Рисунок 10 – Изменение значений индексов для признака «Число пожаров на 10 тысяч населения»

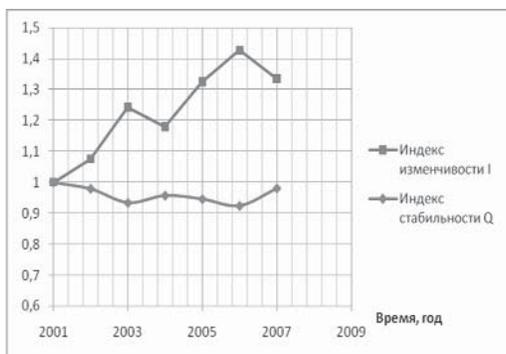


Рисунок 11 – Изменение значений индексов для признака «Число погибших на 10 тысяч населения»

Использование индексов позволяет конструировать интегральные оценки, характеризующие динамику изменения состояния ТС по комплексу признаков:

$$(5) \quad W^{(l)} = \frac{\sum_{k=1}^p a_k \cdot Q_k^{(l)}}{\sum_{k=1}^p a_k},$$

где p – число учитываемых признаков;

a_k – весовой коэффициент, учитывающий значимость k -го признака;

$Q_k^{(l)}$ – значение индекса k -го параметра в l -х условиях исследования.

На рисунке 12 приведен график изменения интегральных индексов, характеризующих комплексное изменение состояния территории Республики Башкортостан по двум характеристическим признакам. Признакам присваивались одинаковые весовые коэффициенты.

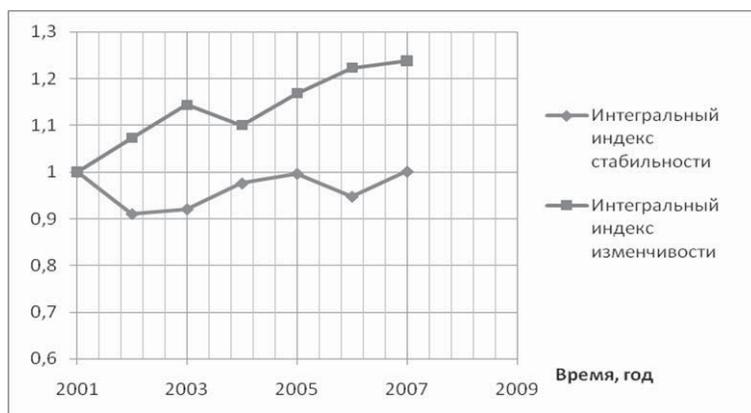


Рисунок 12 – Комплексная оценка изменения состояния территориальной системы

Проведенный таким образом комплексный анализ состояния рассматриваемой территории позволяет отобразить изменение состояния рассматриваемой ТС по ряду характеристических признаков, что делает возможным получить комплексную оценку территориально-временной изменчивости состояния рассматриваемой территории. Результаты анализа являются основой краткосрочного прогнозирования состояния ТС по комплексу характеристик состояния.

Заключение

Разработан метод комплексного анализа территориальных систем на основе разнотипных данных. В основу метода положено системное сочетание процедуры типологической классификации для случая пересекающихся классов состояния и аппарата статистических индексов, что позволяет анализировать состояние рассматриваемой территориальной системы по совокупности гетерогенных характеристик (разных по природе и размерности).

Формализация процедуры построения классификационных шкал и расчета статистических индексов делает возможным реализацию разработанного метода в виде модуля в составе геоинформационной системы. Программная реализация метода позволяет использование его специалистами, не имеющими глубоких знаний в вопросах статистической обработки данных.

Работа поддержана грантом РФФИ 10-08-00359-а.

Список литературы

- [1] А. М. Берлянт Картографический метод исследования. – М.: МГУ, 1988. – 252 с.
- [2] В. Е. Гвоздев, М. Б. Гузаиров, Б. Г. Ильясов, А. Е. Колоденкова Статистическое исследование территориальных систем: монография – М.: Машиностроение, 2008. – 187 с.
- [3] В. Е. Гвоздев, Д. В. Семененко (Блинова) Информационная поддержка анализа состояния территориальных систем по разнотипным признакам // Вестник УГАТУ. Серия «Управление, вычислительная техника и информатика». 2009. Т. 12, № 1 (30). С. 9–15.
- [4] В. Е. Гвоздев, Д. В. Семененко (Блинова) Анализ состояния территориальных систем на основе решения задачи классификации // Мехатроника, автоматизация, управление: 2-я Российск. мультиконф. по пробл. управления : сб. тр. СПб. : ГНЦ РФ ЦНИИ «Электроприбор», 2008. С. 166–169.
- [5] В. Е. Гвоздев, Д. В. Семененко (Блинова) Определение условий применимости унифицированных параметрических моделей закона распределения непрерывной случайной величины // Проблемы машиноведения, процессов управления и критических технологий: Сборник научных трудов. – Уфа: Гилем, 2008г., С. 215 – 219.

МЕТОДЫ МЕНЕДЖМЕНТА КАЧЕСТВА И БЕРЕЖЛИВОГО ПРОИЗВОДСТВА В СИСТЕМЕ УПРАВЛЕНИЯ ПЕРЕВОЗОЧНЫМ ПРОЦЕССОМ НА ЖЕЛЕЗНОДОРОЖНОМ ТРАНСПОРТЕ

О.В. Скачков¹, И.А. Ларин²

¹Поволжское отделение Российской Инженерной академии
443001, Самара, Студенческий пер., 3а, Россия
skachkov@poria.ru

тел: +7 (846) 242-04-39, +7 (846) 242-32-13

²Дирекция управления движением
Куйбышевской железной дороги – филиала ОАО «РЖД»
443030, Самара, Комсомольская площадь, 2/3, Россия
kbshdz1@gmail.com

Ключевые слова: менеджмент качеств, железнодорожный транспорт, бережливое производство, управление движением поездов.

Abstract

In this paper we address the issue of cost reduction and improvement of controllability in the train management. Methods of quality management and lean production were used in project Corporate Integrated management system of “Kuibyshev railway” – branch of “Russian Railways”, in the Traffic Operation Department. The obtained results show high potential and economic benefits of these methods implementation.

Введение

Сеть железных дорог ОАО «РЖД» представляет собой протянувшийся на десятки тысяч километров транспортный конвейер. Сложность и комплексность управления перевозочным процессом, большое число задействованных служб предъявляет высокие требования к точности планов, оперативности реагирования и правильности решений системы управления движением. Поэтому для Дирекции управления движением ОАО «РЖД» всегда актуальным является поиск эффективных методик выявления и реализации имеющихся резервов для повышения управляемости, снижения эксплуатационных расходов, обеспечения надежности и безопасности перевозок.

В статье описаны подходы, методики и инструменты, нашедшие применение в проекте внедрения системы менеджмента качества (далее – СМК) в дирекции управления движением Куйбышевской железной дороги.

1 Характеристика объекта управления и целевые задачи проекта

Дирекция управления движением является вертикально-интегрированной бизнес-единицей холдинга ОАО «РЖД» с тремя уровнями системы управления: центральным с аппаратом управления в Москве; региональным уровнем, включающим 17 региональных дирекций; линейным уровнем, состоящим из примерно 7000 станций.

Проект внедрения СМК в региональной дирекции Куйбышевской железной дороги ведется с 2007 года и затрагивает региональный уровень (аппарат управления, дорожный центр управления перевозками), около 15 пилотных подразделений линейного уровня и, частично, смежные хозяйства Куйбышевской дороги. Консультантом проекта является Поволжское отделение Российской Инженерной академии.

Целью внедрения СМК в соответствии с [1] является достижение наилучших показателей качества, себестоимости и времени за счет внедрения методов менеджмента качества и бережливого производства. На рисунке 1 представлен дом качества региональной дирекции с отмеченной модулей, по которым ведется работа.

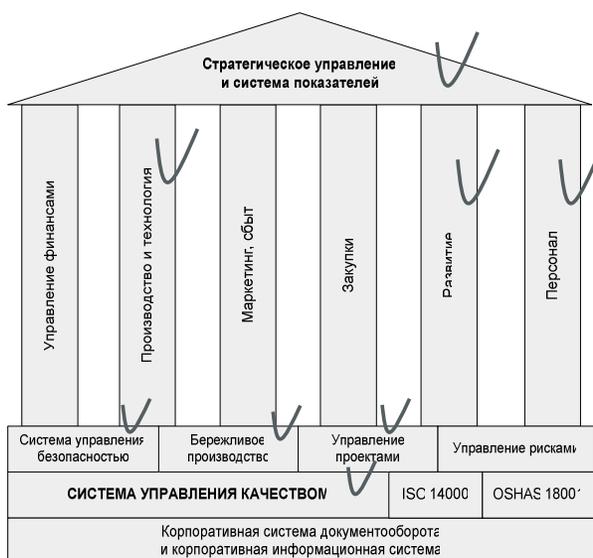


Рисунок 1 – Интегрированная система менеджмента Дирекции управления движением

Задачами проекта внедрения СМК является:

- формирование системы целевых показателей процессов, планирование и контроль достижения целей;
- регламентация процессов взаимодействия с другими подразделениями Куйбышевской дороги, регламентация внутренних процессов дирекции движения (связи между аппаратом управления, дорожным центром управления перевозками, станциями), устранение проблем на стыках;
- непрерывный анализ и устранение потерь в производственных процессах с использованием методов и инструментов менеджмента качества и бережливого производства, повышение уровня качества и безопасности производственных процессов;
- повышение оперативности, достоверности и общей эффективности системы отчетности;
- планирование развития дирекции;
- определение квалификационных требований к персоналу, управление обучением персонала, развитие системы мотивации персонала.

2 Система постановки, планирования и контроля достижения целевых показателей

Согласно стандартам ГОСТ Р ИСО серии 9000 [2,3] для успешного руководства организацией необходимо управлять ею систематически и прозрачным способом. Управление организацией включает менеджмент процессов, который заключается в том, что владелец процесса непрерывно или с установленной периодичностью контролирует ход процесса по установленным *целевым показателям* и принимает управленческие решения в случаях отклонения целевых показателей от значений, установленных для нормального хода процесса [4].

Система показателей Дирекции управления движением приведена на рисунке 2.

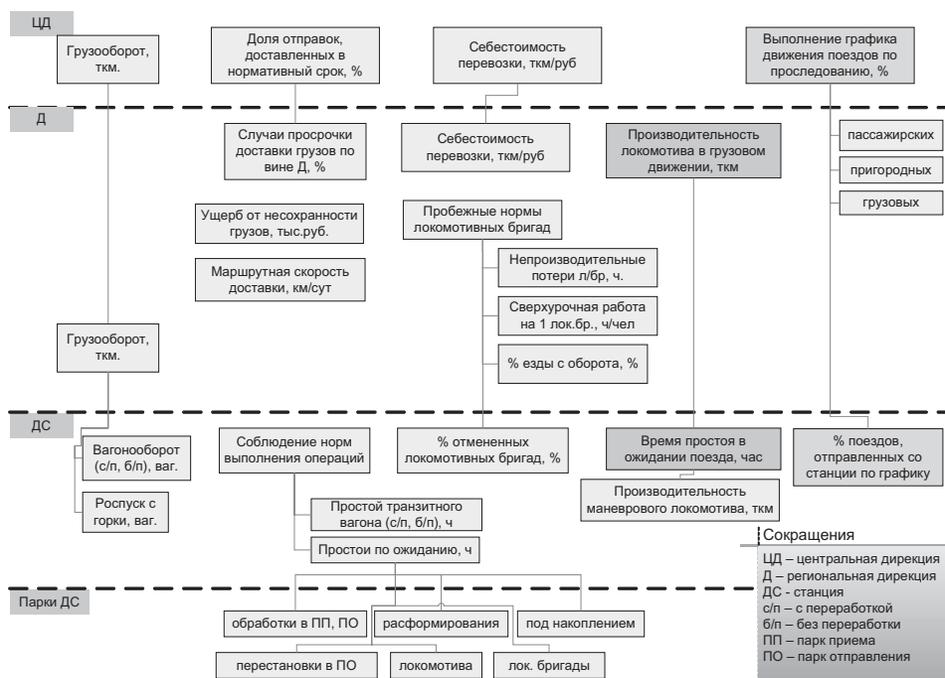


Рисунок 2 – Дерево показателей по качеству и безопасности Дирекции управления движением

Менеджмент процессов осуществляется по циклу PDCA [5].

Первоначально производится постановка целей (утверждаются числовые значения показателей) на основании заданий руководства, наличных ресурсов, прогнозов возможностей улучшения. По каждой цели назначается ответственное лицо из числа высшего руководства подразделения.

При каждом высшем руководителе создаются межфункциональные рабочие группы из специалистов и руководителей среднего звена, задачей которых является анализ деятельности дирекции по своему направлению, грамотное планирование ресурсов при разработке планов мероприятий по достижению целей, оценка возможных рисков и поиск инновационных путей развития.

В ходе реализации плана мероприятий контролируется соответствие показателей запланированным значениям, выявляются все имеющиеся отклонения, причины появления таких отклонений и проводятся корректирующие действия таким образом, чтобы устранить выявленные причины. После реализации корректирующих мероприятий снова проверяется соответствие показателей качества их запланированным значениям. Цикл повторяется до тех пор, пока результат не совпадет с планом.

Для примера рассмотрим планирование и осуществление мероприятий по улучшению процессов оперативного планирование и управления движением поездов.

3 Совершенствование процессов оперативного планирование и управления движением поездов

В перевозочном процессе сортировочные станции можно рассматривать как заводы по производству поездов, а полигон дороги – как огромный транспортный конвейер. Аналогично заводскому конвейеру, множество звеньев (подразделений разных дирекций) должны работать точно в срок, чтобы обеспечить бесперебойное движение поездов. Основой для этого

служит график движения поездов¹. Первым этапом на пути к движению грузовых поездов по нормативному графику движения является планирование отправления поездов со станций по «точкам» графика движения².

Главными показателями процесса на региональном уровне являются:

- Выполнение графика движения поездов по проследованию, %
- Участковая скорость (км/час)
- Производительность локомотива (тыс.ткм брутто)

Для улучшения процесса с января 2009 года на сортировочных и узловых станциях Куйбышевской дороги (Кинель, Дема, Октябрьск, Пенза-3, Сызрань-1, Рузаевка, Ульяновск, Бугульма) внедряется новая технология отправления грузовых поездов на «точки» графика движения.

В технологии применяются следующие принципы бережливого производства (таблица 1).

Таблица 1 – Инструменты бережливого производства в технологии отправления грузовых поездов на «точки» графика движения

Принцип бережливого производства	Примеры по управлению движением
Выравнивание загрузки, исключение потерь и запасов	- Регулирование загрузки перегонов и станций, планирование движения по «точкам» графика движения - Содержание парка локомотивов и явочного контингента бригад исходя из 6-ти часового прогноза потребности станции - Подвязка локомотивов с учетом станции назначения и пробега до ТО-2
Выстраивание потока	- Планирование безостановочного прохождения поезда по участкам полигона дороги (ускоренные поезда) - Планирование использования локомотивов и бригад с оборота
Организация вытягивания	- Подготовка (подвязка) локомотивов и бригад под утвержденные «точки» отправления - Выставка поезда в парк отправления только под согласованную «точку»
Автоматизация Визуальный контроль Предотвращение ошибок	Автоматизация диспетчерского планирования и обмена планами между рабочими местами ДЦУП, станций, депо. Применение наглядных средств оповещения причастных об угрозах срыва плана движения поездов (типа канбан, андон). Цветовая индикация статуса планирования поезда (утвержденного, требующего согласования с поездным диспетчером, критического) Приведение к единому стандарту используемые информационные системы. Протоколирование действий диспетчеров при планировании и управлении, автоматизация отчетности и анализа
Стандартизация	- Регламенты взаимодействия диспетчеров при планировании движения поездов - Шаблоны действий персонала в нестандартных ситуациях

Методы проектирования непрерывных потоков, выравнивания загрузки производственной линии используются при планировании движения ускоренных поездов по полигону дороги – без задержек для пропуска пассажирских и пригородных поездов, с остановками на станциях для выполнения минимально необходимых технологических операций по смене бригад и осмотра составов.

¹ http://ru.wikipedia.org/wiki/График_движения_поездов

² «Точкой» графика называется нормативное время отправления некоего абстрактного грузового поезда со станции. Для конкретной станции, в зависимости от объемов движения, в графике за сутки содержится несколько десятков «точек».

В отличие от пассажирских и пригородных поездов, грузовые поезда не ходят по жесткому и заранее утвержденному графику движения. Задача поездного диспетчера заключается в том, чтобы спланировать и отправить реальный поезд на наиболее подходящую «точку».

На станционном уровне при планировании отправления поезда на «точку» используются методы вытягивающего производства и системы «канбан» [6]. Целью является четкий согласованный план действий всех подразделений в железнодорожном узле - станции, локомотивного и вагонного депо, дома отдыха бригад и других - на период 5-6 часов (рисунок 3).

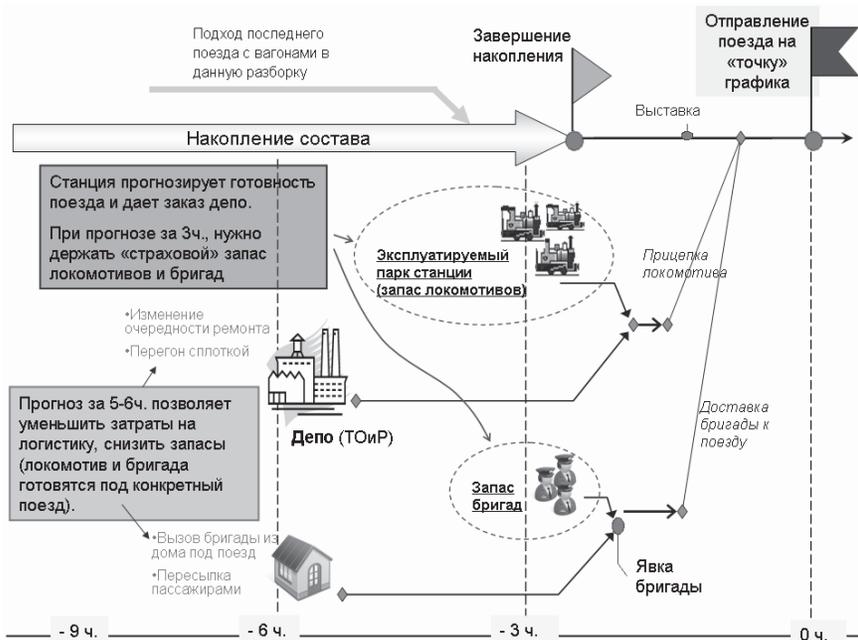


Рисунок 3 – Взаимосвязь операций планирования в единой технологической цепочке

Повышение глубины и точности планов оперативной работы достигается за счет регламентации взаимодействия диспетчерского аппарата ДЦУП, станций, локомотивных депо, согласования временных интервалов выполнения технологических операций, стандартизации порядка обмена информацией и используемых форм данных. Также вносятся дополнения в должностные инструкции и показатели премирования. Для исполнителей разрабатываются технологические карты рабочих мест, регламентирующие порядок выполнения технологических операций, нормы времени на их выполнение, соблюдение требований охраны труда и безопасное производство работ.

Внедрение данной технологии по итогам 2009 года в сравнении с 2008 годом позволило:

- улучшить участковую скорость на 7,26% за счет уменьшения количества задержек поездов на перегонах (по причинам обгона пассажирским и пригородными поездами, остановок под скрещение);
- снизить количество и время задержек грузовых поездов у запрещающих сигналов на 21 177,3 часов за счёт предупреждения пачкообразного подхода поездов, организации пропуска поездов на проход, перераспределения объёма поездной работы между сортировочными станциями;
- увеличить производительность локомотива на 2,29% и среднесуточный пробег локомотива на 9,8% за счет содержания минимального запаса локомотивов на станции (благодаря заблаговременному прогнозу готовности поезда), подвязки локомотивов с учетом их технического состояния, и тем самым предупреждения неплановой отцепки локомотива в пути следования, снижения простоя локомотивов на станциях;

- а также обеспечить информационную эффективность в сфере оперативного управления движением поездов, взаимопонимание, единство представления и восприятия информации для всех работников диспетчерского аппарата.

Технология отправления грузовых поездов на «точки» графика движения в настоящее время является передовой на сети железных дорог, что подтверждено протокольными решениями высшего руководства ОАО «РЖД».

4 Мониторинг, анализ и корректирующие мероприятия

Первичный мониторинг производится в течение смены на станциях и в ДЦУП. В нем собираются данные о причинах допущенных срывов в движении поездов и вина конкретного сотрудника (рисунок 4). Анализ форм позволяет определить причину срыва точки отправления, определить вину конкретного сотрудника и принять к нему адресные действия.

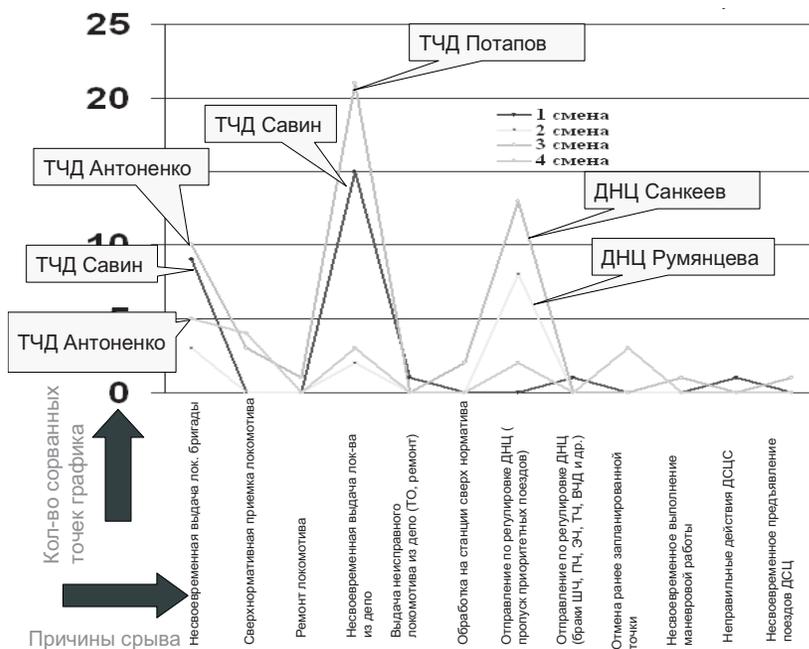


Рисунок 4 – Анализ причин срыва планов отправления поездов в разрезе по сменам (станция Дема)

Мониторинг выполнения показателей производится еженедельно при помощи диаграммы причинно-следственной связи (диаграммы Исикавы), при помощи которой оценивается изменение каждого влияющего фактора за последние 4 недели и используется для принятия оперативных корректирующих решений.

Параллельно отслеживаются тренды показателей в динамике с начала года (рисунок 5). После внедрения мероприятий должно наблюдаться улучшение показателя. При наличии всплесков и падений проводится анализ причин, при необходимости корректируются внедряемые мероприятия. Эта форма позволяет оценить эффективность внедряемых мероприятий, то есть оценить усилия персонала по улучшению процесса, уйти от практики написания формальных мероприятий (что затруднительно при существовавшем ранее практике сравнения значения показателя с аналогичным периодом предыдущего года).

Динамика среднесуточной производительности локомотива (понеделно)

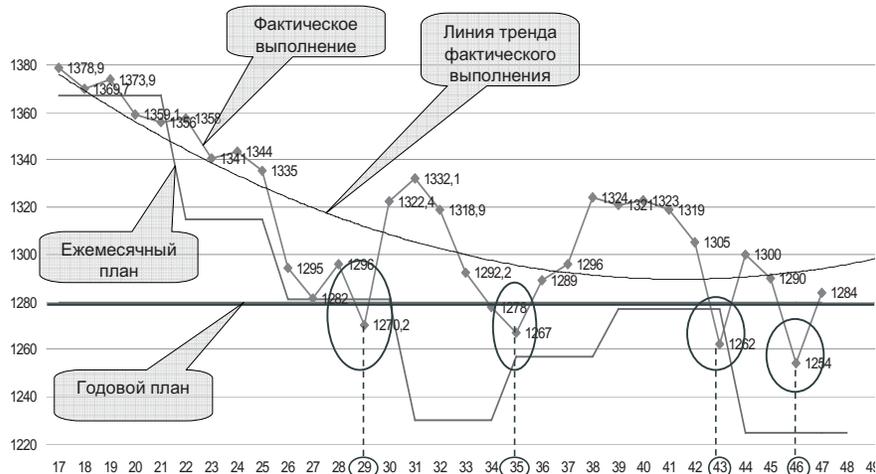


Рисунок 5 – Динамика выполнения среднесуточной производительности локомотива

По итогам года подводится общий результат работы, анализируется какие цели были достигнуты/не достигнуты, какие мероприятия выполнены/не выполнены и почему (по причинам, зависящим от самой службы, по внешним причинам). Данные итогового анализа используются при планировании на следующий год.

Отдельно необходимо отметить, что в существующей прежде системе мониторинга на руководителя предприятия была возложена функция анализа всех данных, тогда как руководители среднего звена (начальники отделов) зачастую просто собирали данные от исполнителей в единый отчет. Изменение системы позволяет вовлечь исполнителей и руководителей среднего звена в процесс анализа и самоконтроля (за счет типовых форм мониторинга у исполнителей, за счет типовых форм анализа, протоколов у руководителей среднего звена), и снизить нагрузку высших руководителей.

Заключение

В статье описана система менеджмента качества в дирекции управления движением, порядок постановки, планирования и контроля достижения целевых показателей. На примерах проиллюстрировано, как производится улучшение процессов оперативного планирования и управления движением поездов с использованием методов бережливого производства.

Список литературы

- [1] Функциональная стратегия управления качеством в ОАО «РЖД» (утв. распоряжением ОАО «РЖД» № 46р от 15.01.2007г.)
- [2] ГОСТ Р ИСО 9000-2008 «Системы менеджмента качества. Основные положения и словарь»
- [3] ГОСТ Р ИСО 9001-2008 «Системы менеджмента качества. Требования»
- [4] Репин В.В., Елиферов В.Г. Процессный подход к управлению: Моделирование бизнес-процессов. – М.: РИА «Стандарты и качество», 2004. – 534с.
- [5] Алексеев В.В. Использование цикла PDCA для улучшения процессов// Методы менеджмента качества. – 2003. – №9. – С. 60.
- [6] Голдсби Т., Мартиченко Р. Бережливое производство и 6 сигм в логистике. – М.: «Гревцов Паблшер», 2009. – 416с.

СИСТЕМНЫЙ АНАЛИЗ ВОСПРИЯТИЯ НАГЛЯДНЫХ АЛГЕБРАИЧЕСКИХ СТРУКТУР

А.М. Кистанов

Самарский Государственный технический университет
443071, Самара, ул. Молодогвардейская, 244, Россия
plast@solid.ru

Ключевые слова: структура, наглядные структуры, отношения наглядности, красота, тесты, субъект испытаний, объект испытаний, связный помеченный граф

Abstract

The system analysis of perception any visual algebraic structures are considered in this article. In practice the offered methods can be used for the education and personal testing.

Введение

В практической деятельности при эксплуатации информационных систем различного уровня сложности приходится сталкиваться с их недостаточной структурированностью, отсутствием эффективной возможности взаимодействия с другими системами и внутренними механизмами развития. Как правило, это связано с недостаточной подготовленностью или отсутствием системного подхода у проектировщиков. Зачастую присутствует исключительно инженерный подход к решению задач или так называемый «ползучий эмпиризм». Следствием является сложность в эксплуатации, короткий жизненный цикл проектируемых или используемых систем и, в конечном счете, неэффективность инвестиций.

В литературе известны многочисленные исследования в области квалиметрии профессиональных знаний. Отличительной чертой данной работы является разработка методики оценки структурно-ориентированной деятельности субъекта посредством анализа восприятия субъектом наглядных алгебраических структур.

Выбор критерия оценки

Критерий оценки – ключевая категория. Неправильный критерий даст неправильные результаты. Он должен быть понятен всем субъектам тестирования. Для критерия необходимо предложить меры и шкалы для получения объективной оценки.

Анализируя различные подходы в этой области, автор пришел к выводу, что таким критерием может служить понятие красоты. Это, на первый взгляд, неожиданное решение имеет под собой весомое основание. С одной стороны, понятие красоты сугубо индивидуальное, личностное, некая «категориальная интуиция» [1], свойство человеческого мозга. С другой стороны, многие ученые рассматривали красоту как непреходящий фактор в своих работах. «Едва ли кто-нибудь из нематематиков в состоянии освоиться с мыслью, что цифры могут представлять собой культурную или эстетическую ценность или иметь какое-нибудь отношение к таким понятиям, как красота, сила, вдохновение. Я решительно протестую против этого костного представления о математике» [2], - говорил Н. Винер в своей работе «Я математик».

Выбор объекта оценки

Наглядное представление различных математических объектов [3] позволяет эффективно исследовать эти объекты, получать новые знания о них, находить требуемые закономерности.

Порой вне наглядных отношений найти такие закономерности не представляется возможным. Получение новых знаний является критерием, по которому оценивается качество наглядных отношений, определяющих структуру математических объектов.

Наиболее четко характеристика эстетической привлекательности математического объекта дана Г. Биркгофом: $M=O/C$, где M — мера красоты объекта, O — мера порядка, а C — мера усилий, затрачиваемых для понимания сущности объекта [4].

Для целей работы объектами оценки были выбраны различные алгебраические структуры, визуально представленные с использованием отношений наглядности [3]. Эти наглядные структуры обладают зеркальной, переносной, поворотной, орнаментальной симметрией, антисимметрией, в них присутствуют соотношение общего и частного, самоподобие, подструктуры и т.д. Критерием правильности выбранных наглядных отношений являются полученные новые знания о структурах и их применение на практике. Поэтому использование данных структур в качестве эталонных объектов выглядит достаточно справедливо.

Кроме того существовало опасение, что предложенные в [3] наглядные методы исключительно индивидуальны. Поэтому дополнительная задача состояла в том, чтобы выяснить, насколько выбранные наглядные отношения отражают понимание структуры у других субъектов.

Выбор субъекта оценки

Были выбраны как индивидуальные, так и коллективные субъекты оценки. Все тесты проводились анонимно. Тестовые задания предлагались студенческим группам, сотрудникам организаций. Разные группы имели различную подготовку, различную предметную область деятельности. Общим для всех субъектов было то, что никто из них не изучал высшую алгебру, не знаком с теорией решеток, поэтому влияние математического образования было минимальным. С другой стороны, это внесло некоторые сложности, поскольку формулировка критерия должна была быть понятна всем, независимо от образования. Поэтому текст тестового задания был ближе к бытовому, чем к научному. Задание №1 звучало так: «Соедините вершины прямыми ребрами, чтобы было красиво». Ключевая фраза: «чтобы было красиво» объяснялась так: «Исходя из вашего внутреннего понимания».

Тестовые задания и результаты

Использовались несколько заданий, которые предлагались трем различным группам, которые условно назовем {студенты 1 гр., студенты 2 гр., специалисты}. Для анализа использовались решетки $L = (L(m,n), \subset, \mu)$ слов m буквенного алфавита длиной не более n символов. Здесь $L(m,n)$ - множество слов, \subset - отношение включения для L , $\mu = (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5)$ - отношения наглядности [3] для L . В дальнейшем, для краткости запись $L(m,n)$ будет означать $L = (L(m,n), \subset, \mu)$. Пусть $P = \{L(4,2), L(4,3), L(4,4)\}$ - множество решеток, предъявленных для тестовых заданий. Пусть $Q = \{\text{студенты 1 гр., студенты 2 гр., специалисты (3 гр.)}\}$ – множество групп субъектов. Тогда множество всех испытаний состоит из элементов декартова произведения $(p,q) \in P \times Q : p \in P, q \in Q$. В случае $L(m,n)$ мы ограничились испытаниями: $(L(4,3), 2 \text{ гр.})$; $(L(4,2), 1 \text{ гр.})$; $(L(4,3), 3 \text{ гр.})$; $(L(4,4), 1 \text{ гр.})$; $(L(4,4), 3 \text{ гр.})$.

На рисунке 1а представлен оригинал $L(4,3)$ для испытания №1, предъявленного студентам 2 гр. Это решетка упорядоченных слов 4-х буквенного алфавита длиной не более 3-х символов. Благодаря введению отношений наглядности μ , решетка обладает рядом свойств, которые характеризуют понятие «красота». Уровневые числа $(1, 1, 2, 3, 4, 4, 5, 4, 4, 3, 2, 1, 1)$ симметричны. Количество ребер, находящееся между смежными уровнями, - симметрично. Решетка разбивается на непересекающиеся подрешетки (рисунки 1б, 1с, 1д). Сами подрешетки упорядочиваются отношением включения. Количество вершин $v(m,n)$ в произвольной решетке $L(m,n)$

равно $v(m, n) = \binom{m+n}{n} = \binom{m+n}{m}$, а ребер $e(m, n) = \binom{m+n-1}{n} \cdot n = \binom{m+n-1}{m} \cdot m$. Это далеко не все свойства, которые характеризуют рисунок 1а.

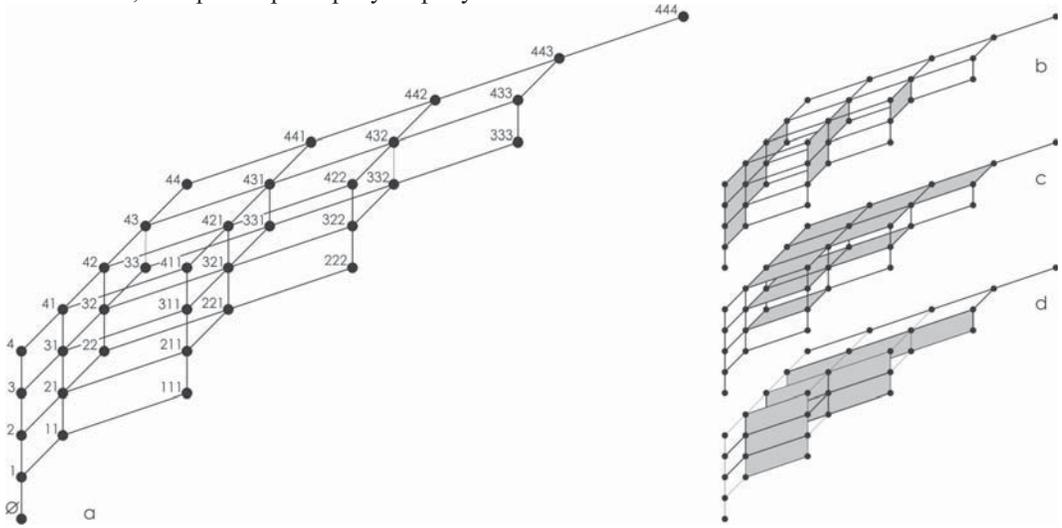


Рисунок 1 – Решетка $L(4,3)$ слева и её непересекающиеся подмножества 1b, 1c, 1d справа

После выбора объекта оценки было сформировано изображение для тестового задания, представленное на рисунке 2b. Здесь из оригинала (рисунок 2а) удалены все ребра. Примеры характерных ответов представлены на рисунках 2c и 2d. Задание следующее: «Соедините вершины прямыми ребрами, чтобы было красиво».

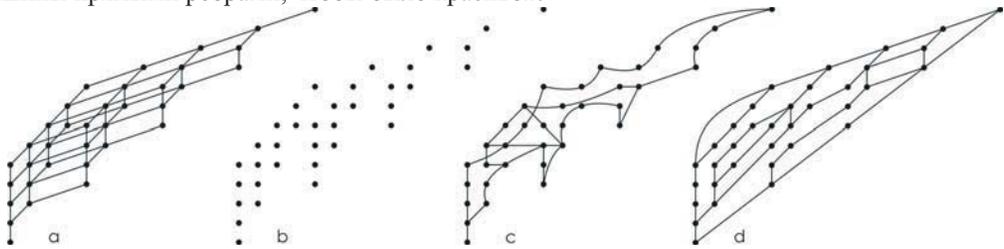


Рисунок 2 – Оригинал 2а, тестовое задание 2b и примеры решений 2c, 2d

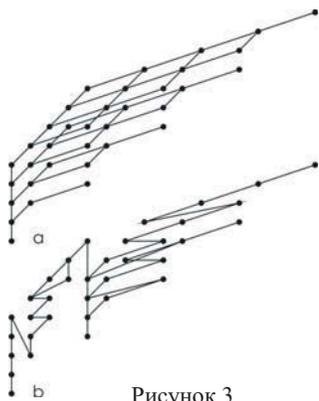


Рисунок 3

На рисунке 3 представлены результаты второго задания, для которого были представлены пронумерованные вершины решетки $L(4,3)$, в соответствии с рисунком 1а. Задание таково: «Соедините вершины ребрами в соответствии с нумерацией, чтобы было красиво». Анализируя результаты обоих заданий, заметим следующее. Наличие нумерации вершин позволило субъекту абсолютно точно воспроизвести логику изображения на рисунке 3а. Субъект, предъявивший рисунок 3b, возможно проигнорировал или не понял структурную нумерацию элементов. С точки зрения описанных выше критериев, рисунок 3а красив и, как следствие, позволяет вычислить ряд численных характеристик решетки. Рисунок 3b назвать красивым затруднительно. Возвращаясь к рисункам 2c и 2d заметим, что воспроизвести логику изображения, как это удалось на рисунке 3а, никому из 20 субъектов не уда-

лось, хотя действия их авторов понять можно. На рисунке 2с понятие красоты связано с задачей соединить все вершины меньшим числом линий. На рисунке 2b сделана попытка соединить вершины параллельными линиями. Что примечательно, в первом случае использовались в основном кривые линии, во втором – прямые.

Рассмотрим следующее задание.

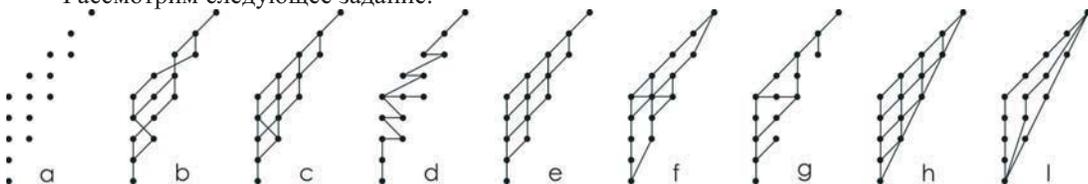


Рисунок 4 – Тестовое задание ($L(4,2)$, 1гр) на рисунке 4а и результаты 4b-4i

На рисунке 4а представлено изображение для тестового задания. Это решетка $L(4,2)$, у которой удалили все ребра. Отметим, что чем меньше количество вершин, тем больше результат совпадает с оригиналом. В частности рисунок 4е полностью совпадает с оригиналом. Во всех представленных рисунках, за исключением рисунка 4с, имеется симметрия относительно центральной оси. Общий подход к составлению теста позволяет произвести коллективную оценку исследуемых субъектов путем объединения и пересечения ребер на графах.

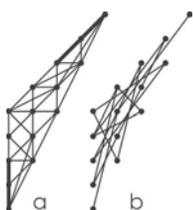
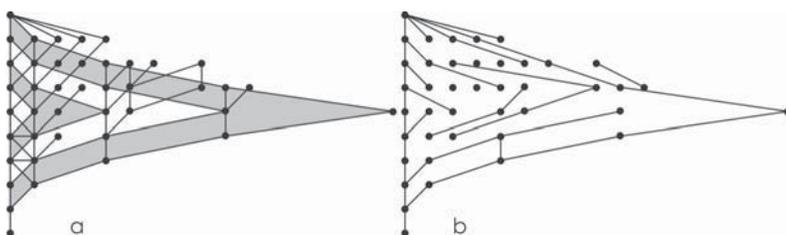


Рисунок 5

На рисунке 5а представлены результаты такого объединения. Мы составляем коллективное решение объединением ребер всех изображений. Не смотря на то что имеются отдельные не симметричные изображения, коллективное изображение обладает симметрией относительно центральной оси. Это наблюдение позволяет предположить, что интегрированные показатели стремятся к некоторому устойчивому состоянию, характеризующему группу субъектов. На рисунке 5b, наоборот, представлены ребра, отсутствующие на рисунке 5а. Из рисунка 5b видно, что в нем отсутствуют ребра удаленные друг от друга. Это своего рода «рационализм» в групповом понятии красоты. Соединяются только близлежащие вершины.

На рисунке 6а представлен фрагмент решетки делителей $P(n, \mu)$. Здесь $n=10$. Наглядные отношения μ такие же, как и для $L(m,n)$. Решетка $P(10, \mu)$ упорядочивает и наглядно изображает все возможные разбиения числа $n=10$. Количество разбиений $p(10)=42$.



Решетка $P(10,\mu)$ симметрична относительно центральной оси. В ней можно выделить подрешетки, упорядоченные отношениями включения (обозначены серым цветом).

Рисунок 6 – Решетка $P(10, \mu)$ на рисунке 6а и результат тестирования - 6b

На рисунке 6b показаны ребра, общие для оригинала и результата тестирования. Числовой мерой оценки схожести может служить отношение числа совпадающих ребер для каждого уровня к числу ребер оригинала. Для оригинала количество ребер равно $(5,7,9,12,9,5,4,2,1)$. Количество совпадающих ребер соответственно $(5,4,6,5,6,4,4,2,1)$. Общий процент совпадения $37/54=69\%$. (На рисунке 6а указаны не все ребра оригинала).

Системный анализ результатов

Как отмечалось ранее, задача данной работы – оценить способность субъекта воспринимать структуру, ее целостность. Такие качества необходимы для проектирования жизнеспособных систем, рассчитанных на длительный жизненный цикл. С этой целью мы предлагали для тестирования объекты, в которых «...симметрия является наиболее общим и наиболее мощным способом изучения структур; более того, симметрия какой-либо структуры, в свою очередь, имеет свою собственную структуру, которая также может изучаться с помощью симметрии и т.д.» [5]. Критерием было выбрано понятие красоты, поскольку «...симметрия обозначает тот вид согласованности отдельных частей, которое объединяет их в единое целое. Красота тесно связана с симметрией» [6, стр. 35].

Анализируя полученные от субъектов результаты, можно сделать следующие выводы.

Понятие красоты связано с понятием симметрии и асимметрии. Это хорошо видно по заданию ($L(4,2)$, 1 гр.).

Понятие красоты связано с понятием рациональности. Никто не изображал лишних дуг, несмотря на отсутствие каких-либо ограничений на их количество. Характерным объяснением этого является рисунок 5b, где изображены дуги, отсутствующие в полученных от субъектов заданиях.

Понятие красоты, как правило, не связано с понятием упрощенности. Все задания можно было выполнить путем соединения вершин только вертикальными и горизонтальными ребрами. Практически никто не воспользовался этой идеей, хотя она была на поверхности.

Многие тесты показали, что интуитивное нахождение самоподобных элементов структур, соотношение целого и его частей в исходном изображении не является приоритетом в понимании красоты.

Неожиданными результатами тестирования было одушевление некоторых изображений. Субъекты приписывали названия: это «елочка», «ежик» и т.д. Несмотря на то, что была четко поставлена задача об использовании только прямых ребер, некоторые субъекты не смогли отказать от изображения кривых ребер, поскольку ряд вершин визуально казались им лежащими на дуге.

Подсчет характеристик испытаний методами наглядного комбинаторного анализа

В конечном итоге для наших визуальных наблюдений и системного анализа результатов необходимо предложить меры и шкалы которыми можно объективно характеризовать результаты испытаний. Подсчет числа вершин и ребер оригинала и результатов – один из таких методов. При подсчете обобщенного результата испытаний группы вхождение ребер в результирующей решетке имеет нечеткий характер. Иногда требуется решить перечислительные задачи теории графов. В частности подсчитать количество связанных помеченных графов на n вершинах (number of connected labeled graphs with n nodes), а также число помеченных связанных графов каждой компоненты разбиения, если граф на n вершинах не связан. Первая задача решается путем вычисления коэффициентов производящей функции [7 с. 138] по формулам (1), (2). Последовательность числа графов представлена в EIS A001187 [8].

$$(1) \quad K(z) = \ln \left(1 + \sum_{n \geq 1} 2^{\binom{n}{2}} \cdot \frac{z^n}{n!} \right).$$

$$(2) \quad K(n) = 2^{\binom{n}{2}} - \frac{1}{2} \sum \binom{n}{n_1, n_2} 2^{\binom{n_1}{2} + \binom{n_2}{2}} + \frac{1}{3} \sum \binom{n}{n_1, n_2, n_3} 2^{\binom{n_1}{2} + \binom{n_2}{2} + \binom{n_3}{2}} - \dots,$$

где $n_1+n_2+\dots+n_k=p$ и $0 < n_j < n$. Для второй задачи предложен метод, основанный на наглядном комбинаторном анализе [3]. Результат для первых $n=2,3, 4,5$ вершин и $k=1,2,\dots,10$ ребер представлен в таблице 1, которая продолжается вправо вверх, используя значения предыдущих столбцов.

Таблица 1

10				1
9				10
8				45
7				120
6			1	205
5			6	222
4			15	125
3		1	16	
2		3		
1	1			
<i>n</i>	2	3	4	5

Таблица составлена с использованием формулы (3) для каждого элемента разбиения числа n .

$$(3) \quad G(S) = \frac{\prod_{i=0}^{K-2} \left(\sum_{j=i}^{K-1} c_j \right)}{f(S,1,1)}.$$

В (3) S – разбиение числа записанное в строковом виде. Например, разбиение $(c_1, c_2, c_3, c_4) = 3+3+2+2$ записывается как $S = (s_1, s_2, s_3, s_4) = \langle 3322 \rangle$. В столбцах таблицы представлены значения для количества связных графов с k ребрами n вершинами. Общее количество связных ориентированных графов равно сумме всех значений соответствующего столбца n . Функция $f(S,1,1)$ вычисляется по рекурсивной формуле (4), представленной в нотации Mathcad. Она подсчитывает произведение факториалов количества вхождений одинаковых символов строки S .

$$(4) \quad f(S, b, F) := \begin{cases} F & \text{if } S = "" \\ f(\text{substr}(S,1, \text{strlen}(S) - 1), 1, F \cdot b!) & \text{if } \text{substr}(S,0,1) \neq \text{substr}(S,1,1) \\ f(\text{substr}(S,1, \text{strlen}(S) - 1), b + 1, F) & \text{otherwise} \end{cases}$$

В нашем примере $f(S,1,1) = 2! \cdot 2! = 4$. Чтобы получить очередной столбец в таблице, все $G(S)$ умножаются на соответствующие значения предыдущих столбцов и складываются. Результат вычитается из количества сочетаний $C(K,k)$, где $K = C(n,2)$.

Заключение

Любые сложные системы проектируют и эксплуатируют люди. Конечная цель работы - придать процессу проектирования системный подход, помочь проектировщику отойти от двумерного восприятия действительности как прямоугольной таблицы.

Проведенные в рамках данной статьи эксперименты говорят о том, что как минимум в рамках предложенных тестов мир не всегда воспринимается структурно. В связи с этим актуальна задача понимания восприятия субъектом окружающей действительность и соответствующая адаптация учебного процесса.

Использование методов наглядного комбинаторного анализа является действенным инструментом решения перечислительных задач.

Список литературы

- [1] Горохов В.Л. Механизмы восприятия когнитивных образов, технические пути их активизации// Труды Международной конференции по мягким вычислениям и измерениям. SCM-2002. СПб.: Гидрометеоздат, 2002 г.
- [2] Винер Н. Я математик. Пер. с англ. Издание второе. Москва, Наука, 1967 г. 253 с.
- [3] Кистанов А.М. Орлов С.П. Наглядный комбинаторный анализ информационных транзакционных систем. – Самара, СНЦ РАН.- 2008. – 206 с.
- [4] Биркгоф Г. Математика и психология. – ЛКИ., 2008г. – 112 с.

- [5] Бутковский А.Г. Философия, методология и математика единой геометрической теории управления//Труды IV международной конференции «Проблемы управления и моделирования в сложных системах». Самара, 17-23 июня 2002 г. с. 1-6.
- [6] Вейль. Г. Симметрия. Изд. 2-е стереотипное. - М.: Эдиториал УРСС, 2003.- 192 с.
- [7] P. Flajolet and R. Sedgewick. Analytic Combinatorics. - Cambridge University Press. 2009, 824 p.
- [8] <http://www.research.att.com/~njas/sequences/>.

АНАЛИТИЧЕСКОЕ КОНСТРУИРОВАНИЕ СИСТЕМ УПРАВЛЕНИЯ МНОГОСВЯЗНЫМИ ДИНАМИЧЕСКИМИ ОБЪЕКТАМИ С УЧЁТОМ ПЕРЕДАТОЧНЫХ НУЛЕЙ

А.З. Асанов, Д.Н. Демьянов

Казанский государственный университет
423812, Набережные Челны, пр-т Сююмбике, 10А, Россия
demyanovdn@mail.ru
тел: +7 (8552) 38-34-05, факс: +7 (8552) 58-88-94

Ключевые слова: многосвязный динамический объект, система управления, передаточный нуль, неминимально-фазовость

Abstract

A method of construction of control systems for nonminimal-phase multiconnected dynamic objects is offered. The method is based on the preliminary correction of the transmission zeros spectrum and the construction of regulator with the help of known principles. One of the variants of control system construction for nonminimal-phase non-proper objects, founded on the multiple-loop and dynamic link decoupling principles is scrutinized.

Введение

Большинство проектируемых и используемых в настоящее время технологических объектов отличаются значительной сложностью, многокомпонентностью и многосвязностью. Одним из способов формального описания особенностей структуры сложного динамического объекта является использование понятия “системный нуль” [1]. Спектр системных нулей характеризует управляемость и наблюдаемость объекта, а также особенности взаимодействия каналов управления.

Для полностью управляемого и наблюдаемого динамического объекта (а именно такие и используются на практике в подавляющем большинстве случаев) понятие системного нуля совпадает с так называемым передаточным нулём – значением комплексной частоты, при котором происходит уменьшение нормального ранга передаточной матрицы [2]. Спектр передаточных нулей характеризует взаимовлияние каналов передачи сигнала, определяет степень взаимодействия отдельных подсистем динамического объекта.

Известно, что принципиальная разрешимость многих задач управления, адаптации и идентификации, а также степень оптимальности получаемых решений во многих случаях зависит от расположения передаточных нулей динамического объекта. Например, при наличии неминимально-фазовых (с положительной вещественной частью) или находящихся в начале координат передаточных нулей невозможно построение устойчивой инвариантной системы управления с большим коэффициентом усиления, нахождение устойчивого регулятора, обеспечивающего динамическую развязку каналов с заданными показателями качества, возникают сложности в задачах слежения и подавления возмущений и т. д. [2-4]. Кроме того, для неминимально-фазовых объектов характерны забросы переходной характеристики в отрицательную область и возникновение так называемых “нулевых” режимов работы, когда на определенной частоте отсутствует реакция на изменение входного сигнала [5].

Таким образом, одним из условий построения качественной системы управления многосвязным динамическим объектом является учёт спектра его передаточных нулей. Если объект является минимально-фазовым, то для синтеза регуляторов можно использовать любой из известных методов. Если же объект является неминимально-фазовым, то необходимо либо

уменьшать требования к показателям качества проектируемой системы управления (например, допустить отрицательное перерегулирование и возможность нулевых режимов работы, лишь частичную развязку каналов), либо разрабатывать специализированные алгоритмы управления подобными объектами.

В представленной работе рассматривается один из возможных подходов к синтезу систем управления неминимально-фазовыми динамическими объектами, заключающийся в предварительной коррекции спектра его передаточных нулей и последующем синтезе регуляторов на основе известных методик.

1 Методы коррекции спектра передаточных нулей

Все известные в настоящее время подходы к изменению спектра передаточных нулей многосвязного динамического объекта можно разделить на две основных группы: методы компенсации и методы сдвига.

Под методами компенсации понимается совокупность алгоритмов, приводящих к уменьшению количества конечных передаточных нулей объекта. При этом можно удалить определённые (как правило, неминимально-фазовые) передаточные нули, или же сформировать объект, полностью их лишенный. К данной группе можно отнести методы, основанные на применении нерегулярных законов управления [6] и некоторые методы, предполагающие модификацию входной или выходной матрицы динамического объекта [7].

На практике методы компенсации используются достаточно редко, так как имеют целый ряд недостатков. Применение нерегулярных законов управления приводит к тому, что передаточная матрица системы управления приобретает матричные делители нуля (левый или правый) и все сопряженные с этим негативные эффекты [6]. Кроме того, расчёт нерегулярных законов управления является достаточно сложной и трудоёмкой задачей.

При использовании методов модификации входной или выходной матрицы основная проблема заключается в негрубости полученных решений в подавляющем большинстве случаев. Малые изменения коэффициентов, неизбежные на практике, приводят к появлению у динамического объекта передаточных нулей, имеющих большое по модулю значение. Без детального исследования вопросов робастности данного метода его использование представляется нецелесообразным.

Под методами сдвига понимается совокупность алгоритмов, приводящих к смещению всех передаточных нулей динамического объекта в заданную область левой полуплоскости. К данной группе относятся методы шунтирования [8], методы, предполагающие модификацию входной или выходной матрицы динамического объекта [9, 10] и методы, основанные на введении обратной связи [11].

Каждый из перечисленных методов обладает определенными достоинствами и недостатками, во многом определяющими его область применения. Использование параллельного компенсатора (шунта) достаточно широко применяется в электрических, гидравлических и пневматических системах. Однако его использование в механических или электромеханических системах ограничено сложностью (иногда и невозможностью) физической реализации. Кроме того, расчет параметров шунта в случае многосвязного объекта является достаточно трудоёмкой задачей.

Методы, предполагающие модификацию входной или выходной матрицы динамического объекта, являются достаточно универсальными. Они позволяют решать широкий круг задач, используя достаточно простые и легко поддающиеся автоматизации алгоритмы. Однако с физической точки зрения такая коррекция структуры и значений коэффициентов матриц динамического объекта соответствует изменению параметров датчиков или исполнительных механизмов. Таким образом, указанные методы более эффективны при формировании требований к вновь проектируемым технологическим объектам.

Методы, основанные на введении обратной связи, могут быть использованы лишь для не собственных динамических объектов – таких, у которых входной сигнал передаётся непосредственно на выход (например, силовая установка летательного аппарата на некоторых режимах работы). Передаточные нули только таких объектов не являются инвариантными к действию статической обратной связи по состоянию [12]. С математической точки зрения, в данном случае модификация спектра передаточных нулей сводится к решению задачи модального управления некоторым объектом и может решаться с использованием классических методов. Недостатком указанного подхода является то, что он применим лишь к не собственным динамическим объектам, матрицы коэффициентов которого удовлетворяют определённым ограничениям [11]. Введение обратной связи для управления спектром передаточных нулей может использоваться в совокупности с методом шунтирования для того, чтобы облегчить процедуру расчёта и упростить структуру параллельного компенсатора.

Сравнительный анализ области применения и возможности практической реализации приведённых выше методов сдвига позволяет выстроить следующую иерархию.

Методы изменения спектра передаточных нулей, основанные на модификации входной или выходной матрицы, наиболее универсальны и применимы к любым типам объектов. Однако их практическая реализация часто требует существенной перестройки объекта управления и, как следствие, значительных затрат. Подобные методы наиболее целесообразно применять на этапе разработки и проектирования технологических объектов, закладывая в них требуемые структурные свойства. В дальнейшем, уже в процессе синтеза системы управления, обеспечение необходимых показателей качества может осуществляться с помощью классических методов.

Методы изменения спектра передаточных нулей, основанные на использовании параллельного компенсатора, являются менее универсальными. Для некоторых групп динамических объектов (например, автомобиль или самолёт) понятие шунта просто не имеет физического смысла. Однако в тех областях, в которых использование параллельного компенсатора возможно (электроника, гидравлические и пневматические устройства), затраты на реализацию данного метода коррекции неминимально-фазовости значительно ниже, чем при перестройке компонентов объекта управления. Введение шунта обеспечивает предварительную коррекцию структурных свойств динамического объекта, а обеспечение требуемых показателей качества осуществляется путем построения регулятора по любой из известных методик.

Методы изменения спектра передаточных нулей, использующие введение обратной связи по состоянию, могут использоваться лишь для отдельного класса объектов, матрицы коэффициентов которых удовлетворяют изложенным в работе [11] критериям. При этом данный метод отличается простотой реализации (введение обратной связи для обеспечения требуемых характеристик является одним из классических подходов в инженерной практике) и требует минимальных затрат. Используя идеи многоконтурности [4] можно отдельно строить внутренний контур, обеспечивающий необходимые структурные свойства, и внешний контур, обеспечивающий требуемые показатели качества. В некоторых случаях расчёт обоих контуров может вестись совместно (например, построение регулятора, обеспечивающего устойчивость и минимально-фазовость). Для упрощения расчётов и расширения области применения указанный метод может использоваться в совокупности с методом шунтирования: путём введением шунта достаточно простой структуры объект преобразуется в не собственный, и дальнейший синтез законов управления осуществляется на основе приведённых выше принципов.

2 Синтез алгоритмов управления неминимально-фазовым объектом

Основываясь на приведённой выше классификации методов сдвига, можно предложить алгоритм синтеза системы управления неминимально-фазовым объектом, основанный на предварительной коррекции спектра его передаточных нулей.

Если исходный объект является несобственным и удовлетворяет ограничениям, изложенным в [11], то структурная схема его системы управления имеет вид, представленный на рисунке 1.

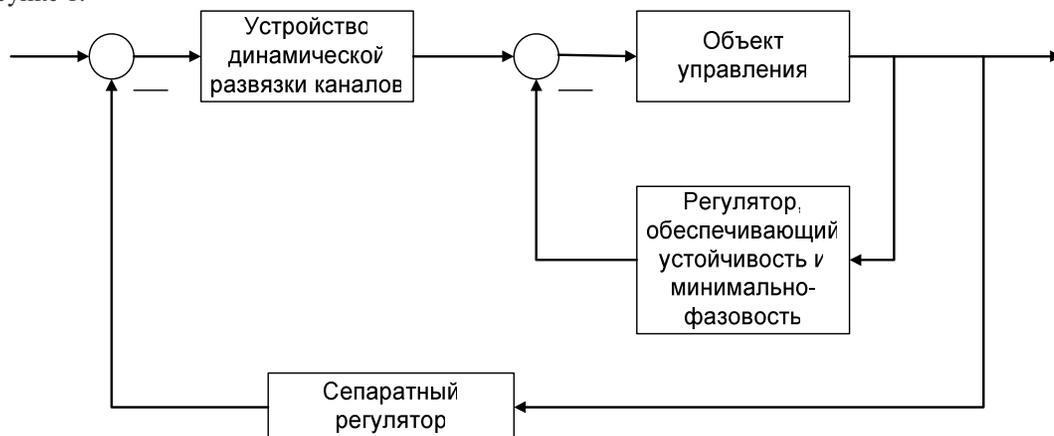


Рисунок 1 – Структурная схема системы управления неминимально-фазовым несобственным объектом

Предлагаемая структура системы управления базируется на идеях многоконтурности [4] и динамической развязки каналов [13].

Внутренний контур обратной связи обеспечивает минимально-фазовость и, если нужно, устойчивость. В работе [11] показано, что процедура синтеза статического регулятора в этом случае может быть сведена к задаче квадратичной стабилизации нескольких объектов и решена, например, с применением аппарата линейных матричных неравенств [14]. Таким образом, внутренний контур обратной связи обеспечивает системе требуемые структурные свойства.

После этого для устойчивого и минимально-фазового объекта синтезируется устройство динамической развязки каналов [13], которое также будет устойчивым и минимально-фазовым [4]. Для автоматизации расчёта коэффициентов передаточной матрицы может быть использован изложенный в [15] алгоритм, основанный на применении технологии канонизации матриц. Введение устройства динамической развязки каналов приводит к тому, что исходный многосвязный объект разделяется на совокупность независимых подсистем.

Для каждой из независимых подсистем строится свой контур обратной связи с использованием классических методов управления скалярными объектами. Их совокупность образует внешний контур обратной связи, обозначенный на рисунке 1 как сепаратный регулятор. Внешний контур обратной связи обеспечивает требуемые показатели качества по каждому из каналов управления.

Таким образом, сложная задача синтеза системы управления многосвязным динамическим объектом может быть разбита на несколько этапов, на каждом из которых расчёт осуществляется с использованием известных методик. Недостатком предлагаемой схемы является то, что она применима лишь для несобственных объектов.

Если исходный объект не удовлетворяет изложенным в работе [11] требованиям, то для него проверяется возможность построения параллельного компенсатора [8]. Если физическая природа объекта такова, что для него можно использовать методы шунтирования, то структурная схема его системы управления имеет вид, представленный на рисунке 2.

Введение параллельного компенсатора позволяет скорректировать некоторые структурные свойства динамического объекта, в частности, спектр его передаточных нулей. Однако это не позволяет обеспечить все требуемые показатели качества (например, невозможно только лишь с помощью шунта осуществить робастную стабилизацию или развязку каналов неустойчивого объекта).

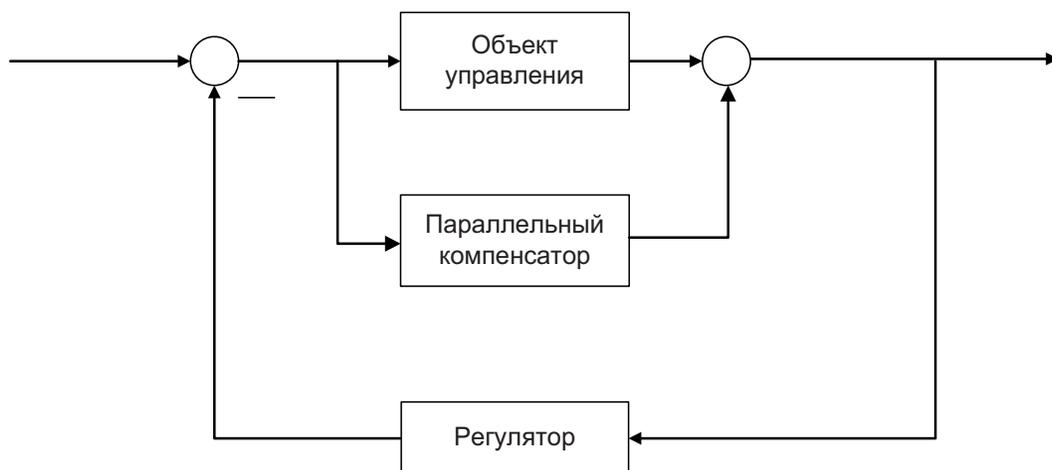


Рисунок 2 – Структурная схема системы управления неминимально-фазовым объектом, допускающим введение параллельного компенсатора

Минимально-фазовость шунтированного объекта позволяет использовать для дальнейшего синтеза системы управления специализированные методы, разработанные для многосвязных случаев. Синтез регулятора может осуществляться, например, с помощью технологии вложения систем [6], основанной на применении новейших достижений матричной математики. При этом на представленной структурной схеме возможно также введение дополнительного предкомпенсатора, предназначенного для первоначальной обработки входного сигнала.

Таким образом, в рассмотренном случае этап коррекции спектра передаточных нулей является подготовительным этапом, предшествующим непосредственно построению регулятора. Шунт служит лишь для обеспечения требуемых структурных свойств объекта управления, позволяя обходить ограничения известных методов синтеза. При расчёте самого регулятора приходится использовать алгоритмы, разработанные для многосвязных объектов. Это усложняет процедуру расчёта по сравнению с предыдущей схемой. Однако процедура коррекции спектра передаточных нулей путём введения параллельного компенсатора может быть использована для большего количества динамических объектов.

Если обе приведённых выше структуры системы управления не могут быть использованы (не выполняются ограничения [11] и объект не допускает введения параллельного компенсатора), то единственным способом коррекции спектра передаточных нулей является изменение входной или выходной матрицы динамического объекта. Представленный в работе [9] алгоритм можно рассматривать в качестве подготовительного этапа, предшествующего синтезу системы управления по одной из известных методик. Показано, что в некоторых случаях для обеспечения минимально-фазовости объекта управления может потребоваться его существенная перестройка. В таком случае процесс модификации входной и выходной матрицы может осуществляться последовательно в итерационном режиме. Для уменьшения затрат на последующую доработку представляется целесообразным формулировать требования к структуре ещё на этапе проектирования технологических объектов, закладывая такую совокупность передаточных нулей, которая не потребует дальнейшей коррекции.

Модификация спектра передаточных нулей путём изменения матриц входа или выхода в общем случае является достаточно универсальным подходом, применимым к любым объектам управления. Однако, затраты на его реализацию являются, как правило, весьма существенными и оправданными лишь при построении высокоточных мехатронных систем управления. Указанный подход может рассматриваться в качестве крайней меры, когда никакие другие методы не дают приемлемых результатов.

Заключение

В работе рассмотрены различные подходы к управлению спектром передаточных нулей многосвязного динамического объекта, проанализированы их достоинства, недостатки и область применения. Предложен подход к синтезу систем управления неминимально-фазовыми объектами, основанный на предварительной коррекции спектра их передаточных нулей. Рассмотрена структурная схема и последовательность расчёта системы управления неминимально-фазовым несобственным объектом, базирующаяся на идеях многоконтурности и динамической развязки каналов.

Благодарности

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант 08-08-00536).

Список литературы

- [1] Смагина Е.М. Вопросы анализа линейных многомерных объектов с использованием понятия нуля системы. – Томск: Изд-во Том. ун-та. 1990. – 160 с.
- [2] Смагина Е.М. Нули линейных многомерных систем. Определение, классификация, применение. // Автоматика и телемеханика. – 1985. – № 12. – С. 5–33.
- [3] Васильев В.И., Шаймарданов Ф.А. Синтез многосвязных автоматических систем методом порядкового отображения. М: Наука. 1983. – 126 с.
- [4] Методы классической и современной теории автоматического управления: Учебник в 5-и тт.; 2-е изд. Т.3: Синтез регуляторов систем автоматического управления / Под ред. К.А. Пупкова и Н.Д. Егупова. М.: Издательство МГТУ им. Н.Э. Баумана. 2004. – 616 с.
- [5] Конструирование входных/выходных матриц модели многосвязной динамической системы с заданием некоторых инвариантов / Асанов А.З., Демьянов Д.Н. // Идентификация систем и задачи управления SICPRO'09: Труды VIII Международной конференции. – Москва: ИПУ РАН, 2009. – С. 1390-1421.
- [6] Буков В.Н. Вложение систем. Аналитический подход к анализу и синтезу матричных систем – Калуга: изд. научной лит. Н.Ф. Бочкаревой. 2006. – 720 с.
- [7] Обеспечение заданного спектра передаточных нулей многосвязной динамической системы путём формирования выходной матрицы / Асанов А.З., Демьянов Д.Н. // Проблемы управления и моделирования в сложных системах: Труды X Международной конференции. – Самара: Самарский НЦ РАН, 2008. – С. 226-231.
- [8] Фрадков А.Л. Адаптивная стабилизация минимально-фазовых объектов с векторным входом без измерения производных от выхода // Докл. РАН. – 1994. – Т. 337. – № 5. – С. 592–594.
- [9] Синтез вход/выходных матриц многосвязной динамической системы по заданным передаточным нулям / Асанов А.З., Демьянов Д.Н. // Известия РАН. Теория и системы управления. – 2008. – № 6. – С. 5–18.
- [10] Вопросы анализа и синтеза несобственных динамических систем с учетом передаточных нулей / Асанов А.З., Демьянов Д.Н. // Изв. вузов. Авиационная техника. – 2009. – №2. – С. 24-28.
- [11] Одновременная стабилизация и обеспечение минимально-фазовости многосвязного динамического объекта / Асанов А.З., Демьянов Д.Н. // Мехатроника, автоматизация, управление – 2009: Материалы Международной научно-технической конференции. – Таганрог: ТТИ ЮФУ, 2009. – С. 27-30.
- [12] Редукция матрицы Розенброка при анализе инвариантных нулей линейной МИМО-системы / Мисриханов М.Ш., Рябченко В.Н. // Автоматика и телемеханика. – 2008. – № 10. – С. 31–47.
- [13] Метод динамической развязки каналов управления в многосвязных объектах / Филимонов А.Б., Филимонов Н.Б. // Изв. вузов. Приборостроение. – 2002. – №7 – С. 28 -34.
- [14] Баландин Д.В., Коган М.М. Синтез законов управления на основе линейных матричных неравенств. М.: ФИЗМАТЛИТ. 2007. – 280 с.
- [15] Алгоритм обращения полиномиальных матриц и его использование в задаче динамической развязки каналов / Асанов А.З., Демьянов Д.Н. // Мехатроника, автоматизация, управление – 2009: Материалы Международной научно-технической конференции. – Таганрог: ТТИ ЮФУ, 2009. – С. 24-27.

СИНТЕЗ СИСТЕМЫ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ ДЛЯ МНОГОСВЯЗНОГО ОБЪЕКТА С ЗАПАЗДЫВАНИЯМИ ПО УПРАВЛЕНИЮ НА ОСНОВЕ КОМПЕНСАЦИОННО-НАБЛЮДАТЕЛЬНОЙ СХЕМЫ

А.З. Асанов, В.С. Каримов

Казанский государственный университет
423812, Набережные Челны, пр-т Сююмбике, 10А, Россия
kvs-chelny@mail.ru
тел: +7 (8552) 38-34-05, факс: +7 (8552) 58-88-94

***Ключевые слова:** система автоматического управления, многосвязный объект, запаздывания по управлению, компенсационно-наблюдательная схема, технология вложения систем*

Abstract

Design of automatic control system for the multivariable controlled plant with time delays on input vector was considered. For reproduction state vector of controlled plant without time-delays using compensating observation scheme. Design of system on forced component closed-loop dynamic system was carried out on basis of system embedding technology. Compute the algebraical expressions which to determine solution set of problem design and existence conditions this solution.

Введение

Запаздывания, встречающиеся в объектах управления, негативно влияют на процессы регулирования и могут приводить к потере устойчивости системы управления, ухудшению качества процессов управления и т.д. Поэтому их так важно учитывать при проектировании систем автоматики. Проблема управления этими объектами усложняется в нашем случае тем, что они являются многосвязными.

В настоящее время проблему управления объектами с запаздываниями по входу зачастую решают следующим образом. Звено запаздывания в матричной передаточной функции объекта управления заменяют аппроксимацией Паде и затем решают задачу синтеза системы классическим методом, например, модальным. Однако, этот способ дает лишь приближенное решение задачи синтеза и ведет к увеличению размерности математической модели объекта управления, а, следовательно, усложнению решения, в соответствии с числом запаздываний в нем.

В данной работе предлагается точное аналитическое решение задачи синтеза системы автоматического управления для многосвязного объекта с запаздываниями по управлению на основе технологии вложения систем. Предполагается, что фазовый вектор состояния объекта недоступен для непосредственного наблюдения. Поэтому на первом этапе решения задачи рассматривается компенсационно-наблюдательная схема управления, которая воспроизводит фазовый вектор состояния объекта без запаздываний. На втором этапе определяется с помощью технологии вложения систем [1, 2] матричная передаточная функция регулятора или условия ее определяющие, при которой поведение системы управления будет описываться желаемой матричной передаточной функцией. Синтез проводится по вынужденной составляющей замкнутой динамической системы. Находятся соотношения, определяющие множество решений задачи синтеза, и условия существования этих решений.

1 Постановка задачи исследования

Пусть наблюдаемый и управляемый линейный стационарный объект с сосредоточенными запаздываниями по управлению может быть представлен в виде дифференциально-разностных уравнений:

$$(1) \quad \begin{aligned} \dot{x}(t) &= Ax(t) + \sum_{i=0}^r B_i u(t - \tau_i), \\ y(t) &= Cx(t) + \sum_{i=0}^r D_i u(t - \tau_i), \end{aligned}$$

где $\tau_0 = 0$, $0 < \tau_1, \tau_2, \dots, \tau_l$ – постоянные времена запаздываний, $i = 0, 1, \dots, l$, $u(t) \in R^s$ – вектор входных переменных, $y(t) \in R^m$ – вектор выходных переменных, $x(t) \in R^n$ – фазовый вектор объекта управления. В нашем случае матрица A имеет размер $n \times n$, $B_i - n \times s$, $C - m \times n$, $D_i - m \times s$. Матрицы B_i являются постоянными матрицами при временах запаздывания τ_i .

Слагаемые $B_i u(t - \tau_i)$ описывают запаздывания сигналов на время τ_i во внутренних каналах ОУ.

Начальные условия зададим с учетом запаздывания сигналов в объекте управления (ОУ) – формально будем рассматривать отрицательные моменты времени $t < 0$, предполагая, что в объекте происходили динамические процессы до начального момента времени:

$$x(t) = \varphi_x(t), \quad t_0 - \tau \leq t \leq t_0,$$

где τ – наибольшее время запаздывания.

Решение задачи управления предлагается разделить на два этапа.

Предположим, что фазовый вектор состояния объекта недоступен для непосредственного наблюдения, тогда на первом этапе необходимо скомпенсировать запаздывания по управлению путем построения наблюдателя, вычисляющего текущий фазовый вектор $x(t)$ объекта по информации о сигналах $u(t)$ и $y(t)$.

На втором этапе необходимо найти матричную передаточную функцию (МПФ) регулятора $K(p)$ или условия ее определяющие, при которой поведение системы автоматического управления (САУ) будет описываться желаемой матричной передаточной функцией $E_y^v(p)$.

Согласно [3], объект управления должен быть устойчивым, в противном случае можно получить неустойчивую САУ.

2 Компенсация запаздываний по управлению путем применения компенсационно-наблюдательной схемы

Пусть компенсационно-наблюдательная схема управления, вычисляющая текущий фазовый вектор $x(t)$ объекта может быть представлена уравнениями:

$$(2) \quad \begin{aligned} \dot{\hat{x}}_i(t) &= A\hat{x}_i(t) + \sum_{i=0}^r B_i u(t - \tau_i), \\ y_i(t) &= LC\hat{x}_i(t) + L \sum_{i=0}^r D_i u(t - \tau_i), \\ \hat{\hat{x}}(t) &= A\hat{x}(t) + Ly(t) - y_i(t) + \sum_{i=0}^r B_i u(t), \end{aligned}$$

где $\hat{x}(t)$ – фазовый вектор состояния наблюдателя, $x_i(t)$ – фазовый вектор состояния модели объекта, $y_i(t)$ – вектор выхода модели объекта, L – матрица наблюдателя.

Проанализируем действие наблюдателя, рассмотрев уравнения (1) и (2). Для этого найдем МПФ $F_{\hat{x}}^u(p)$ от управляющего сигнала $u(t)$ к фазовому вектору состояния наблюдателя $\hat{x}(t)$.

Применим преобразования Лапласа к уравнениям (1) и (2):

$$\begin{aligned} px(p) &= Ax(p) + \sum_{i=0}^r B_i e^{-\tau_i p} u(p), \\ y(p) &= Cx(p) + \sum_{i=0}^r D_i e^{-\tau_i p} u(p), \\ px_m(p) &= Ax_m(p) + \sum_{i=0}^r B_i e^{-\tau_i p} u(p), \\ y_m(p) &= LCx_m(p) + L \sum_{i=0}^r D_i e^{-\tau_i p} u(p), \\ p\hat{x}(p) &= A\hat{x}(p) + Ly(p) - y_m(p) + \sum_{i=0}^r B_i u(p). \end{aligned}$$

Далее, производя несложные преобразования, получим выражение:

$$\begin{aligned} (pI_n - A)\hat{x}(p) &= L(C(pI_n - A)^{-1} \sum_{i=0}^r B_i e^{-\tau_i p} u(p) + \sum_{i=0}^r D_i e^{-\tau_i p} u(p)) - \\ &- LC(pI_n - A)^{-1} \sum_{i=0}^r B_i e^{-\tau_i p} u(p) + L \sum_{i=0}^r D_i e^{-\tau_i p} u(p) + \sum_{i=0}^r B_i u(p). \end{aligned}$$

Сократив одинаковые члены, уравнение будет иметь вид:

$$\hat{x}(p) = (pI_n - A)^{-1} \sum_{i=0}^r B_i u(p).$$

Следовательно, МПФ $F_{\hat{x}}^u(p)$ можно представить следующим образом:

$$F_{\hat{x}}^u(p) = \frac{\hat{x}(p)}{u(p)} = (pI_n - A)^{-1} \sum_{i=0}^r B_i,$$

которая может соответствовать объекту управления

$$\dot{x}(t) = Ax(t) + \sum_{i=0}^r B_i u(t).$$

Таким образом, применение компенсационно-наблюдательной схемы, описываемой уравнениями (2), позволяет полностью компенсировать запаздывания по управлению и воспроизводить фазовый вектор состояния $x(t)$ объекта.

3 Синтез САУ на основе технологии вложения систем

Применение преобразования Лапласа к уравнениям (1) дает операторную форму описания объекта:

$$(3) \quad \begin{aligned} px(p) &= Ax(p) + \sum_{i=0}^r B_i e^{-\tau_i p} u(p), \\ y(p) &= Cx(p) + \sum_{i=0}^r D_i e^{-\tau_i p} u(p). \end{aligned}$$

Пусть закон управления в общем случае описывается матричным уравнением:

$$(4) \quad v(p) = K(p)\hat{x}(p) + u(p),$$

где $K(p)$ – МПФ регулятора размера $n \times s$, $v(p) \in R^s$ – вектор управления на входе системы.

Пусть желаемое поведение САУ задается МПФ $E_y^v(p)$ от управляющих воздействий v к выходу y .

Для ОУ (1) и закона управления (4) необходимо найти МПФ регулятора $K(p)$ или условия ее определяющие, при которых поведение САУ (рисунок 1) будет описываться желаемой МПФ $E_y^v(p)$.

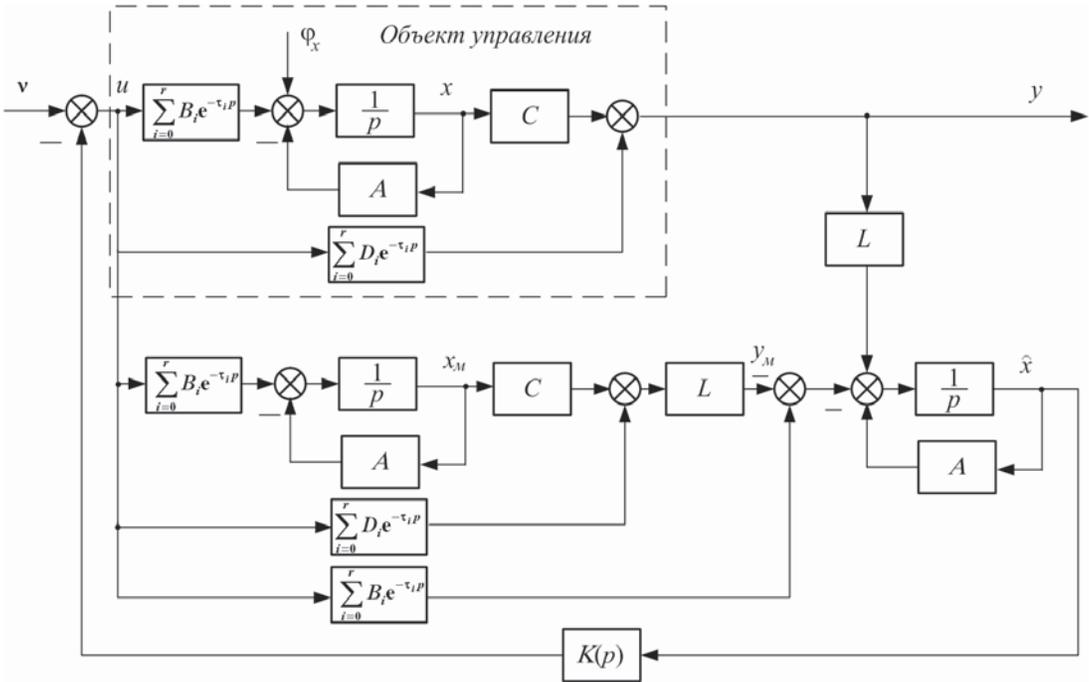


Рисунок 1 – Система автоматического управления с запаздываниями

Далее при синтезе системы управления будем использовать технологию вложения систем [1, 2]. С учетом уравнений (1), (2), (4) и выполнения процедур метода вложения систем проблемная матрица (проматрица) рассматриваемой задачи будет иметь вид:

$$\Omega(p) = \begin{bmatrix} (pI_n - A) & 0 & 0 & 0 & 0 & -\sum_{i=0}^r B_i e^{-\tau_i p} & 0 \\ 0 & (pI_n - A) & 0 & 0 & 0 & -\sum_{i=0}^r B_i e^{-\tau_i p} & 0 \\ 0 & 0 & (pI_n - A) & -L & I_m & -\sum_{i=0}^r B_i & 0 \\ -C & 0 & 0 & I_m & 0 & -\sum_{i=0}^r D_i e^{-\tau_i p} & 0 \\ 0 & -LC & 0 & 0 & I_m & -L \sum_{i=0}^r D_i e^{-\tau_i p} & 0 \\ 0 & 0 & K & 0 & 0 & I_s & -I_s \\ 0 & 0 & 0 & 0 & 0 & 0 & I_s \end{bmatrix}.$$

Репроматрица системы в обобщенном виде представляется следующим образом

$$\Omega^{-1}(p) = \begin{bmatrix} E_x^{\varphi_x}(p) & * & * & * & * & * & E_x^v(p) \\ E_{x_m}^{\varphi_x}(p) & * & * & * & * & * & E_{x_m}^v(p) \\ E_{\bar{x}}^{\varphi_x}(p) & * & * & * & * & * & E_{\bar{x}}^v(p) \\ E_y^{\varphi_x}(p) & * & * & * & * & * & E_y^v(p) \\ E_{y_m}^{\varphi_x}(p) & * & * & * & * & * & E_{y_m}^v(p) \\ E_u^{\varphi_x}(p) & * & * & * & * & * & E_u^v(p) \\ 0 & 0 & 0 & 0 & 0 & 0 & I_s \end{bmatrix}.$$

Здесь $E_j^i(p)$ – МПФ от параметра i к параметру j . Блоки, не представляющие особого интереса, отмечены звездочками.

Матрицы вложения α и β , используемые при вложении систем имеют вид:

$$\alpha = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ I_s]^0, \quad \beta = [0 \ 0 \ 0 \ I_m \ 0 \ 0 \ 0] \quad \text{при } \omega = E_y^v(p),$$

где ω – образ синтезируемой системы – желаемая передаточная функция системы.

После выполнения процедур технологии вложения – последовательной факторизации матриц $\Omega = \Sigma \Xi$, $\alpha = \Sigma \delta$, $\beta = \pi \Xi$, $\omega = \pi \delta$, можно получить уравнения, которым должна удовлетворять МПФ регулятора $K(p)$ при синтезе по вынужденной составляющей $E_y^v(p)$ замкнутой динамической системы.

Для решения матричных уравнений, которые получаются в результате применения процедур вложения, возможно применение аппарата канонизации матриц [4].

Применение технологии вложения при синтезе по вынужденной составляющей дает следующее уравнение относительно искомой передаточной матрицы $K(p)$:

$$(5) \quad C(pI_n - A)^{-1} \sum_{i=0}^r B_i e^{-\tau_i p} + \sum_{i=0}^r D_i e^{-\tau_i p} - E_y^v(p) = E_y^v(p) K(p) (pI_n - A)^{-1} \sum_{i=0}^r B_i.$$

Используя результаты [2, 3] множество регуляторов в этом случае можно описать формулами:

$$(6) \quad \left\{ K(p) \right\}_{\mu, \eta} = (E_y^v(p) \tilde{\sim}) [C(pI_n - A)^{-1} \sum_{i=0}^r B_i e^{-\tau_i p} + \sum_{i=0}^r D_i e^{-\tau_i p} - E_y^v(p)] \times \\ \times ((pI_n - A)^{-1} \sum_{i=0}^r B_i \tilde{\sim} + E_y^v(p) \mu(p) + \eta(p) (pI_n - A)^{-1} \sum_{i=0}^r B_i),$$

где $\eta(p)$, $\mu(p)$ – произвольные дробно-полиномиальные матрицы соответствующих размеров. Условия разрешимости уравнения (5), а значит и существования множества решений (6) имеют вид:

$$(7) \quad \overline{E_y^v(p)}^L [C(pI_n - A)^{-1} \sum_{i=0}^r B_i e^{-\tau_i p} + \sum_{i=0}^r D_i e^{-\tau_i p} - E_y^v(p)] = 0, \\ [C(pI_n - A)^{-1} \sum_{i=0}^r B_i e^{-\tau_i p} + \sum_{i=0}^r D_i e^{-\tau_i p} - E_y^v(p)] ((pI_n - A)^{-1} \sum_{i=0}^r B_i) \overline{\quad}^R = 0,$$

или

$$[C(pI_n - A)^{-1} \sum_{i=0}^r B_i e^{-\tau_i p} + \sum_{i=0}^r D_i e^{-\tau_i p} - E_y^v(p)] = E_y^v(p) (E_y^v(p))^{-R} \chi(p),$$

$$(8) \quad [C(pI_n - A)^{-1} \sum_{i=0}^r B_i e^{-\tau_i p} + \sum_{i=0}^r D_i e^{-\tau_i p} - E_y^v(p)] \underbrace{((pI_n - A)^{-1} \sum_{i=0}^r B_i)}_R =$$

$$= \xi(p) ((pI_n - A)^{-1} \sum_{i=0}^r B_i)^L ((pI_n - A)^{-1} \sum_{i=0}^r B_i),$$

где $\xi(p)$, $\chi(p)$ – произвольные дробно-полиномиальные матрицы соответствующего размера.

Таким образом, получено соотношение (6) с условиями существования этого решения (7) или (8), которые позволяют найти множество регуляторов, удовлетворяющих закону управления (4).

4 Алгоритм синтеза САУ на основе компенсационно-наблюдательной схемы управления

Алгоритм проводится по вынужденной составляющей движения замкнутой динамической системы.

- 1) Определяются математическая модель объекта управления в виде дифференциально-разностных уравнений вида (1) и закон управления (4).
- 2) Формируется желаемая МПФ $E_y^v(p)$.
- 3) Осуществляется проверка условий существования решения матричного уравнения по формулам (7) или (8). Если условия разрешимости не выполняются, то переход на этап 2) и коррекция желаемой МПФ $E_y^v(p)$.
- 4) По формуле (6) вычисляется регулятор $K(p)$ с $\mu = 0$ и $\eta = 0$ при синтезе по вынужденной составляющей.
- 5) При необходимости возможно получение множества решений регулятора $K(p)$ по формуле (6) с последующим решением задачи оптимизации [5] на найденном множестве регуляторов и предкомпенсаторов.

5 Синтез системы автоматического управления газотурбинным двигателем самолета

На примере синтеза системы управления газотурбинным двигателем (ГТД) самолета покажем эффективность применения предложенного решения задачи синтеза на основе компенсационно-наблюдательной схемы по сравнению с классическим методом, использующим аппроксимацию Паде.

Математическая модель газотурбинного двигателя, описывающая один из режимов работы двигателя, представлена в литературе [7] в виде матричной передаточной функции:

$$(9) \quad W(p) = \begin{bmatrix} \frac{0.59}{0.45p + 1} & \frac{0.85}{0.45p + 1} \\ \frac{0.27(1.1p + 1)}{0.45p + 1} & \frac{0.29}{0.45p + 1} \end{bmatrix}.$$

Входной вектор $u(t)$ ГТД, описанного в литературе [6], включает в себя такие составляющие, как расход топлива G_T через основную камеру сгорания и площадь F_C реактивного сопла. Выходными величинами вектора $y(t)$ являются частота n вращения турбокомпрессора и температура T_3^* газов перед турбиной:

$$u(t) = \begin{bmatrix} G_T(t) \\ F_C(t) \end{bmatrix}, \quad y(t) = \begin{bmatrix} n(t) \\ T_3^*(t) \end{bmatrix}.$$

В литературе [7] отмечено, что в переходных режимах изменение температуры газа на выходе из камеры сгорания перед турбиной происходит с запаздыванием во времени относительно изменения соответствующего расхода топлива. Следовательно, изменение расхода топлива во времени можно записать как $G_T(t - \tau)$, где τ – время запаздывания.

Одним из возможных представлений математической модели ГТД (9) в пространстве состояний могут быть следующие матрицы:

$$A = \begin{bmatrix} -2.222 & 0 \\ 0 & -2.222 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 0.6556 & 0.9444 \\ -0.4333 & 0.3222 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 \\ 0.66 & 0 \end{bmatrix}.$$

Допустим, что время запаздывания при изменении температуры газа на выходе из камеры сгорания перед турбиной относительно изменения соответствующего расхода топлива будет составлять $\tau = 0.2$ с.

Пусть желаемая передаточная функция $E_y^v(p)$, с более высокой скоростью переходных процессов, чем у исходного объекта управления, будет иметь вид:

$$E_y^v(p) = \begin{bmatrix} \frac{1.31e^{-0.2p}}{(p+4.022)} & \frac{1.89}{(p+3.022)} \\ \frac{0.66(p+0.91)e^{-0.2p}}{(p+4.022)} & \frac{0.644}{(p+3.022)} \end{bmatrix}.$$

Осуществим проверку условий существования решения матричного уравнения по формулам

(7). Убедимся в их выполнении, найдя $\overline{E_y^v(p)}^L$ и $\overline{((pI_n - A)^{-1} \sum_{i=0}^r B_i)}^R$ тождественно равные

нулю. По формуле (6) найдем регулятор $K(p)$ при $\mu = 0$ и $\eta = 0$: $K(p) = \begin{bmatrix} 0.9 & 0 \\ 0 & 0.4 \end{bmatrix}$.

Переходные характеристики объекта управления, системы управления, полученной с помощью метода модального синтеза и замены звена запаздывания Паде аппроксимацией 2-го порядка, системы управления полученной с помощью компенсационно-наблюдательной схемы управления при $\tau = 0,2$ с показаны на рисунке 2.

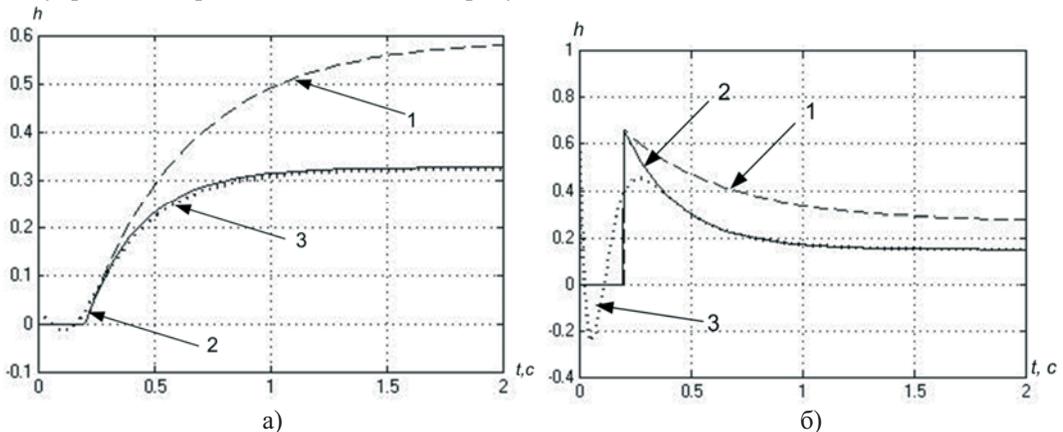


Рисунок 2 – Переходные характеристики исходного объекта управления (1), САУ, полученной на основе вышеприведенного метода (2), САУ, полученной на основе классического метода с аппроксимацией Паде 2-го порядка (3) по каналам 1-1 а) и 1-2 б)

Необходимо отметить ряд преимуществ рассмотренного решения задачи синтеза на основе компенсационно-наблюдательной схемы управления перед синтезом системы модальным методом с наблюдателем и представлением звена запаздывания Паде аппроксимацией:

- 1) Синтез системы проводится по желаемой матричной передаточной функции от входа к выходу (учитываются не только полюса, но и нули системы управления);
- 2) Нахождение МПФ регулятора является математически точным, без применения аппроксимационных методов;
- 3) Размер МПФ регулятора у рассмотренного решения задачи синтеза всегда равен $n \cdot s$ (n – количество переменных состояния объекта, s – количество входов объекта), а при синтезе модальным методом $(n + 2 \cdot q) \cdot s$ (q – количество времен запаздываний по входу объекта управления);
- 4) Соответственно размер матрицы состояния модели объекта при синтезе модальным методом с наблюдателем возрастает до $(n + 2 \cdot q) \cdot (n + 2 \cdot q)$;
- 5) При увеличении времени запаздывания в объекте управления точность синтеза модальным методом с наблюдателем и Паде аппроксимацией существенно падает.

Заключение

Был проведен синтез системы автоматического управления для многосвязного объекта с запаздываниями по управлению. На первом этапе решения задачи синтеза была предложена компенсационно-наблюдательная схема, позволяющая воспроизводить фазовый вектор состояния объекта без учета запаздываний. На втором этапе было найдено множество решений задачи синтеза и условия разрешимости задачи с помощью применения технологии вложения систем. Приведен алгоритм синтеза САУ на основе компенсационно-наблюдательной схемы управления. На примере синтеза системы управления ГТД была показана эффективность рассмотренного решения задачи перед синтезом системы модальным методом с наблюдателем и представлением звена запаздывания Паде аппроксимацией.

Благодарности

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант 08-08-00536).

Список литературы

- [1] Асанов А.З. Технология вложения систем и ее приложения: учебное пособие / А.З. Асанов; Уфимск. гос. авиац. техн. ун-т – Уфа: УГАТУ, 2007 – 227 с.
- [2] Буков В.Н. Вложение систем. Аналитический подход к анализу и синтезу матричных систем.– Калуга: Издательство научной литературы Н.Ф. Бочкаревой, 2006. – 720 с.
- [3] Кирьянен А.И. Устойчивость систем с последствием и их применения. СПб.: СПбГУ, 1994.-236 с.
- [4] Буков В. Н., Рябченко В. Н., Косьянчук В. В., Зябин Е. Ю. Решение линейных матричных уравнений методом канонизации // Вестник Киевского университета. Серия: Физико-математические науки.– Вып. 1.–2002.– С.19–28.
- [5] Буков В. Н., Рябченко В. Н., Зубов Н. Е. Вложение и оптимизация линейных систем // Автоматика и телемеханика. – 2002. –№ 5.– С. 12–23.
- [6] Основы теории многосвязных систем автоматического управления летательными аппаратами: Учеб. Пособие / С.Ф. Бабак, В.И. Васильев, Б.Г. Ильясов и др.; Под ред. М.Н. Красильщикова. – М.: Изд-во МАИ, 1995. – 288 с.
- [7] Черкасов Б.А. Автоматика и регулирование воздушно-реактивных двигателей: Учебник для вузов по специальности «Авиационные двигатели». 3-е изд., перераб. и доп.– М.: Машиностроение, 1988.– 360 с.

МЕТОДИКА КОРРЕКТИРОВКИ УЧЕБНОГО ПЛАНА ПРОФЕССИОНАЛЬНОЙ ПОДГОТОВКИ С УЧЕТОМ ТРЕБОВАНИЙ ЗАИНТЕРЕСОВАННЫХ СТОРОН

А.З. Асанов, И.Ю. Мышкина

Казанский государственный университет
423812, Набережные Челны, пр. Суюмбике, 10А, Россия
mirinau@mail.ru
тел: +7 (8552) 38-34-05, факс: +7 (8552) 58-88-94

Ключевые слова: нечеткая логика, профессиональные знания, учебный план

Abstract

In this article the method of adjusting the professional training curriculum focusing on the stakeholders' requirements is considered. The method is based on fuzzy calculations and a method of the analysis of hierarchies.

Введение

Важной составляющей менеджмента качества высшего образовательного учреждения является учет мнений всех заинтересованных лиц в организации обучения – государства (в лице министерства образования и науки), работодателей, персонала и администрации самого вуза и т.д. Профессиональные знания выпускников являются основным продуктом деятельности вуза. Главным инструментом, позволяющим регулировать качество этого «продукта», является учебный план специальностей, который требует периодической корректировки для обеспечения соответствия приобретаемых студентами знаний требованиям потребителей процесса обучения.

В работе предлагается алгоритм принятия решений по корректировке учебного плана, основанный на нечеткой логике и учитывающий требования заинтересованных сторон, а также обеспеченность вуза необходимыми ресурсами.

В работе рассматривается одна из компонент компетенций специалиста – профессиональные знания. Основными методами получения количественных оценок в данном случае являются рейтинг учебных дисциплин, экспертные оценки, результаты тестирования. Важность рассматриваемой компоненты компетенций специалиста для управления системой качества вуза связана с тем, что ее формирование обусловлено, в первую очередь, организацией процесса обучения в вузе, тогда как формирование других компетенций (социальных и т.п.) начинается в значительно более раннем возрасте. Кроме того, оценка профессиональных знаний выпускника напрямую определяет качество обучения в вузе и является важной составляющей при принятии решений о внесении возможных изменений в учебные планы вуза. Существующие работы, посвященные оценке компетенций выпускников, направлены прежде всего на принятие решений об обновлении учебных планов вуза типа «добавить дисциплину» или «увеличить изучение дисциплины в 1,3 раза», при этом не определяются количественные оценки компетенций конкретного выпускника и всего выпуска вуза, не описываются конкретные алгоритмы корректировки учебных планов, не учитываются мнения всех потребителей вуза – работодателей и выпускников, не учитываются ограничения в организации обучения в вузе – требования ГОС ВПО.

1 Постановка задачи

Информационная модель учебного плана специальности представлена на рисунке 1. Неопределенность в проблеме корректировки учебного плана, требующей оценки соответствия учебного плана специальности внешним требованиям групп заинтересованных лиц, заключается в невозможности точно сформулировать требования к структуре учебного плана и взаимосвязям его основных элементов типа «высокая значимость дисциплины» и т.п. Как правило, возможно определение только границы значений оцениваемой характеристики. Для оценки определяемых экспертно характеристик, отраженных в модели, предлагается использовать соответствующие нечеткие числа, определяемые гауссовой функцией принадлежности (В и А – параметры гауссовой функции принадлежности).



Рисунок 1 – Информационная модель учебного плана специальности

Профессиональные знания (ПЗ) можно представить в виде: $ZN = ZN_{UchP} \cup ZN_{\overline{UchP}}$, ZN_{UchP} – профессиональные знания, входящие в учебный план вуза $ZN_{UchP} = \langle T, D, R, H_D \rangle$, где T – тип знания (теоретическое или практическое), D – дисциплина, в рамках которой приобретается знание, R – раздел дисциплины, в рамках которого приобретается знание, H_D – число часов по дисциплине согласно учебному плану. $ZN_{\overline{UchP}}$ – профессиональные знания, не входящие в учебный план вуза $ZN_{\overline{UchP}} = \langle T, D, R \rangle$.

В работе приводится методика решения рассматриваемой задачи в следующей постановке. Разработать алгоритм принятия решений по корректировке процесса обучения в вузе по некоторой специальности на основе изменения учебного плана, который должен учитывать требования студентов, работодателей, профессиональных стандартов и ГОС ВПО, и приводить к корректировке объемов часов, отводимых на дисциплину, в рамках которой приобретаются профессиональные знания, входящие в профессиональные компетенции специалиста.

Перечень всех профессиональных знаний (ПЗ) достаточно большой, для анализа следует рассматривать ПЗ, которые чаще всего работодатели рассматривают в качестве требований к специалистам.

2 Определение значимости профессиональных знаний

Функциональная модель корректировки учебного плана представлена на рисунке 2.

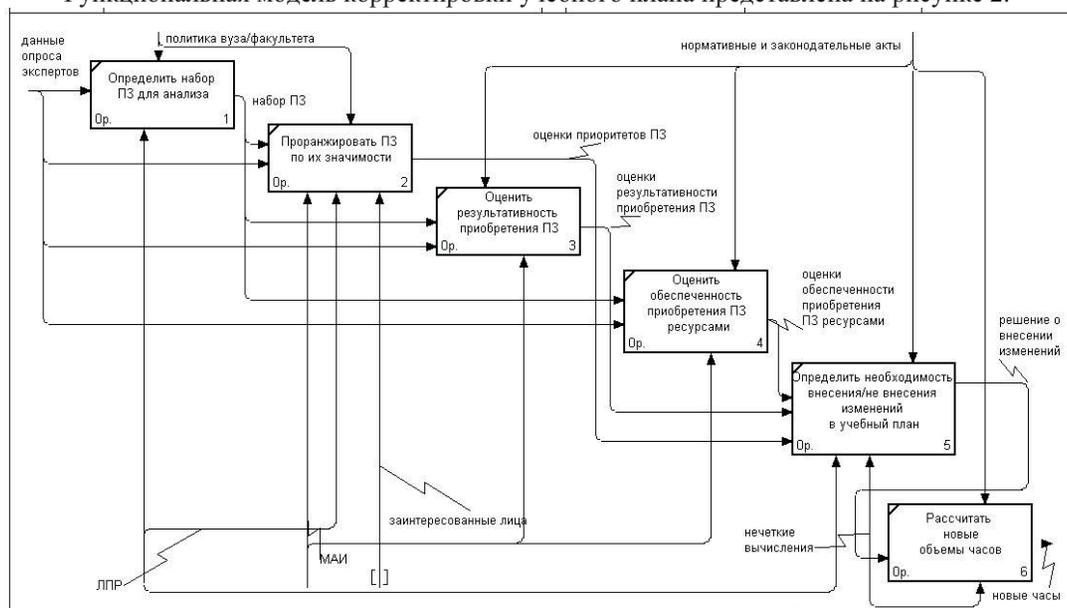


Рисунок 2 – Функциональная модель корректировки учебного плана

На первом этапе для того, чтобы определить наиболее важные/значимые ПЗ для дальнейшего принятия решения о корректировке учебного плана, следует решить задачу многокритериального выбора альтернатив – ПЗ. Причем необходимо не просто провести ранжирование ПЗ по их значимости, отражающее необходимость их приобретения как для изучения всех дисциплин, предусмотренных учебным планом (речь идет о так называемой внутренней значимости), так и для будущей профессиональной деятельности (т.н. внешняя значимость [1]), но и получить некоторую количественную оценку этой значимости, которую в дальнейшем можно использовать для принятия окончательного решения.

Для получения достоверной комплексной оценки значимости знаний должны учитываться мнения и требования всех заинтересованных сторон в организации процесса обучения, что соответствует требованиям к организации системы менеджмента качества любой организации согласно стандартам ISO серии 9001. Согласно ISO 9001-2001, заинтересованная сторона: лицо или группа, заинтересованные в деятельности или успехе организации. Заинтересованные стороны можно разделить на пять групп [2]: потребители, акционеры («управление»), общество («окружающая среда»), поставщики и персонал организации. Для вуза это:

- 1) Потребители – абитуриенты, родители (попечители), предприятия (организации).
- 2) «Управление» – Министерство образования и науки РФ.
- 3) Общество – администрации субъектов Федерации, советы ректоров вузов регионов, службы занятости, средства массовой информации и другие.
- 4) Поставщики – учреждения среднего и средне-профессионального образования.
- 5) Персонал – сотрудники вуза.

Примерный перечень заинтересованных лиц с точки зрения рассматриваемой задачи для некоторой специальности представлен в таблице 1.

Таблица 1 – Заинтересованные лица

Кластер «Заинтересованные лица»	Группа (согласно 5 групп, указанных в [2])	Элементы кластера	Интересы/цели	Критерии оценок
Государство (министерство образования и науки)	«Управление»	Государственный образовательный стандарт (ГОС) по специальности	Соответствие учебного плана и рабочих программ требованиям ГОС	Место дисциплины, включающей ПЗ, в ГОС (включена/не включена в ГОС; цикл дисциплин, обязательная/вариативная часть, число часов (если включена в ГОС))
		Спрос на специальность	Обеспечение государства квалифицированными специалистами, соответствующими потребностям экономики	Рейтинг специальности, прогнозируемые характеристики спроса на специальность
		Профессиональные стандарты	Подготовка специалиста, отвечающего требованиям будущей профессиональной деятельности	Место ПЗ в профессиональных стандартах (включена/не включена, число включений ПЗ в стандарты по различным профессиям, отвечающим исследуемой специальности и уровню высшего профессионального образования)
Предприятия (производство)	Потребители	Работодатель	Подготовка специалиста, отвечающего требованиям и потребностям конкретного производства	Анализ соответствия требуемых ПЗ и учебных планов, уровень знаний в требуемой области, оценка работы сотрудников-выпускников рассматриваемого вуза
Администрация вуза	Персонал	Учебно-методическое объединение факультета	Формирование учебных планов, соответствующих стратегии развития вуза и факультета, обеспечение его соответствия требованиям ГОС и имеющимся ресурсам, реализация основных принципов обучения: преемственность, системность и т.д.	Экспертная оценка по нескольким критериям
		Научный отдел вуза	Формирование учебных планов, соответствующих научным направлениям вуза, вовлечение студентов в научную работу	Экспертная оценка по нескольким критериям
		Профессорско-преподавательский состав	Формирование учебных программ, обеспечение ее соответствия требованиям ГОС и имеющимся ресурсам	Экспертная оценка по нескольким критериям

Необходимость учета требований всех заинтересованных сторон значительно затрудняет экспертные оценки. В данном случае целесообразно ранжирование значимостей ПЗ проводить по двум уровням: во-первых, оценка экспертами вуза важности требований каждой заинтересованной стороны при оценке ПЗ, во-вторых, оценка значимости ПЗ для каждой заинтересованной стороны экспертами вуза и внешними экспертами (лицами, занимающимися подбором персонала на предприятиях — потенциальных работодателях). С учетом этого задачу определения значимости ПЗ можно представить в виде иерархии, представленной на рисунке 3. Для решения задач, подобных сформулированной, хорошо зарекомендовал себя метод анализа иерархий (МАИ), обоснованный математически [3] и тесно связанный с особенностями организации мышления и памяти человека. Одним из основных преимуществ данного метода является то, что иерархическое представление задачи принятия решений позволяет описывать влияние элементов иерархии одного уровня на элементы другого уровня.



Рисунок 3 – Иерархия для определения значимости ПЗ

Анализ элементов 2 уровня иерархии – экспертов – производится лицом принимающим решение (ЛПР). Рассчитанные оценки приоритетов для каждого эксперта будут означать значимость каждого эксперта для окончательного принятия решения. Оценки значимости каждого эксперта учитываются при вычислении оценок приоритетов элементов на каждом уровне иерархии.

Приоритеты элементов 3 уровня иерархии – кластеры заинтересованных лиц: Государство, Предприятия и Администрация вуза – вычисляются каждым экспертом. С учетом этих оценок рассчитываются приоритеты элементов кластеров (например, при наличии нескольких работодателей требуется учесть их мнения, которые могут значительно отличаться).

Следует отметить, что если число рассматриваемых ПЗ больше 9, то для вычисления приоритетов следует их объединить в кластеры (группы) [3], что упрощает работу экспертов по составлению матриц парных сравнений и в случае, если добавляется новая альтернатива ПЗ.

В случае если ПЗ входит в учебный план (ПЗ входит в ZN_{UchP}), то на решение об изменении учебного плана влияет результативность изучения соответствующих дисциплин, куда входят результаты обучения по дисциплинам (средний рейтинг изучения дисциплины), мн-

ния работодателей об уровне знаний уже работающих выпускников на предприятии и удовлетворенность студентов уровнем обучения в вузе. Анализ результативности может быть проведен, исходя из анализа соответствующей иерархии, представленной на рисунке 4.

В случае если ПЗ не входит в учебный план (ПЗ входит в ZN_{UchP}), то на решение об изменении учебного плана (внесение изучения ПЗ в рабочую программу дисциплины или внесение в учебный план новой дисциплины) влияет обеспеченность ресурсами (профессорско-преподавательский состав и материально-техническая база), необходимыми для изучения соответствующей дисциплины. Данный анализ может быть проведен, исходя из анализа иерархии, представленной на рисунке 5.

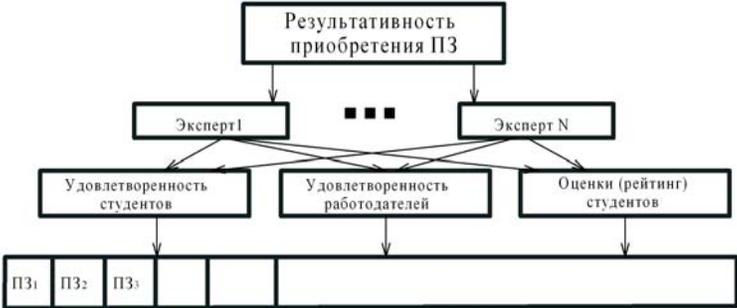


Рисунок 4 – Иерархия для определения результативности приобретения ПЗ



Рисунок 5 – Иерархия для определения обеспеченности приобретения ПЗ ресурсами

Следующим этапом принятия решений (после получения оценок значимости ПЗ (З), оценок обеспеченности ресурсами (ОР) приобретения ПЗ и оценок результативности (Р) приобретения ПЗ) является построение решающих правил для принятия решений по корректировке учебных планов. Полученные числовые оценки следует спроецировать на лингвистическую шкалу «низкий», «средний» и «высокий» уровень. Таким образом, характеристики З, ОР и Р следует рассматривать как лингвистические переменные с тремя термами «низкий», «средний» и «высокий», задаваемые, например, треугольной функцией принадлежности (треугольную функцию принадлежности можно задать тройкой чисел (a, b, c) , причем значение параметра b – значение оценки характеристики, которое, по мнению эксперта, более достоверно, а a и c – полной уверенности эксперта, что данные оценки не соответствуют действительности.). Для получения лингвистической оценки некоторой характеристики выбирается терм, функция принадлежности которого для соответствующей числовой оценки принимает наибольшее значение.

Базы правил для ПЗ, входящих в ZN_{UchP} и в $ZN_{\overline{UchP}}$ различны, так как имеют различные входные переменные (З, ОР и Р). Возможные правила имеют вид:

ЕСЛИ Значимость «низкая» И Обеспеченность ресурсами «низкая» И «ПЗ не входит в ГОС» ТО «Не вносить изменения в учебный план»

3 Методический пример

Рассмотрим применение описанного алгоритма в случае четырех ПЗ: ПЗ₁ — входит в ГОС, $ПЗ_1 \in ZN_{UchP}$; ПЗ₂ — не входит в ГОС, $ПЗ_2 \in ZN_{\overline{UchP}}$; ПЗ₃ — входит в ГОС, $ПЗ_3 \in ZN_{UchP}$, ПЗ₄ — входит в ГОС, $ПЗ_4 \in ZN_{\overline{UchP}}$. Пусть в оценках на всех уровнях участвуют по одному эксперту.

Экспертом построена матрица парных сравнений элементов 3 уровня иерархии, представленной на рис. 2. Т.е. попарно сравниваются значимости учета мнений Администрации вуза, Предприятий и Государства для определения значимости ПЗ:

Значимость ПЗ	АВ	П	Г
Администрация вуза (АВ)	1	1/4	1/9
Предприятия (П)	4	1	1/9
Государство (Г)	9	9	1

Так, например, в данном случае принято, что Государство имеет абсолютную значимость перед остальными элементами – соответствующие значения в матрице равны 9 (см. таблицу 2).

Определен собственный вектор полученной матрицы, соответствующий ее максимальному собственному значению, называемый вектором приоритетов, при этом сумма компонент этого вектора равна единице: $w_{\max}^1 = (W^{AB}, W^P, W^G) = (0.06, 0.14, 0.80)$

Аналогично составляются матрицы парных сравнений элементов на всех уровнях по силе их влияния на предыдущий уровень и вычисляются собственные векторы.

Оценка приоритетов значимостей ПЗ рассчитана по формуле:

$$z(ПЗ_i) = W^P \cdot w_{\max}^P(ПЗ_i) + W^{HO} \cdot w_{\max}^{HO}(ПЗ_i) + W^{YMO} \cdot w_{\max}^{YMO}(ПЗ_i) + W^{ППС} \cdot w_{\max}^{ППС}(ПЗ_i) + W^{GOC} \cdot w_{\max}^{GOC}(ПЗ_i) + W^{CC} \cdot w_{\max}^{CC}(ПЗ_i) + W^{PC} \cdot w_{\max}^{PC}(ПЗ_i),$$

где $i = \overline{1,3}$.

$$z(ПЗ_1) = 0.40, \quad z(ПЗ_2) = 0.15, \quad z(ПЗ_3) = 0.15, \quad z(ПЗ_4) = 0.30.$$

Таким образом, наибольшую значимость имеет ПЗ₁.

Аналогично рассчитываются результативность (r) и обеспеченность ресурсами (or) приобретения ПЗ. Пусть $r(ПЗ_1) = 0,2$, $r(ПЗ_3) = 0,8$, $or(ПЗ_2) = 0,5$, $or(ПЗ_4) = 0,5$.

Пусть параметры треугольных функций принадлежности термов лингвистических переменных значимость, результативность и обеспеченность ресурсами следующие:

	«низкий»	«средний»	«высокий»
значимость	(0,0,0.5)	(0.25,0.5,0.75)	(0.5,1,1)
результативность	(0,0,0.5)	(0.25,0.5,0.75)	(0.5,1,1)
обеспеченность ресурсами	(0,0,0.5)	(0.25,0.5,0.75)	(0.5,1,1)

В итоге получены следующие лингвистические оценки ПЗ:

	ПЗ ₁	ПЗ ₂	ПЗ ₃	ПЗ ₄
значимость	средняя	низкая	низкая	средняя
результативность	низкая	—	высокая	—
обеспеченность ресурсами	—	средняя	—	средняя

Пусть фрагмент базы правил для ПЗ, входящих в ZN_{UchP} следующий:

ЕСЛИ Значимость «средняя» И Обеспеченность ресурсами «средняя» И «ПЗ входит в ГОС» ТО «Включить ПЗ в учебный план»

ЕСЛИ Значимость «низкая» И Обеспеченность ресурсами «низкая» И «ПЗ не входит в ГОС» ТО «Не вносить изменения в учебный план»

ЕСЛИ Значимость «низкая» И Обеспеченность ресурсами «средняя» И «ПЗ не входит в ГОС» ТО «Не вносить изменения в учебный план»

Пусть фрагмент базы правил для ПЗ, входящих в ZN_{UchP} следующий:

ЕСЛИ Значимость «низкая» И Результативность «высокая» И «ПЗ входит в ГОС» ТО «Не вносить изменения в учебный план»

ЕСЛИ Значимость «средняя» И Результативность «средняя» И «ПЗ входит в ГОС» ТО «Не вносить изменения в учебный план»

ЕСЛИ Значимость «средняя» И Результативность «низкая» И «ПЗ входит в ГОС» ТО «Увеличить время изучения ПЗ в учебном плане»

Вывод – 1) по дисциплине, в рамках которой приобретается ПЗ₂, следует увеличить время изучения, 2) ПЗ₄ следует включить в рабочую программу соответствующей дисциплины, 3) в дисциплины, в рамках которых приобретаются ПЗ₁, ПЗ₃, изменения вносить не следует.

Расчет скорректированного, нового объема часов может быть проведен с помощью базы нечетких правил с входными переменными: значимость дисциплины, в рамках которой приобретается ПЗ, объем часов по ГОС, фактический объем часов, объем тезауруса.

Заключение

Среди достоинств рассмотренного в работе подхода можно отметить следующие:

- 1) Учет мнений всех заинтересованных сторон в организации образовательного процесса в вузе.
- 2) Получение числовых оценок различных характеристик ПЗ.
- 3) Возможность на основе полученных количественных оценок регулировать объемы изменений учебного плана вуза с помощью, например, нечеткого логического вывода.

Дальнейшая работа по данной тематике видится в адаптации предложенного подхода к организации процесса обучения по конкретной специальности.

Благодарности

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (грант 08-08-00536).

Список литературы

- [1] Трофимова О.К. Автоматизация процесса составления учебных планов вузов : автореферат дис. канд. тех. наук.- Москва: МГУПП, 1999.
- [2] Управление качеством образования: монография / Т.П.Костюкова и др.; Уфимск. гос. авиац. техн. ун-т. – Уфа: УГАТУ, 2008.
- [3] Саати Т. Л. Принятие решений. Метод анализа иерархий. М.: Радио и связь, 1993.

СИНТЕЗ ЛИНЕЙНЫХ СИСТЕМ УПРАВЛЕНИЯ С ИСПОЛЬЗОВАНИЕМ ЭТАЛОННОЙ МОДЕЛИ

Р.Н. Курманалиева

Кыргызский государственный университет строительства,
транспорта и архитектуры им. И. Исанова
720020, Бишкек, ул. Малдыбаева, 346, Кыргызская Республика
nasbekova@mail.ru

Ключевые слова: управление, эталонная модель, структурный синтез, быстродействие и точность

Abstract

There is problem of control multivariate dynamic with object in this article. As quality of basic character system control use the direct parameters of quality. Method of synthesis of automatic system control, ensure it's desire dynamic properties definite etalon model offered in this article.

Введение

При проектировании систем автоматического управления (САУ) широкое применение находят эталонные модели [1,2,3], определяющие желаемое поведение управляемого объекта в переходных процессах. Построение таких моделей часто базируется на оптимизации квадратических оценок качества управления [1,4]. В данной статье предлагается метод синтеза регуляторов САУ (рисунок 1), основанный на использовании эталонных процессов, удовлетворяющих первичным инженерным требованиям к проектируемой системе, таким как время регулирования и перерегулирование. Последнее, как известно, определяют быстродействие и точность САУ.

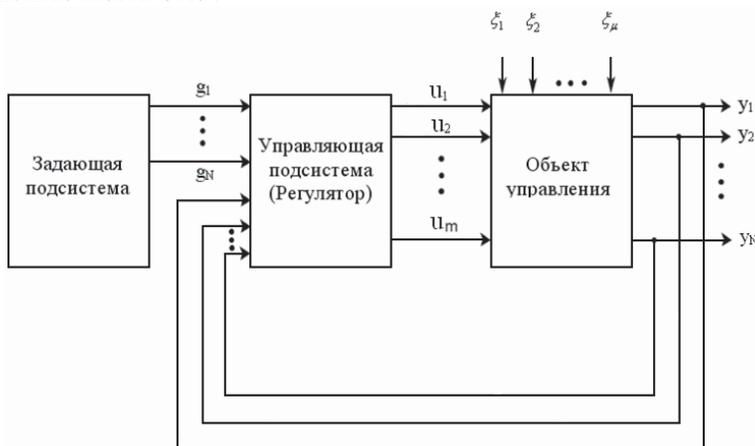


Рисунок 1 – Структура системы автоматического управления

Постановка задачи. Рассмотрим многомерный объект управления, описываемый векторным уравнением в отклонениях:

$$(1) \quad \dot{x}(t) = Ax(t) + Bu(t) + \xi(t), \quad x(t_0) = x^0, \quad t \in [t_0, t_k],$$

где $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$ – n -мерный вектор состояния объекта;

$u(t) = [u_1(t), u_2(t), \dots, u_m(t)]^T$ – m -мерный вектор управления;

$\xi(t) = [\xi_1(t), \xi_2(t), \dots, \xi_r(t)]^T$ – r -мерный вектор внешних возмущений;

A и B – вещественные матрицы: $A = \{a_{ij}\}_{n \times n}$, $B = \{b_{il}\}_{n \times m}$.

Пусть объект обладает свойством управляемости /4/, а компоненты векторов состояния $x(t)$ и возмущений $\xi(t)$ доступны для измерения. Вектор управляемых переменных (выход объекта) $y = [y_1, y_2, \dots, y_N]^T$ определяется соотношением

$$(2) \quad y(t) = Cx(t),$$

где C – $N \times n$ -мерная вещественная матрица. Далее предполагается, что $N \leq m$.

Целью управления является поддержание выхода объекта $y(t)$ в соответствии с заданной вектор-функцией $g(t) = [g_1(t), g_2(t), \dots, g_N(t)]^T$, которая представляет собой желаемую траекторию движения системы. Функцию управления выполняет *регулятор*, который на основе вычисления вектора ошибки управления

$$e_i(t) = g_i(t) - y_i(t), \quad i = \overline{1, N}.$$

формирует управляющие воздействия $u(t)$ на объект.

Задача синтеза заключается в *определении структуры и параметров регулятора*, обеспечивающего гарантированное выполнение условий требуемого качества переходных процессов. Пусть эти условия заданы проектировщиком системы управления в виде следующих критериальных ограничений (рисунок 2):

$$(3) \quad |e_i(t)| \leq \bar{b}_i(t), \quad t \in [t_0, t_k], i = \overline{1, N},$$

где $\bar{b}_i(t)$ – положительные непрерывно дифференцируемые функции, задающие границы соответствующих допустимых множеств

$$(4) \quad E_i(t) = \left\{ e_i \in R^1 : |e_i(t)| \leq \bar{b}_i(t) \right\}, \quad i = \overline{1, N}.$$

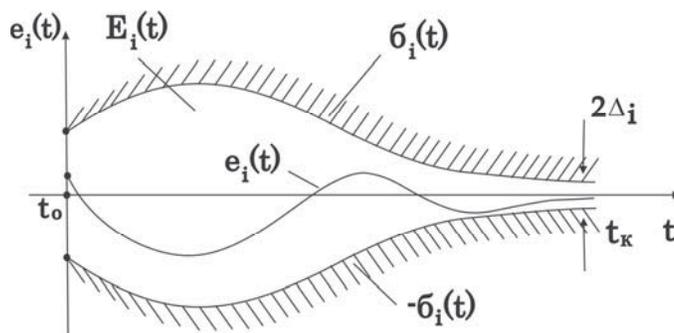


Рисунок 2 – Критерий допустимого качества управления

Анализ математического описания управляемого объекта (1) показывает, что при синтезе САУ можно выделить следующие случаи/5/:

- 1) все уравнения динамики объекта содержат компоненты вектора управления $u(t)$ /6/;
- 2) некоторые уравнения объекта **не содержат ни одну** компоненту $u(t)$.

Рассмотрим особенности решения задачи структурного синтеза для второго случая. Анализ показывает, что в этом случае эталонную модель целесообразно задавать в виде

$$(5) \quad \tilde{e}(t) = M\tilde{e}(t) + \tilde{\varepsilon}(t),$$

где N -мерная вектор-функция $\tilde{\varepsilon}(t) = [\tilde{\varepsilon}_1(t), \tilde{\varepsilon}_2(t), \dots, \tilde{\varepsilon}_N(t)]^T$.

Компоненты $\tilde{\varepsilon}(t)$ задаются по следующему правилу:

$$(6) \quad \tilde{\varepsilon}_v(t) = \begin{cases} 0, & \text{если } v\text{-ое уравнение объекта содержит хотя бы одну компоненту } u(t); \\ \varepsilon_v(t), & \text{в противном случае.} \end{cases}$$

При этом функции $\varepsilon_v(t)$ должны удовлетворять условию

$$(7) \quad |\varepsilon_v(t)| \leq \varepsilon_v^*(t), \quad \text{для всех } v \in I_1,$$

где $\varepsilon_v^*(t)$ – положительная непрерывно дифференцируемая монотонно убывающая функция; I_1 – множество тех индексов v , которые соответствуют номерам уравнений объекта, не включающих компонентов вектора управления $u(t)$.

На основе утверждения /5/, как и в первом случае, для определения элементов матрицы M , обеспечивающих $\tilde{\varepsilon}_i(t) \in \tilde{E}_i(t)$, $i = \overline{1, N}$, получаем следующие соотношения:

$$(8) \quad \sum_{\substack{j=1 \\ j \neq \mu}}^n |m_{\mu j}| \tilde{\delta}_j(t) \leq \tilde{\delta}_\mu(t) - m_{\mu\mu} \tilde{\delta}_\mu(t), \quad \mu \in I_0,$$

$$(9) \quad \sum_{\substack{j=1 \\ j \neq v}}^n |m_{vj}| \tilde{\delta}_v(t) + \varepsilon_v^*(t) \leq \tilde{\delta}_v(t) - m_{vv} \tilde{\delta}_v(t), \quad v \in I_1,$$

где I_0 – множество тех индексов μ , которые соответствуют номерам уравнений объекта, включающих компоненты вектора $u(t)$.

Уравнения синтеза, соответствующие управляемым переменным $Y_\mu(t)$, $\mu \in I_0$, формируются аналогично первому случаю /5,6/. Для получения остальных уравнений синтеза вначале необходимо построить функции $\varepsilon_v(t)$, который имеет следующий вид:

$$(10) \quad \varepsilon_v(t) = \sum_{j=1}^n \tilde{a}_{vj} x_j(t) + \tilde{\xi}_v(t) - \sum_{j=1}^N m_{vj} \tilde{e}_j(t), \quad v \in I_1,$$

где \tilde{a}_{vj} , $\tilde{\xi}_v(t)$ – элементы матрицы \tilde{A} и вектора $\tilde{\xi}(t)$.

Анализ показывает, что с целью упрощения решения задачи синтеза регулятора с учетом условий (7) закон изменения функций $\varepsilon_v(t)$, в частности, можно задать в следующем виде:

$$(11) \quad \dot{\varepsilon}_v(t) = \gamma_v \varepsilon_v(t), \quad v \in I_1,$$

где параметры γ_v определяются из условия выполнения ограничений (7). Нетрудно видеть, что использование утверждения приводит к неравенствам:

$$(12) \quad \gamma_v \varepsilon_v^*(t) \leq \dot{\varepsilon}_v^*(t), \quad v \in I_1.$$

Если принять, что

$$\dot{\varepsilon}_v^*(t) = \alpha_v \varepsilon_v^*(t), \quad \alpha_v < 0,$$

то параметры γ_v должны удовлетворять неравенствам

$$(13) \gamma_\nu \leq \alpha_\nu, \quad \nu \in I_1.$$

Производные $\dot{\epsilon}_\nu(t)$ определяются по формулам (10):

$$(14) \dot{\epsilon}_\nu(t) = \sum_{j=1}^n \tilde{\alpha}_{\nu j} \dot{x}_j(t) + \dot{\xi}_\nu(t) - \sum_{j=1}^N m_{\nu j} \dot{\tilde{e}}_j(t), \quad \nu \in I_1.$$

В результате с учетом выражений (10), (14) и динамики объекта (4) на основе соотношений (8) получаем оставшуюся часть уравнений синтеза искомого закона управления $u(t)$.

Алгоритм структурного синтеза регулятора включает следующие основные этапы:

Шаг 1. Задание модели объекта управления в форме векторного уравнения (1).

Шаг 2. Задание критериальных ограничений (3) и граничных функций $\tilde{b}_i(t)$, $i = \overline{1, N}$.

Шаг 3. Определение матрицы M эталонной модели, т.е. построение желаемого процесса $z(t)/6/$.

Шаг 4. Формирование уравнений синтеза с учетом рассмотренных выше случаев /5,6/.

Шаг 5. Определение вектора управления $u(t)$ как решение уравнений синтеза.

Таким образом, на основе принципа гарантируемой динамики предложен алгоритм структурного синтеза автоматических систем с использованием эталонной модели.

Список литературы

- [1] Бойчук Л.М. Метод структурного синтеза нелинейных систем автоматического управления. – М.: Энергия, 1971. – 112 с.
- [2] Тимофеев А.В. Построение адаптивных систем управления программным движением. – Л.: Энергия, 1988. – 88 с.
- [3] Смольников Л.П., Бычков Ю.А., Гудкова Н.В. Расчет систем управления. – Л.: Энергия, 1979. – 112 с.
- [4] Квакернаак Х., Сиван Р. Линейные оптимальные системы управления. – М.: Мир, 1977. – 650 с.
- [5] Оморов Т.Т., Курманалиева Р.Н. Многокритериальный синтез систем управления по показателям качества и сложности – Бишкек: Илим, 2007. – 136 с.
- [6] Курманалиева Р.Н., Оморов Т.Т. Синтез алгоритмов управления линейными многомерными системами. – Вестник КОМАЭ им.А.Эйнштейна. 2006.№2. –С.109-114.

МЕТОД ОЦЕНКИ НАДЕЖНОСТИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ПО ИНФОРМАЦИИ, СОХРАНЯЕМОЙ СИСТЕМОЙ УПРАВЛЕНИЯ ВЕРСИЯМИ

М.Н. Филиппов

Институт автоматизации и электрометрии СО РАН
630090, Новосибирск, пр-т Академика Коптюга, 1, Россия
michael@idisys.iae.nsk.su
тел: +7 (383) 333-26-25

***Ключевые слова:** надежность программного обеспечения, система управления версиями, остаточное количество ошибок, средняя наработка до отказа*

Abstract

A method of the software reliability estimation based on using data, automatically stored by a version control system, is presented. The number of remaining errors, mean time to failure and also time on the software testing and fixing defects can be estimated.

Введение

Известен ряд методов оценки надежности программного обеспечения (модели Холстеда, Нельсона, Джелински–Моранды и др.), учитывающих широкий спектр различных параметров, таких как размер исходных текстов и используемые языки программирования, количество модулей и связей между ними, интенсивность выявления ошибок в процессе тестирования, эффективность организации управления проектом и уровень квалификации сотрудников и т.п., но при этом в большинстве моделей не уделяется должного внимания существенным факторам, играющим свою роль во время создания данного конкретного продукта, в первую очередь, это полнота предоставленных спецификаций и частота появления новых требований в процессе реализации и тестирования, тип выявленных ошибок и потребовавшееся на исправление время, скорость внесения изменений в текст программы и прочие характеристики индивидуальной истории разработки

С другой стороны, создание масштабного программного продукта коллективом разработчиков сегодня практически невозможно представить без использования современных систем управления версиями (Version Control Systems или Revision Control Systems, таких как Subversion, Git, Mercurial, Bazaar и др.), существенно облегчающих работу с изменяющейся информацией, автоматизируя хранение исходных текстов и сопутствующей документации.

Значительная часть необходимой информации об индивидуальной истории создания конкретного программного продукта автоматически сохраняется в журнале используемой системы управления версиями, который в рамках разработанного метода и предлагается использовать в качестве источника данных для анализа и численной оценки основных характеристик надежности программного обеспечения.

1 Математический аппарат и пример практического применения

Разработанный метод оценки основных характеристик надежности программного обеспечения заключается в выявлении зависимости количества значимых изменений кода от фактического времени разработки по предоставляемой журналом системы управления версиями информации и экстраполяции полученной функции с целью получения числовых оценок остаточного количества ошибок и скорости их обнаружения.

В качестве иллюстрации представлен пример практического применения предложенного метода для оценки характеристик программного обеспечения системы управления движением поездов станции «Заяльцовская» Новосибирского метрополитена. При разработке программного обеспечения использована хорошо зарекомендовавшая себя открытая централизованная система управления версиями Subversion.

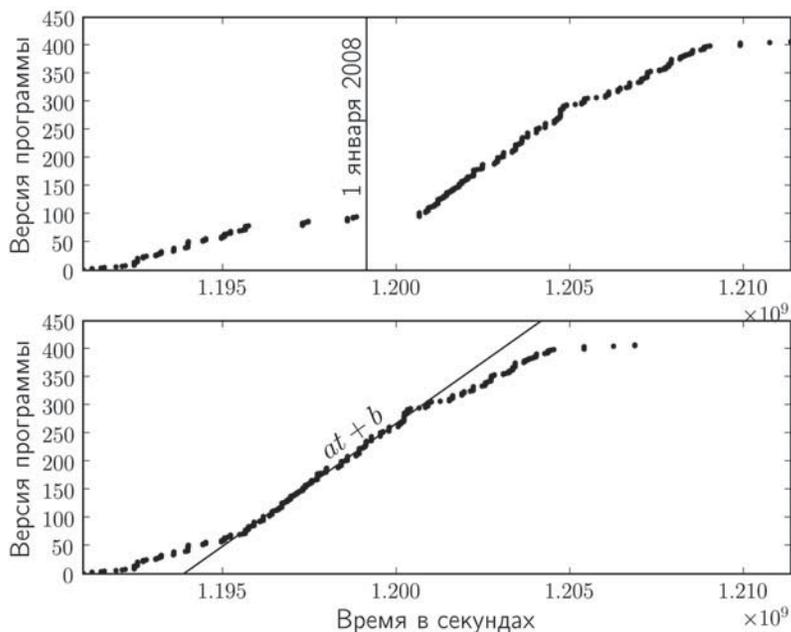


Рисунок 1 – Версия программы от времени разработки

На рисунке 1 приведён построенный по записям журнала системы управления версиями график времени выпуска версии программы, начиная с момента получения предварительных спецификаций 28 сентября 2007 г. и заканчивая установкой программы на станции в режим опытной эксплуатации 21 мая 2008 г.; всего за этот период в исходный текст программы было внесено 407 изменений. В нижней части рисунка для удобства дальнейшего рассмотрения исключены наиболее заметные интервалы времени, в течение которых разработка программы была приостановлена по различным причинам, таким как участие сотрудников в других проектах и конференциях, праздничные дни, дни отпуска и т.п.

При анализе данного графика целесообразно выделить три основные части: первая соответствует начальному периоду разработки, во время которого проводится уточнение спецификаций, планируется внутренняя архитектура программного обеспечения и создаются прототипы основных модулей. На данном этапе скорость внесения изменений относительно невелика, затем она постепенно увеличивается и достигает максимального значения к средней части графика.

В течение следующего периода, который соответствует времени активной разработки, кривая подчиняется простой линейной зависимости $(at + b)$. Объясняется это тем, что, несмотря на готовый план работы и четкие спецификации программных интерфейсов, скорость внесения изменений ограничена производительностью труда коллектива разработчиков.

Наконец, последняя часть графика вновь демонстрирует уменьшение скорости внесения изменений и отражает плавный переход от активной разработки к процессу тестирования, описываемому принципиально другой зависимостью (рисунок 2), которая и представляет наи-

большой интерес с точки зрения оценки основных характеристик надежности поступающего в эксплуатацию программного обеспечения.

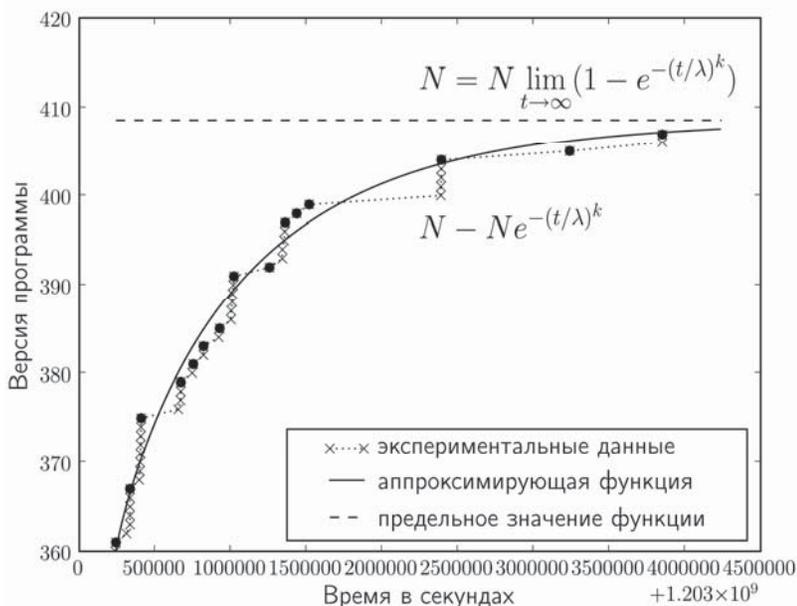


Рисунок 2 – Версия программы от времени тестирования

На представленном графике видно, что изменения обычно вносятся группами по 3–5 исправлений подряд. Отчасти это связано с особенностями работы в системе Subversion, но в основном объясняется тем, что требования об изменениях, предоставляются по окончании очередного периода тестирования в виде списка найденных ошибок без указания точного времени их обнаружения. Поэтому для построения аппроксимирующей функции использованы точки, описывающие конечные состояния периодов интенсивного исправления, между которыми прошло не менее одного рабочего дня.

В результате анализа широкого класса функций выяснено, что наблюдаемые данные на этом участке хорошо аппроксимируются зависимостью $N(1 - e^{-(t/\lambda)^k})$.

Исследование ряда других программ с сохранившимися журналами системы управления версиями подтвердило, что выявленная закономерность с хорошей степенью точности описывает период заключительного тестирования программного обеспечения, непосредственно предшествующий сдаче его в эксплуатацию.

При этом получаемые по методу наименьших квадратов коэффициенты (N, λ, k) позволяют оценить количество ошибок на момент начала тестирования, эффективность выбранных методик тестирования и практическую скорость выявления ошибок, а также качество конечного продукта по оценке количества ошибок, оставшихся невыявленными к моменту окончания работ, которая на основании существующего предела полученной функции имеет вид

$$(1) \quad n(t) = Ne^{-(t/\lambda)^k}.$$

На основе приведенной зависимости получено выражение для ожидаемого количества ошибок $n(T_1, T_2) = N(e^{-(T_1/\lambda)^k} - e^{-(T_2/\lambda)^k})$, обнаруживаемых за период времени $T_1 - T_2$. В частности для рассматриваемого в качестве примера программного обеспечения станции «Заельцовская» была получена оценка, что в течение запланированных 2 месяцев выявляются 98.5% оставшихся ошибок при условии сохранения интенсивности тестирования.

Для получения количественных характеристик надежности, в частности среднего времени наработки до отказа, использована модель нестационарного пуассоновского процесса: введены вспомогательные функции интенсивности обнаружения ошибок $\mu(t)$ и количества остающихся не обнаруженными ошибок $n(t)$, тогда $dn(t) = -n(t)\mu(t)dt$, откуда $n(t) = Ne^{-\int_0^t \mu(\tau)d\tau}$, где N – суммарное количество ошибок в программе на момент начала тестирования. После ряда преобразований полученная из зависимости (1) интенсивность обнаружения ошибок имеет вид $\mu(t) = k / t^{1-k} \lambda^k$.

Поскольку в анализируемом случае $k < 1$, интенсивность обнаружения ошибок $\mu(t)$ со временем уменьшается. Качественно это объясняется тем, что к моменту времени t наиболее очевидные и часто встречающиеся ошибки уже найдены и исправлены, а для обнаружения оставшихся требуется совпадение большого количества условий или возникновение редко проявляющихся факторов.

Для передаваемой в эксплуатацию программы можно оценить среднее время наработки до отказа (*MTTF*), воспользовавшись тем, что на анализируемом участке $\mu(t)$ является медленно убывающей функцией и с достаточной степенью точности ее можно считать постоянной до момента обнаружения следующей ошибки: $\mu(t) = \mu(T_0)$, где T_0 – время начала эксплуатации, откуда $MTTF = n(T_0)\mu(T_0)$. На основе предоставляемой журналом системы управления версиями информации также можно определить среднее время ремонта (*MTTR*) и оценить коэффициент готовности разрабатываемого программного комплекса $A = MTTF / (MTTF + MTTR)$, а также время, необходимое для поиска и исправления ошибок с целью достижения требуемых эксплуатационных характеристик.

Необходимо отметить, что изменение текста программы в процессе тестирования не обязательно означает исправление обнаруженной критической ошибки: практика показывает, что заметная часть исправлений даже на этом этапе связана не столько с функциональностью или надежностью, сколько с изменениями пользовательского интерфейса, поэтому при оценке полученные значения A и *MTTF* должны быть нормализованы с учетом соответствующих пропорций. Требуемые коэффициенты могут быть получены также автоматически из журнала системы управления версиями при соблюдении разработчиками общих соглашений по оформлению комментариев (так, в рассматриваемом примере исправленные ошибки, которые могли повлиять на надежность и безопасность системы управления, помечаются модификатором «[CRITICAL]»).

2 Сравнение с другими методами

Для сравнения с результатами, получаемыми известными методами оценки надежности программного обеспечения, ниже приведены краткие описания наиболее часто используемых математических моделей, включая модели Нельсона (упрощенный вариант которой используется в ГОСТ 28195-89 «Оценка качества программных средств») и Джелински–Моранды, а также рассмотрены некоторые эмпирические методы.

Модель Нельсона (имеет несколько вариаций, описание приведено по [1]) была разработана с учетом основных свойств программ и может использоваться для расчета надежности программного обеспечения на всех этапах его жизненного цикла. В данной модели предполагается, что область, которой могут принадлежать входные данные программы, разделена на k непересекающихся областей Z_i , $i \in [1, k]$. Пусть p_i – вероятность того, что для очередного выполнения программы будет выбран набор данных из области Z_i (значения p_i определяются по статистике входных данных в реальных условиях работы программного обеспечения). Положим, что к моменту оценки надежности было выполнено n_i прогонов программного обеспече-

ния на наборах данных из области Z_i , и \tilde{n}_i из этих прогонов закончились отказом, тогда надежность программного обеспечения оценивается по формуле $R = 1 - \sum_{i=0}^k \frac{\tilde{n}_i}{N_i} p_i$.

Применяя предлагаемую в модели Нельсона линейную зависимость при оценке времени, которое требуется для обнаружения ошибки, и предполагая, что это время пропорционально количеству испытательных прогонов, на основе данных, использованных при построении графика зависимости версии программы от времени разработки программного обеспечения системы управления движением поездов станции «Заельцовская» (рисунок 2), выявлено более чем 60%-ое отклонение от значений, полученных опытным путем (для сравнения: результат, полученный в соответствии с предложенным методом, на аналогичных входных данных демонстрирует 5%-ю погрешность). Таким образом, основным недостатком модели Нельсона является ее неточность, вызванная необходимостью большого количества испытательных прогонов при фиксированных условиях, а также отсутствие возможности учитывать исправление ошибок при прогнозировании последующих значений показателей надежности.

Модель Джелински–Моранды (описание приведено по [2]) основана на предположении об экспоненциальном распределении времени до следующего отказа, при этом интенсивность отказов считается пропорциональной количеству оставшихся в программе ошибок. В этом случае вероятность безотказной работы программы оценивается как $R(t_i) = e^{-\lambda_i t_i}$, где интенсивность отказов $\lambda_i = C_D(N - (i - 1))$. Здесь C_D – коэффициент пропорциональности; N – первоначальное количество ошибок; отсчет времени начинается от момента последнего $(i - 1)$ -го отказа программы. При известных значениях t_1, t_2, \dots, t_k можно найти значения параметров модели C_D и N , а затем интенсивность отказов, время от последнего до следующего отказа t_{k+1} , и т.д.

На анализируемых экспериментальных данных (рисунок 2) оценка, полученная на основе модели Джелински–Моранды, продемонстрировала хорошую точность (около 20%), но все же заметно меньшую, чем оценка, полученная в соответствии с разработанным методом. В первую очередь это объясняется тем, что степень при t не предполагается постоянной (как в модели Джелински–Моранды), а вычисляется методом наименьших квадратов с учетом накопленной информации (например, для точки t_0 степень $k = 0.88$). Качественно меньшую точность получаемых на основе модели Джелински–Моранды оценок можно объяснить тем, что в общем случае не выполняются ни используемое в данной модели предположение о линейной зависимости между количеством ошибок и интенсивностью отказов, ни предположение о том, что при исправлении обнаруженных ошибок в программу не вносятся новые. Поэтому использование вычисляемых компенсирующих коэффициентов в соответствии с предложенным методом и позволяет получить более точную оценку.

Таким образом, важнейшим условием применимости модели Джелински–Моранды для оценки надежности конкретного программного продукта является соответствие результатов тестирования положенному в основу модели предположению об уменьшении интенсивности отказов после устранения каждой ошибки: получаемые в процессе тестирования и эксплуатации данные должны подтверждать факт увеличения интервалов времени (количества тестов), необходимых для обнаружения каждой последующей ошибки.

Данное замечание о применимости касается и других моделей, опирающихся на схожие предположения, например, модели Шика–Волвертона [3], интенсивность обнаружения ошибок в которой, как и в случае модели Джелински–Моранды, считается постоянной в течение интервала времени t_i между изменениями программы и пропорциональной числу ошибок, оставшихся в программе по истечении $(i - 1)$ -го интервала тестирования, а отличие состоит в том, что дополнительно учитывается среднее время выполнения программы в текущем интервале тестирования (вероятность безотказной работы оценивается как $R(t_i) = e^{-\lambda_i t_i^2 / 2}$).

Хотя для рассматриваемого в качестве примера программного обеспечения станции «Заельцовская» условие уменьшения со временем интенсивности обнаружения ошибок выполняется, для многих современных программных продуктов это не так, а следовательно применение модели Джелински—Моранды и ее модификаций может привести к неверным оценкам. Например, для одного из наиболее качественно сопровождаемых программных продуктов Microsoft Windows NT 4.0 в [4] приведены данные, демонстрирующие рост интенсивности обнаружения ошибок при выпуске обновлений и многочисленные факты внесения новых ошибок при исправлении найденных.

Еще один класс представляют так называемые эмпирические модели надежности, основанные на анализе накопленной информации о характеристиках ранее разработанных программ. Наиболее простые из них связывают число ошибок в программном обеспечении с его объемом. Например, фирма IBM использует модель, которая оценивает число ошибок в различных редакциях операционной системы как $N = 23M_{10} + 2M_1$, где M_{10} – число модулей, потребовавших 10 и более исправлений, а M_1 – число модулей, в которых было обнаружено менее 10 ошибок [5]. Аналогично, модель Холстеда оценивает количество ошибок в программе как $N = V / 3000$, где V – «объем» программы, рассчитанный по формуле $V = (N_1 + N_2) \log_2(n_1 + n_2)$, где N_1 – общее количество операторов программы; N_2 – общее количество операндов в программе; n_1 – число отличающихся операторов; n_2 – число отличающихся операндов [6]. Экспериментальная проверка этих моделей для оценки количества ошибок использующегося в качестве примера программного обеспечения станции «Заельцовская» показала значительную неточность данных методов.

Очевидно, что накопленные данные о предыдущих разработках не всегда применимы к последующим, поскольку при создании очередного программного продукта используются другие спецификации, возникают отличающиеся по сложности новые задачи, меняется коллектив разработчиков и т.д., поэтому заложенные в расчетах постоянные коэффициенты и приводят к такой низкой точности результата. Разработанный метод оценки надежности по информации, сохраняемой системой управления версиями, лишен этого недостатка, позволяя учитывать уникальную историю разработки конкретного программного продукта.

Заключение

Предложенный метод позволяет получить числовые оценки основных характеристик надежности программного обеспечения, таких как количество невыявленных ошибок, среднее время наработки на отказ, коэффициент готовности системы, вероятность опасного отказа за интересующий период времени и т.д., по информации, сохраняемой системой управления версиями; дополнительно продемонстрирован способ оценки количества времени, необходимого для поиска и исправления ошибок с целью достижения требуемых эксплуатационных характеристик.

Список литературы

- [1] Т. Тейер, Р. Липов, Э. Нельсон. Надежность программного обеспечения. – М.: «Мир», 1981. – 325 с.
- [2] Jelinski Z. and Moranda P. Software reliability research // Statistical computer performance evaluation. – 1972. – Vol. No 5. – P. 475–478.
- [3] Shick G.J. and Wolverton R.W. An analysis of computing software reliability models // IEEE Transactions on Software Engineering. – 1978. – July. – Vol. SE-4. – P. 104–120.
- [4] Прудник В.В. Оценка изменения надежности программного обеспечения в период эксплуатации // Проблемы информационной безопасности. Компьютерные системы. – 2000. – Т. 2 (<http://www.ssl.stu.neva.ru/ssl/publications/magazine/2000/2/2/prudnik.pdf>).
- [5] Бозм Б., Браун Дж., Каспар Х. Характеристики качества программного обеспечения. – М.: Мир, 1981. – 208 с.
- [6] Холстед М. Начала науки о программах. – М.: Финансы и статистика, 1981. – 128 с.

ОПИСАНИЕ, АНАЛИЗ И ПРЕОБРАЗОВАНИЕ СТРУКТУРЫ СЛОЖНЫХ МНОГОУРОВНЕВЫХ ПРОЦЕССОВ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ С ИСПОЛЬЗОВАНИЕМ ГРАФОВ С ОБОЛОЧКОЙ

Л.Л. Чекалов

ООО «Транс-Мобил Спедишн»
443030, Самара, Желябова д. 3 - 45, Россия
ltrans777@mail.ru
тел: +7 (963) 115-00-77

Ключевые слова: структурное моделирование, многоуровневые графы

Abstract

Development of methodology for analysis and design of structural models of systems is one of the areas for improvement of automated systems creating tools. Proposed the use of graphs with a shell to describe, analysis and transformation of structural models of complex nested and multi-level processes and objects in different domains. Using graphs with the shell we can transform the structural model of the process or object, simplifying and complicating its presentation by folding and unfolding (composition and decomposition) subgraphs shells in the apex with the same name. In graphs with a shell, you can use all the mathematical tools of graph theory in the design and analysis of structural models.

Введение

Автоматизированная система (АС) – это организованная совокупность средств, методов и мероприятий, используемых для регулярной обработки информации для решения поставленных задач, и состоит из персонала, комплекса средств автоматизации и регламентов работы.

Главной целью создания АС является не упрощение, но категоризация и **стандартизация автоматизируемого процесса или объекта**, что позволяет обеспечивать стабильность работы системы, прозрачность её контроля и анализа слабых мест и основания для её развития либо свёртывания (списания, замены).

Жизненный цикл автоматизированных систем охватывает все стадии ее создания, сопровождения и развития:

- анализ и разработка требований к системе;
- проектирование системы, в том числе:
- разработка архитектуры и структурной модели системы;
- разработка математических моделей предметной области;
- разработка и анализ макета системы (эскизный проект);
- разработка технического проекта;
- разработка (программирование) системы;
- интеграция, сборка и испытания системы;
- эксплуатация системы, сопровождение, модернизация.

Существуют различные подходы к организации работ, которые связаны с разными моделями жизненного цикла. Наиболее распространенными являются каскадная и спиральная модели разработки прикладных информационных систем.

Каскадная модель жизненного цикла информационных систем отражает схему разработки в виде последовательности завершённых этапов. Данная модель характерна для доста-

точно полно специфицированных систем, эксплуатация которых предполагается в неизменных условиях, а доработки сводятся к исправлениям ошибок программирования и незначительным корректировкам проектных решений.

Спиральная модель жизненного цикла характерна для развивающихся систем. Эта модель предусматривает возможность их **модернизации**. При этом модернизация системы затрагивает все стадии разработки: анализ новых возникших задач, проектирование и программирование новой версии системы. Преимущество спиральной модели для разработки информационных систем очевидно, но они могут быть в полной мере реализованы только при **высокой степени автоматизации этапов проектирования, разработки и сопровождения системы**.

1 Развитие средств разработки автоматизированных систем

Рост сложности создаваемых АС и повышение требований к их надежности вело к развитию средств их разработки. С одной стороны, появились объектно-ориентированные языки программирования, отдельно можно отметить разработку языка Ада [1], создаются средства концептуального программирования в терминах предметной области. Также на рынке программных продуктов появляется большое количество инструментальных программных средств автоматизации разработки программных приложений, так называемых CASE – средств (Computer-Aided Software Engineering), к которым обычно относят любое программное средство, автоматизирующее те или иные этапы жизненного цикла.

С другой стороны, часть специалистов старались решить сложную задачу создания крупномасштабных АС, включающих как людей и машины, так и программное обеспечение, развивая методологии структурного анализа и проектирования, позволяющие внести больше упорядоченности, начиная с ранних стадий разработки АС.

Целью структурного моделирования является построение спецификаций основных функций разрабатываемой системы, ее архитектуры, состава компонент, структуры данных, блок-схем и других объектов в наглядной и доступной форме.

В настоящее время получили распространение несколько методологий:

- SADT (Structured Analysis and Design Technique) [2];
- Структурного системного анализа и проектирования Йодана/Де Марко (Yourdon/De Marko);
- Анализа и проектирования систем реального времени Уорда-Меллора (Ward-Mellor) и Хатли (Hatley);
- Информационного моделирования Мартина (Martin).

В методологии SADT (Structured Analysis And Design Technique) [2] существуют два основных направления: *функциональные модели* выделяют события в системе, *модели данных* выделяют объекты (данные) системы, связывающие функции между собой и с их окружением. В обоих случаях используется **один и тот же графический язык блоков и дуг** (но блоки и дуги меняются ролями).

Функциональный вариант данной методологии в правительственной стандартизированной версии, получившей название *IDEF0* – это иерархически организованная совокупность диаграмм. Диаграмма содержит 3-6 блоков. Каждый из блоков потенциально может быть детализирован на другой диаграмме. Разделение блока на его структурные части (блоки и дуги) называется *декомпозицией*. Декомпозиция формирует границы, то есть блок и касающиеся его дуги определяют точную границу диаграммы, представляющей декомпозицию этого блока. Эта диаграмма называется *диаграммой-потомком*. Декомпозируемый блок называется *родительским блоком*, а содержащая его диаграмма – *родительской диаграммой*. Название диаграммы-потомка совпадает с функцией родительского блока.

Таким образом, IDEF0-диаграмма является декомпозицией некоторой ограниченной функции (субъекта). Принцип ограничения субъекта встречается на каждом уровне. Один блок и несколько дуг на самом верхнем уровне используются для определения границы всей системы. Этот блок описывает общую функцию, выполняемую системой. Дуги, касающиеся этого блока, описывают главные управления, входы, выходы и механизмы этой системы. Диаграмма, состоящая из одного блока и его дуг, определяющая границу системы, называется *контекстной диаграммой модели*. Все, что лежит внутри блока, является частью описываемой системы, а все, лежащее вне его, образует *среду системы*.

Существует ряд готовых инструментальных систем, которые автоматизируют структурное проектирование прикладных информационных систем и оперируют геометрическими изображениями разных типов диаграмм, стрелок указателей, бирок и прочее, как например, VPwin и Erwin – средства для разработки АС [3].

Приведенное выше **подробное** описание методов структурного проектирования систем показывает сходство представления диаграмм и обычных графов. Но, в данном случае, недостаток представления обычных графов – отсутствие возможности создавать вложенные фрагменты диаграмм для их декомпозиции и композиции.

В данной работе предлагается использовать представление структурных моделей в виде графов с оболочкой, полученных из обычного графа.

2 Графы с оболочками

Рассмотрим обычный граф, представленный вершинами V и ребрами R на рисунке 1. Такое представление графа может описывать сколь угодно сложные структурные модели объектов или процессов, если ребра сделать ориентированными.

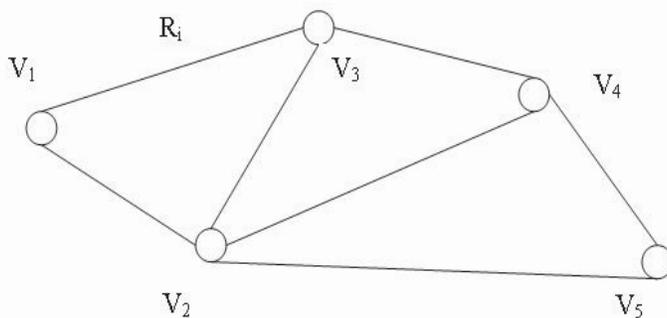


Рисунок 1 – Обычный граф с вершинами V и ребрами R

В современной науке и технике чаще всего рассматриваются очень сложные многоуровневые процессы и объекты, которые сложно охватить одним взглядом и тем более сложно вести описание всех частей одновременно. Для описания объекта или процесса по частям на различных уровнях вложенности предлагается использовать немного измененное представление графа, названное графом с оболочкой. Граф с оболочкой получается из обычного графа путем проведения вокруг одной или нескольких вершин замкнутого контура – оболочки, где пересечения ребер с этим контуром образуют вершины оболочки – рисунок 2.

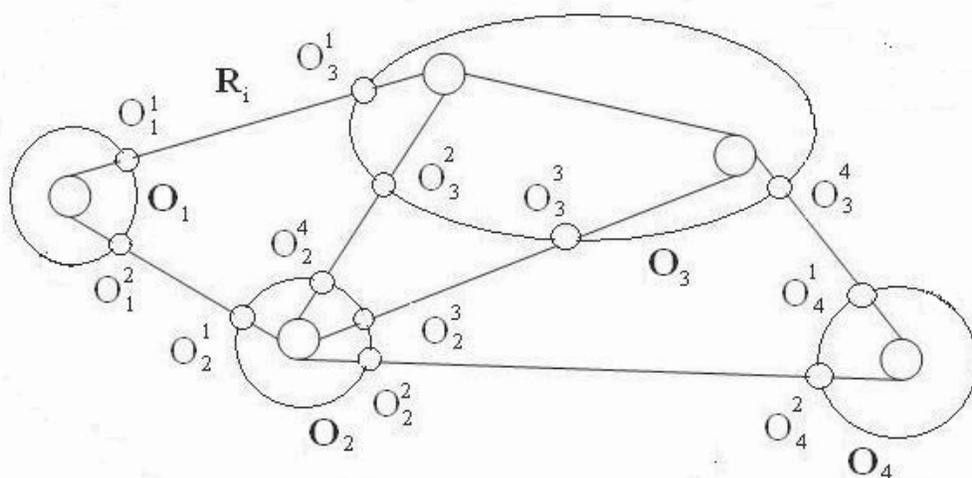


Рисунок 2 – Преобразование обычного графа в граф с оболочкой

В результате образовались новые вершины O_1 - O_4 . Вершины V_1 , V_2 и V_5 обычного графа соответствуют вершинам O_1 , O_2 и O_4 графа с оболочкой, а вершины V_3 , V_4 – вершине O_3 , как показано на рисунке 3. В графе с оболочкой у каждой вершины O_n появились вершины оболочки O_n^1 , O_n^2 и т.д. Также граф может быть вложен в собственную внешнюю оболочку O_0 , которая будет обычной вершиной с оболочкой на более высоком уровне вложенности, где ребра W_j будут внешними ребрами графа.

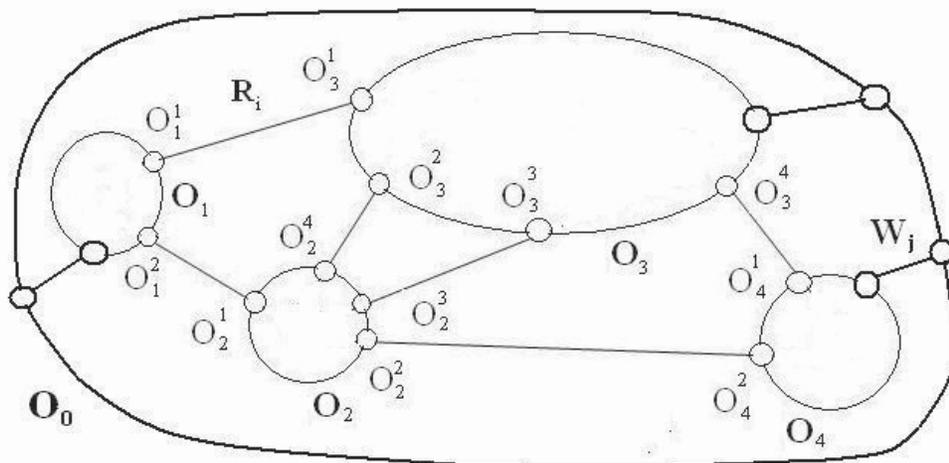


Рисунок 3 – Граф с внешней оболочкой O_0 и с вершинами O_n и ребрами R_i

Вершина с оболочкой может быть как вершиной последнего низшего уровня вложенности, так и скрывать в себе бесконечное множество вложенных подграфов. Вершина с оболочкой становится универсальным способом описания сложных многоуровневых объектов и процессов. Вершина с оболочкой позволяет при необходимости сворачивать сложные подграфы, упрощая их рассмотрение, а потом разворачивать до первоначального состояния без потери связей, используя единый алгоритм преобразования на всех уровнях.

Если обычный граф может быть представлен списком вершин и ребер, а ребра R могут быть заданы парами $V_1 V_2$, то граф с оболочкой представляется множеством вершин O_n , каждая из которых представляется множеством вершин оболочки O_n^m и ребер R_i , где R_i задается парами вершин оболочки $O_1^1 O_3^1$ и т.д.. Граф с оболочкой или любой его подграф может быть отображен в обычный граф.

Граф с оболочкой позволяет описывать распределенные процессы. Граф с оболочкой так же позволяет организовать многослойное представление предметной области, в зависимости от решаемой в данный момент задачи. В этом случае, с каждой оболочки вершины может сниматься только определенный образ, требуемый в конкретном случае, если например, различным специалистам необходима только их касающаяся информация об одном и том же объекте или процессе.

С другой стороны, можно при необходимости для упрощения представления структуры сворачивать фрагменты графа, или в дальнейшем вырезать их из общей структуры и рассматривать отдельно, а потом можно встроить назад.

3 Анализ графа с оболочкой на наличие сильно связанных подграфов и их преобразование в слабо связанные

Если рассматривать разработку сложного процесса с любой его точки, то можно предположить, что в окончательном виде граф будет иметь сильно связанные замкнутые итерационные области, не позволяющие получить «поток». Один из вариантов борьбы с этим – сворачивание сильно связанных подграфов графа в одну вершину. Но это может привести к неоправданному упрощению структурной модели процесса, и все равно остается вопрос, что делать со свернутым сильно связанным подграфом.

Второй вариант предлагается автором, как метод разделения вершин сильно связанного подграфа [4], позволяющий за конечное число разделения вершин преобразовать любой сильно связанный подграф в слабо связанный. При разделении вершины V_1 на рисунке 4 образуются две смежные вершины V_1^1 и V_1^2 и дополнительное ребро S , размыкающее сильно связанный граф. Предметная составляющая этого процесса разделения вершин решается из особенностей задач предметной области.

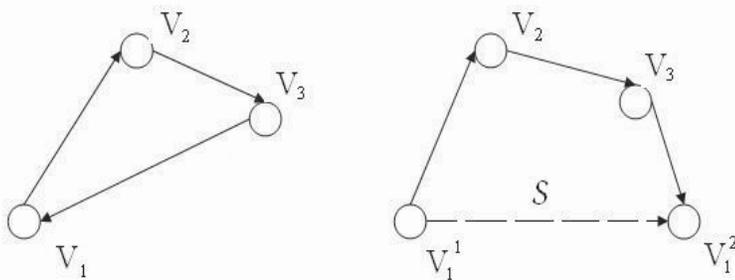


Рисунок 4 – Разделение вершины V_1 сильно связанного графа

Применение метода разделения вершин и получение в результате слабо связанного графа позволяет задавать относительную последовательность или время для вершин графа.

Заключение

Графы с оболочкой позволяют:

- описывать структурные модели сложных вложенных и многослойных процессов и объектов в различных предметных областях;

- проводить анализ структурной модели процесса на наличие замкнутых сильно связанных областей и преобразовывать сильно связанные области в слабо связанные методом разделения вершин;
- проводить как отдельную разработку частей структурной модели процесса с последующей ее композицией, так и проводить при необходимости декомпозицию структурной модели процесса;
- преобразовывать структурную модель процесса или объекта, упрощая и усложняя ее представление путем сворачивания и разворачивания (композиция и декомпозиция) оболочек подграфов;
- автоматизировать описание, анализ и преобразование структурной модели процесса и выполнять перечисленные действия в режиме реального времени, когда нет заранее сформированной структурной модели процесса;
- использовать весь математический аппарат теории графов при разработке и анализе структурных моделей.

Графы с оболочкой могут быть частью инструментальных средств разработки АС, приближающих эти средства к понятиям структуры предметной области и позволяющих использовать единые математические средства для формирования структурной модели предметной области.

Список литературы

- [1] Бар Р. Язык Ада в проектировании систем: Пер. с англ. – М.: Мир, 1988. – 320 с.
- [2] Марка Д.А., МакГоуэн К. Методология структурного анализа и проектирования SADT. – М., 1993.
- [3] Маклаков С.В. ВРwin и ERwin: Case-средства для разработки информационных систем. Изд.: Диалог-МИФИ. 256 стр.
- [4] Чекалов Л.Л. Метод разработки структурной многоуровневой модели технологии автоматизированного проектирования внутриплощадочных технологических объектов обустройства нефтяных месторождений. Научно-технические проблемы разработки и обустройства нефтяных месторождений: Сб. науч. тр. Институт «Гипровостокнефть». – Куйбышев, 1990. – с.114-118.

Biography

1999 – present time	“Trans-Mobile Spedition Ltd.”, Samara, logistic firm, director
1993 – 1999	“Samteks Ltd.”, Samara, custom clearance department head
1990 – 1993	“GiproVostokNefit” design and research institute, Samara, CAD laboratory
1987 – 1990	“Sentral Reseach Institute of Economy and CAD-Systems for Civil Engineering”, Moscow, post-graduate, specialization – CAD systems (In 1991 received the degree of candidate of technical sciences)
1985 – 1987	“GiproVostokNefit” design and research institute, Samara, CAD laboratory
1980 – 1987	“Institute of Civil Engineering”, Samara, student

ИНФОРМАЦИОННАЯ МОДЕЛЬ СОЗНАНИЯ И ОПИСАНИЕ ЭВОЛЮЦИИ ЕЁ ФУНКЦИЙ ОТ ЖИВОТНОГО К ЧЕЛОВЕКУ

Л.Л. Чекалов

ООО «Транс-Мобил Спедишн»
443030, Самара, Желябова д. 3 - 45, Россия
ltrans777@mail.ru
тел: +7 (963) 115-00-77

Ключевые слова: эволюция сознания, информационная модель сознания, информация, обработка информации

Abstract

The model of mind and evolution of its functions from animal to the man in information terms. This article gives a broader definition of the concepts of information and information processing, also notes the invariance of information to substantive component of knowledge. Development of the ability to process information took place in the evolution of living matter from animal to man. The fragmentation of information influences the external perception of the objectivity of the information processing by the mind.

Введение

Сознание человека постоянно воспринимает окружающий мир, накапливает опыт и принимает решения, управляет сложными системами, но сам процесс этот остается за кадром и не всегда поддается логическому объяснению. Попытаться понять возможности сознания современного человека достаточно сложно без описания эволюции способностей живой материи к обработке информации и принятию решений, начиная от животного до человека. За начальную точку описания берутся возможности животных, потом появление человека и развитие способностей его сознания до настоящего времени.

Сознание в данной работе рассматривается как черный ящик – по результатам его работы. Рассматривается эволюция только тех функций сознания, в части способности живой материи к восприятию, хранению и обработке информации.

1 Информационная модель сознания и описание эволюции её функций от животного к человеку

1.1 Описание понятий

Для описания информационной эволюции сознания уточняются понятия памяти, информации, обработки информации.

Всё в этом мире находится в движении и взаимодействует постоянно и неизбежно. Результат взаимодействия – взаимное изменение формы или физико-химического состояния или энергетического состояния или еще чего-то. Образуются обоюдные следы, остающиеся от взаимодействия объекта с объектами внешней среды.

Память – это следы взаимодействия, оставшиеся на объектах, т.е. часть свойств одного объекта передается на какое-то очень короткое время или на весь период жизни другому объекту. Что-то вроде отражения внешнего мира в себе (объекте). Память это всеобщее свойство материи, как и постоянное взаимодействие. Поэтому все объекты постоянно меняются, при-

нимая на себя часть свойств других объектов. Память – естественное свойство объектов и это относится как к живой, так и не живой материи.

Например, удар двух твердых тел оставляет следы (информацию) на обоих. Тела сохраняют на себе память о взаимодействии. Память – прошлый опыт объектов или субъектов.

Информация переводится с латинского языка (informatio) обычно как «представление», «понятие» (о чем-либо), «изложение», «сведения», «осведомление», «сообщение» и т.п. Термин этот интуитивно ясный, обладает широким смысловым полем и поэтому трудно поддается строгому определению [1]. В данной работе мне хотелось бы определить **информацию** как часть свойства (сведений) объектов, передаваемые между ними в результате взаимодействия. Полной передачи свойств объектов не происходит, иначе это будет просто второй такой же объект.

Информация находится между двумя её предельными видами – **дискретной и непрерывной**.

С определением понятия дискретной информации особых проблем быть не должно - это информация, имеющая определенные границы в пространстве и времени, в простейшем виде 0-1, да - нет. Но это не значит, что непрерывные функции относятся к непрерывной информации. Это тоже дискретная информация. Самое простое предельное состояние дискретной информации - это точка.

Бесконечное множество различных точек не ограниченных в пространстве и во времени предельное с другой стороны состояние информации – непрерывное. Самый понятный вариант непрерывной информации - это зрительный образ.

Ежесекундно через зрение в сознание поступает огромное количество информации и мы «мгновенно» распознаем все образы, связанные с прошлым опытом.

Дискретная информация обратима во времени, непрерывная информация не необратима во времени. Необратимость непрерывной информации - гарантия течения времени. Таким образом, всё понимание термина информации укладывается между двумя предельными ее формами – дискретной и непрерывной.

Также необходимо отметить, что в этом смысле информация – понятие инвариантное к конкретным предметным знаниям.

Сознание человека соответственно обладает возможностью **дискретной и непрерывной обработки информации**. Дискретная обработка информации это когда человек пошагово начинает считать, это любая обработка информации компьютером в соответствии с любым заданным алгоритмом. Непрерывная обработка информации это – один из вариантов – когда при отсутствии алгоритма мгновенно и ежесекундно производится распознавание зрительного образа.

Обработка информации происходит в течении времени:

- время непрерывной обработки непрерывной информации сознанием стремится к нулю;
- время дискретной обработки непрерывной информации стремится к бесконечности.

Понятие алгоритма появляется при дискретной обработке дискретной или непрерывной информации. Алгоритм это протяженное во времени преобразование информации по определенным правилам (законам).

В неживой материи роль алгоритма при взаимодействии играет закон природы. В живой материи при обработке информации роль закона переходит к алгоритму. Это правило взаимодействия или преобразования объектов или информации, составленное сознанием.

При разворачивании информации требуется алгоритм, но это всё та же дискретная информация, просто в видоизмененной форме. Информация срастается с алгоритмом. Как закон природы связан с материей. Материя разворачивается (взаимодействует) во времени по определенным физическим законам.

ДНК превращается в человека тоже по алгоритму или закону. Это одно и то же свойство трансформации информации имеющее одним из своих параметров время, другим – алгоритм или закон.

Также необходимо ввести понятие субъекта – как носителя памяти и способности к восприятию и обработке информации.

1.2 Текущая информационная модель сознания человека

Теперь попробуем нарисовать информационную модель сознания современного человека, используя обычный практический опыт доступный каждому человеку.

Человек через органы чувств получает информацию о внешнем мире. Большую часть около 80% через органы зрения. Луч внимания человека может переключаться между органами чувств, но информационные процессы, в общем, сходны между собой для всех органов чувств. Поэтому наиболее удобно и для всех интуитивно понятно рассматривать зрительное восприятие мира.

Информационную модель сознания человека условно можно представить в виде цилиндра, который является памятью, а один из торцов связан с текущим временем (рисунок 1). В этом торце выделен «луч внимания» и вокруг него зона бокового зрения или в дальнейшем будем называть боковым вниманием. Луч внимания может по усмотрению человека сжиматься и расширяться, поэтому введено понятие радиуса луча внимания, а радиус бокового внимания в основном зависит от физиологического строения органов чувств, а в данном случае зрения.

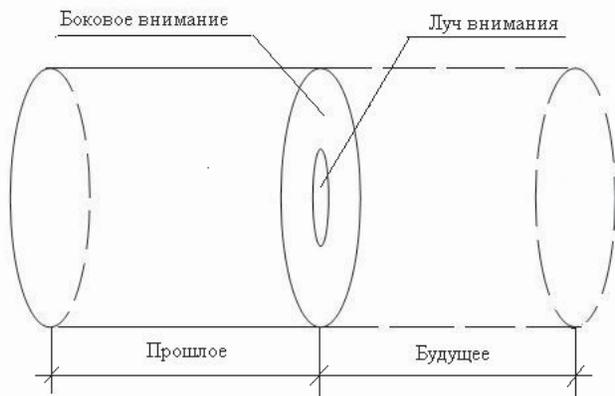


Рисунок 1 – Информационная модель сознания человека в настоящее время

Луч внимания обладает некоторой свободой выбора и может быть направлен на любой объект внешнего мира. Но, наверное, эволюционно обусловлено постоянное обязательное перескакивание луча внимания по объектам. Попробуйте задержать внимание на одной точке хотя бы на несколько секунд – скорее всего не получится.

Зрительный (непрерывный информационный) образ внешнего мира мгновенно распознаётся при наличии зоны памяти (опыта) по отдельным дискретным объектам и тут же уходит в зону памяти, откуда снова вернуть его в зону луча внимания почти невозможно.

Но, как известно, в состоянии гипноза человек вспоминает забытые жизненные ситуации. Это также часто используется в криминалистике для восстановления подробностей происшествия.

Таким образом, можно предположить, что в зоне памяти сохраняется **огромное количество информации**, прошедшее не только через луч внимания, но и боковое внимание. Доступ к этой информации закрыт, но как мы говорили, взаимодействие зоны текущего внимания и зоны памяти происходит мгновенно и без участия свободы воли человека, т.е. способность к

непрерывной обработке информации работает **объективно и почти независимо от воли человека**, как в зоне луча внимания, так и в зоне памяти.

В 2008 году группа ученых провела опыты [2] с нажатием кнопок с использованием современной технологии – функциональной магнитно-резонансной томографии (МРТ). МРТ позволяет исследовать изменение активности различных участков мозга, наблюдая за изменением тока крови (наиболее активные участки мозга требуют больше кислорода). Испытуемых сажали перед экраном, на котором менялись буквы. Испытуемый должен был запоминать, при появлении какой буквы они совершали выбор между двумя кнопками. Ученые пытались определить, возбуждение каких участков мозга содержит наибольшую информацию о том, какой выбор совершит человек: нажмет он на левую или на правую кнопку. С учетом всех статистических поправок повышение активности мозга позволяла предугадывать выбор человека до того, как он его осознавал. В ряде условий прогноз удавалось осуществлять за 10 секунд до момента принятия осознанного решения испытуемым! Нейрофизиолог Джон-Дилан Хейнс (John-Dylan Haynes) и коллеги, участвовавшие в этом исследовании, пришли к выводу, что сеть управляющих участков мозга, ответственных за принятие решений, начинает формироваться задолго до того, как мы начинаем это осознавать.

Обработка информации - это свойство зоны памяти, как и хранение информации, **сформированное в процессе эволюции** и работающее по законам, трудно поддающимся влиянию воли субъекта, т.е. это во многом объективный природный процесс. Обработка информации в зоне памяти в большей степени непрерывная, а вот в зоне луча внимания сознание человека пытается заниматься дискретной обработкой дискретной информации и с большим напряжением складывает и умножает большие числа или выполняет какие-либо логические действия. Эта зона в большей мере управляема человеком при обработке информации.

1.3 Изменение функций модели сознания в процессе эволюции

В животном мире изменение функций обработки информации происходило в процессе эволюции от дискретного восприятия мира и дискретной обработки поступившей информации к непрерывному восприятию мира и непрерывной обработке информации, как к эволюционно более сложному процессу.

Так, например, земноводные воспринимали мир в типично дискретном виде и так же дискретно обрабатывали информацию и память была соответственно дискретной. Что-то вроде темная тень сверху - хищник, быстро движущаяся точка – добыча [3].

В зоне луча внимания воспринимаются дискретные образы, такими же они попадают в память и такими же они обрабатываются. Здесь можно говорить одновременно как о неразвитости органов чувств, так и о неразвитости мозга и соответственно его функций, хотя органы чувств можно назвать выдвинутым на поверхность тела мозгом.

Постепенно в процессе эволюции живые организмы стали воспринимать непрерывную картину мира и стали непрерывно обрабатывать ее. Животный мир (млекопитающие) достиг возможности непрерывного восприятия мира и непрерывной обработки информации в зоне восприятия информации. Полученная и обработанная информация используется животными **для управления движением тела во внешнем пространстве**. Млекопитающие уже достаточно полно и красочно воспринимают внешний мир, это уже не простые образы земноводных, в простых формах; воздействуют на внешний мир - копают норы, строят хатки (бобры), запоминают десятки команд.

В результате информационная модель у высших животных по виду такая же, как у человека, кроме функций зоны текущего внимания. У человека в зоне текущего внимания активно развивается функция управления, направленная на изменение и создание новых внешних объектов. Т.е. живая материя, развиваясь от восприятия и обработки дискретной информации о мире к непрерывной для управления телом, достигла своего предела из которого появилось

сознание человека, который начал второй цикл развития обработки информации и опять от дискретной к непрерывной, но теперь это уже связано с управлением внешними объектами.

Также по емкости дискретной информации зона памяти у животных существенно меньше, чем у человека. У человека она постепенно увеличивалась в процессе эволюции. При обучении обезьяна запоминает до 300 слов или несложные алгоритмы, но дальше дело не идет.

Если взять для сравнения наскальные рисунки и картины эпохи Возрождения, то ясно виден переход от дискретного к непрерывному при направленном воздействии на внешние объекты. Человек развивал эту функцию сознания, обучаясь управлять кистью своей руки. Большая часть верхней коры головного мозга, ответственного за мышление, также управляет кистью руки [4]. **Через управление частью тела – рукой – происходит обучение управлением внешними объектами и их изменением.**

Таким образом, человек учится строить более длинные алгоритмы воздействия на внешний мир и за счет этого получает в зоне луча внимания гораздо больше свободы воли, чем у животных. Это еще один аспект эволюции, который связан с обработкой информации. Животное «автоматически» использует способность к обработке информации, сформированную живой материей в процессе эволюции. У человека появляется в луче внимания некоторая свобода воли для построения «длинных и сложных» алгоритмов и соответственно возможность для отказа от решений, выдаваемых из зоны памяти, как у животного. Но с другой стороны, человек создал среду обитания и знания, которые он загружает в зону памяти, и решения, принимаемые в этой зоне, на самом деле могут быть более объективными, чем в зоне луча внимания.

С появлением свободы воли появилась возможность управления информацией и управления миром. Информация от органов чувств, попав в память, сохраняется и продолжает обрабатываться в основном без участия луча внимания, но луч внимания не всегда может попасть туда и вспомнить «все». Одна из причин в том, что эволюционно сложилось так, что в памяти осуществляется так же управление внутренними органами организма и если «беспокойный» луч внимания попадет туда, то для организма это может кончиться плачевно.

В работах Г.Селье, Н.П.Бехтеревой говорится о том, что за последние 50 лет человеческое сознание подошло к пределу своих возможностей по функционированию в перегруженной информационной среде. Основной вывод этих работ: если человек должен принимать решение в условиях большого объема информации, повышенной значимости или мотивации, в короткий промежуток времени, то возникает повышенная вероятность сбоя управлением работы внутренних органов. Т.е. процесс обработки информации попадает в зону управления внутренними органами и если возникает сильное психологическое напряжение или отрицательные эмоции, то это ведет к повреждению организма.

Это один из признаков того, что человек подошел к пределам информационных возможностей своего сознания в его текущем состоянии, и возможно нас ждет новый этап эволюции.

1.4 Объективность отражения мира сознанием и творческое сознание

Объективность восприятия внешнего мира и принятие поведенческих решений в результате обработки информации такой же фактор естественного отбора у животных. Если субъектом не объективно воспринимается внешняя ситуация, не корректно происходит управление телом, то либо животное остается голодным и умрет, либо его съедает.

Еще один вариант не объективности – это муха, не замечающая стекла, и бьющаяся об него. Муха необъективно воспринимает информацию о мире и бьется об объективный закон бытия.

Таким образом, объективность работы сознания в части восприятия мира, обработки информации, принятия решений и управления телом – составная часть эволюционного процесса по развитию функций сознания. Человеку от природы достался инструмент, которым он пользуется в меру жизненной необходимости, в основном не думая о его возможностях и функциях.

Ещё раз возвращаемся к информационной модели сознания:

- поступающая в сознание информация проходит через луч внимания и боковое внимание и «исчезает» в памяти;
- луч внимания перескакивает на следующий объект и всё повторяется снова.

Внешний мир представлен в луче внимания как огромное множество фрагментов, «исчезнувших» в памяти. Представьте популярную игру – сложите картину из пазлов. Множество фрагментов разбросано по столу, а картина не собрана. В результате сознание не видит связи, которые очевидны на общей картине. Сознание в луче внимания видит мир сильно фрагментированным. На фрагментах не видно общей картины мира в данный момент. Объективность в данном случае сильно фрагментирована и связи, которые видны на общей картине – не заметны на фрагментах.

На курсе «Принятие решений для руководителей» Зегер Дегрэйв (Zeger Degraeve) профессор London Business School демонстрировал психологический эксперимент, на котором нескольким испытуемым показывали на компьютере фрагмент игры в баскетбол и просили посчитать количество ударов мяча. По окончании просмотра каждый назвал какое-то количество ударов. А потом испытуемых спросили – не заметили ли они чего-нибудь необычного? Никто ничего не заметил. Показали фрагмент игры ещё раз. В середине игры на площадку вышла обезьяна, помахала зрителям и ушла. И этого никто не заметил в первый просмотр.

Таким образом, один из вариантов объективности восприятия мира связан со складыванием фрагментов в целостную картину в луче внимания. Это способность талантливого сознания. Для этого необходимо расширять луч внимания и держать его в спокойном состоянии. Кстати, из обычных видов деятельности этому способствует рисование.

Вспоминая высказывание А.Экзепюри: «Последний школяр на отделении точных наук знает о законах природы куда больше, чем знали Декарт и Паскаль. Но способен ли школяр мыслить, как они?» - ещё раз хочу подчеркнуть инвариантность информации к предметной составляющей.

Сознание по отношению к отдельным видам знания является универсальным инструментом. **Неважно о чем думает ученый – важно как.** Очень редко встречается врожденная гениальность, но это отдельная тема для обсуждения. Обычно, присутствует длительный процесс **дискретного** изучения какой-либо области знания, а потом, если удастся охватить вниманием созданную в памяти «картину», то легче заметить научные закономерности. Я бы назвал это **непрерывным** мышлением в зоне луча внимания. Поэтому очень часто ученым удается быть универсальными и добиваться успехов в различных областях знания и даже в искусстве. Все это относится к категории, которую можно назвать талантом.

Ещё одним из аспектов таланта, наряду с расширением радиуса луча внимания, можно считать необходимость удержания его в спокойном состоянии, чтобы не исказить картину мира.

Заключение

Сознание, являясь инструментом познания инвариантно к предметным знаниям. Сознание поддается тренировке, и такие виды деятельности, связанные с непрерывной обработкой информации в зоне луча внимания, как например рисование, литература способствуют расширению луча внимания и более целостному восприятию.

В начале эволюции живой материи движение внешней среды в большей мере управляет поведением субъекта, через реакции на поступающую информацию. В настоящий момент у человека наоборот гораздо больше возможностей управлять и трансформировать внешнюю среду, и осуществлять менее зависимое от нее поведение.

Таким образом, человек в зоне луча внимания, в процессе эволюции обучаясь изменять рукой внешнюю среду, получил свободу выбора (воли), которая может быть направлена как на внешнюю среду, так и на самого себя.

Увеличение свободы воли и средств воздействия на внешнюю среду ведет к опасности не заметить и не вписать движение цивилизации во внешние законы природы. Фрагментарное сознание не обращает на это внимание, а целостное сознание гораздо острее чувствует несоответствие своей деятельности окружающей ситуации. Возможно, это становится вопросом выживания человечества и его дальнейшей эволюции.

Список литературы

- [1] Корогодина В. И., Корогодина В. Л. Информация как основа жизни. – Дубна: Издательский центр «Феникс», 2000. – 208 с.
- [2] John-Dylan Haynes. Unconscious determinants of free decisions in the human brain // *Nature Neuroscience* 11, 543–545 (1 May 2008) | doi:10.1038/nn.2112.
- [3] Эндрю А. Искусственный интеллект: Пер. с англ./Под ред. и с предисл. Д.А. Поспелова. -М: Мир, 1985. 264с.
- [4] Ярошевский М.Г. История психологии от античности до середины XX в. Учеб. пособие. – М., 1996 – 416 с.

КАТЕГОРНЫЕ МЕТОДЫ КОМПЛЕКСНОГО ПРЕДСТАВЛЕНИЯ И СТРУКТУРИЗАЦИИ РАЗНОРОДНЫХ ДАННЫХ В ИНФОРМАЦИОННЫХ СИСТЕМАХ АНАЛИЗА СЛОЖНЫХ ОБЪЕКТОВ¹

В.И. Батищев, Н.Г. Губанов

Самарский государственный технический университет
443100, Самара, ул. Молодогвардейская, 244, Россия
vib@list.ru
тел: (846)3371371, (846)3371274

Ключевые слова: *сложные технические объекты, многомодельные комплексы, логический вывод*

Abstract

Are given the methods of constructing the categoring models of complex objects on the basis of inductive logical conclusion. Questions of the practical application of data of models in the analytical systems are examined.

Введение

Анализ современных тенденций, связанных с автоматизацией оценки состояния сложных технических объектов показал, что существенное усложнение объекта анализа приводит к тому, что результаты, полученные с использованием отдельных, разрозненных, хотя и глубоко исследованных, методов системного анализа, управления, интеллектуальных технологий не приносят вполне удовлетворительного результата. Это подтверждается активизацией, на настоящий момент, исследований, опирающихся на синтез различных методов, формирования некоторых надпредметных метасистем, формирующих семантические конструкции, обобщающие накопленный опыт, методологии и информационные технологии системных исследований.

Сложность объекта анализа, объективно, в силу необходимости выполнения условия адекватности, порождает и сложность информационной системы анализа состояния. Данный факт определяет важность и актуальность разработки конструктивных методологий анализа, синтеза и управления информационными системами анализа вкупе с объектами анализа.

Постоянно увеличивающаяся стоимость информационных ресурсов в современной экономике, обуславливает прагматические требования к максимальному использованию накопленных баз данных, баз знаний, фактографических и документальных систем. Данный факт определяет актуальность вопросов отчуждения, адаптации и повторного использования накопленных моделей, технических решений. Несмотря на обилие научно-технических идей, методов и информационных технологий переноса результатов исследования системных закономерностей с одного объекта исследования на другой, вопрос остаётся до сих пор не вполне решённым из-за уникальности объекта анализа. Данный факт ведёт к необходимости дополнительных процедур декомпозиции, некоторых допущений, формирования и анализа гипотез для выявления и проверки условий эквивалентности объектов анализа.

Общесистемные свойства сложного технического объекта, как объекта анализа приводят к необходимости использования полимодельных комплексов, в качестве основы анализа состояния. Анализ проблем построения данных структур выявил ряд общих задач: согласование разнородных моделей, согласование многомодельного комплекса с другими денотатами (сре-

¹ Работа выполнена в рамках гранта РФФИ (проект № 09-08-00459-а).

дой; лицами, принимающими и обосновывающими организационные и технические решения; объектом анализа). Данные задачи лежат в русле общенаучной проблематики – проблеме квалитметрии моделей.

Анализ показал перспективу локальной организации для построения систем данного класса, это обусловлено относительной самостоятельностью, элементов сложноструктурированного объекта анализа, а также исходя из постановки задачи исследования – существенный дефицит информации об организации объекта анализа в целом и наличия, тоже время, данных отдельных элементов о других подсистемах собранных с других объектов.

1 Комплексный подход к формированию алгоритмов логического вывода для повышения эффективности информационных систем анализа состояния сложных объектов

Перспективен подход к формализации системы, который позволял бы объединить (дедуктивные) методы достоверного вывода с методами правдоподобного вывода [1]: обобщения (индуктивными), аргументации (абдуктивными).

Практика подтверждает эффективность синтеза различных подходов вывода в системах анализа, в частности дедуктивно-индуктивным подходом к моделированию систем, по сути, являются аппроксимационные методы [2]. В основе построения баз знаний используют синтез индуктивных и абдуктивных методов логического вывода. Абдукция, как процесс формирования объясняющей гипотезы служит методологической основой построения алгоритмов правдоподобного вывода. Функционально, абдуктивный вывод заключается в принятии решения по выбору оптимального объяснения наблюдения на основе заданной теории [1]. Для данного исследования абдукция интересна как средство решения следующих классов задач: задача распознавания целей и стратегий деятельности субъекта, задача формирования моделей по наблюдениям за объектом, задача накопления и усвоения знаний, описывающих новые явления, моделируя тем самым, использование накопленного опыта для обработки новых знаний.

Индуктивный вывод, в сложных системах, позволяющий строить обобщенные модели знаний, основан на построении некоторого общего правила, на основании анализа конечного множества наблюдаемых фактов. Качество обобщенных моделей зависит от полноты набора фактов, которым он пользуется при формировании гипотез. Процедура, процесс индуктивного вывода сложноформализуем и заключается в машинном построении новых гипотез на основе наблюдаемых фактов. Индукционный вывод позволяет решать следующие классы задач: задача индуктивного формирования понятий, с целью выделения наиболее общих или характерных фрагментов знания; задача машинного обучения, где на основе анализа обучающей выборки даётся прогноз о новых объектах; задача распознавания, заключающаяся в формировании решающего правила, относящего объект к определённому классу.

В настоящее время существуют конструктивные методы автоматического формирования алгоритмов мониторинга и управления структурной динамикой сложных технических объектов (СТО) [3], которые базируются на оперативном формировании операторных цепочек, последовательного отображения измерительных и вычислительных параметров СТО для достижения цели анализа или управления. Данные технологии в частности базируются на обобщённых вычислительных моделях, являющихся развитием недоопределённых моделей.

На принципах индуктивного логического вывода, статистической обработки информации, а также информационных технологиях DM, OLAP, KDD основан целый класс аналитических систем, которые, по мнению аналитиков рынка программных продуктов, составляют существенную часть стоимости сложных технических систем (СТС) в целом [3]. Полученные современные результаты и алгоритмы, позволившие автоматизировать решение таких задач как: синтез схем программ как задача доказательства теоремы в формально-дедуктивной системе, методология проверки на модели (model checing), алгоритмы распараллеливания вывода;

показывают перспективность использования в рамках одной системы комплексного подхода [1], включающего индукции для формирования и оценки правил вывода, и дедукции для прогнозирования перспективного состояния системы.

2 Категорные методы представления и структуризации информации в информационных системах анализа состояния сложных объектов

Комплексная обработка информации в современных информационно-вычислительных системах проходит в несколько этапов: сбор априорной информации и формирование модели данных; структуризация собранных моделей и формирования баз знаний; формирование многомодельной структуры объекта анализа; формирование множества алгоритмов вычисления целевых параметров.

В соответствии с целями исследования построен формально-математический аппарат на основе категорного подхода, инвариантный к видам обрабатываемой информации и этапам обработки данных.

Анализ средств представления данных для формирования многомодельных комплексов показал перспективу использования категорно-функторного аппарата, который, основываясь на гомоморфном (структурно эквивалентном) отображении, позволяет описывать объекты инвариантно их внутренней структуре через морфизмы (отличия) их друг от друга. Предложен формально-математический аппарат на основе категорного подхода для представления данных, основанный на включении сформированной категории обобщённых вычислительных моделей в продукционные системы представления знаний, что позволило с единых позиций подойти к формализации данных и процессов на различных этапах обработки информации.

Как было отмечено выше, конструктивным путём снижения неопределённости в процессах предварительной подготовки данных процедур поддержки принятия решений является синтез дедуктивных, индуктивных и абдуктивных методов логического вывода [3,4,7,8]. Данные методы основаны на, таких общенаучных понятиях, как отношения «общее-частное».

Структура и функционирование аналитических систем – S зависит от следующих информационных сущностей: объекта анализа (СТО)- Q ; цели функционирования аналитической системы – G , определяемой конкретной задачей принятия решения; полимодельного комплекса, задающего структуру системы – M ; среды, определяющей параметры системы – C , а также отношений между данными структурами $R = (r_{Q,M}, r_{Q,C}, r_{Q,G}, r_{M,C}, r_{Q,M}, r_{Q,C})$.

Пусть $d = \langle n_i^d, C_d \rangle$, где n_i^d - имя параметра объекта анализа СТО; C_d - значение параметра n_i^d . Обобщая фреймовую модель представления декларативных знаний представим информационные источники (ИИ) в следующем виде: $D = \langle D^N, D^L, D^K \rangle$, где D база данных; D^N – данные с измерительного тракта системы; D^L – данные, накопленные в процессе мониторинга параметров жизненного цикла объектов анализа данного класса; D^K – данные, аккумулированные в документальных системах данного объекта.

В случае D^N выполняется следующее условие $D^N = \langle X_k^{in} = X_k^{out}, F_j^{X_k^{in} \rightarrow X_k^{out}} = \emptyset \rangle$. D^L - данные представляют собой «чёрный ящик» предметом мониторинга является множество входных и множество выходных параметров, когда по входным и выходным значениям восстанавливается (подбирается) модель (отображение) входных параметров в выходные $D^L = \langle X_k^{in}, X_k^{out}, F_j^{X_k^{in} \rightarrow X_k^{out}} = \emptyset \rangle$; D^K данные (знания), полученные как результат научно-исследовательской и опытно-конструкторской деятельности $D^K = \langle X_k^{in}, X_k^{out}, F_j^{X_k^{in} \rightarrow X_k^{out}} \rangle$. Приняв во внимание, что составляющие системы – программные объекты и взяв за основу систему

продукций, рассмотрим следующую формальную модель $P_i = \langle M, M', R_i, O_i \rangle$, где в качестве множества заданных литералов M_i продукционной системы и множества формируемых литералов M'_i продукционной системы определены обобщённые вычислительные модели, R_i – множество продукций i -го вида, O_i – множество процедур присвоения i -го вида. $M = \langle A, F_M \rangle$, где $A = \{\alpha_i, i = 1, \dots, n\}$ – конечное множество параметров состояния объекта, F_M – конечное множество отношений на множестве параметров из A . $F = \langle f_i, i = 1, \dots, k \mid \alpha \in A \rangle$ отношение на множестве параметров $f = \langle gr(f) \rangle$. Множество всех отображений ω для всех отношений $f \in F_M$, входные $in(\omega) = Z_A^{in}$ параметры для оператора ω , выходные $out(\omega) = Z_A^{out}$ параметры для оператора ω .

В качестве объектов категории M определены вычислительные модели $Ob(M_i^k)$; для каждой пары объектов $Ob(M_i^k)$ и $Ob(M_j^k)$ определены множество, морфизмов $Hom(M_i^k, M_j^k)$; для любой тройки объектов определена композиция $Ob(M_i^k)$, $Ob(M_j^k)$ и $Ob(M_l^k)$; морфизмов $\varphi \in Hom(M_i^k, M_j^k)$ и $\psi \in Hom(M_j^k, M_l^k)$ определена композиция $\varphi\psi \in Hom(M_i^k, M_l^k)$; для каждого объекта $Ob(M_i^k)$ определен единичный морфизм $1_x \in Hom(M_i^k, M_i^k)$.

3 Методы структурирования данных и формирования моделей информационных систем анализа состояния сложных технических объектов

На основе предложенного формального аппарата создана методология целевого формирования полимодельного комплекса, который характеризуется как разнородная сетевая структура с иерархическими элементами. Созданы методы автоматического формирования множества алгоритмов вычисления целевых параметров на основе анализа предварительной обработки данных. Созданы и исследованы алгоритмы реализации разработанных методов индуктивного вывода.

Взяв за основу подход [5], к построению порядка на множестве образцов, сформируем правила структурирования полученной информации: если M_b^k – базис модели, (M_0^k, ψ) – исходная ситуация, а (M_i^k, λ) – производная ситуация, $\lambda \in Hom(M_i^k, M_b^k)$ & $\exists \omega \in Hom(M_0^k, M_i^k)$ – условие сопоставимости то:

$$\exists (\mu \in Hom(M_i^k, M_j^k)) \rightarrow ((M_i^k, \varphi) N (M_j^k, \nu)), \text{ где } N \text{ – операция наследования.}$$

Условие обобщения двух моделей выразится в следующем виде: $(M_i^k, \varphi); (M_j^k, \nu); \exists \eta: M_i^k \rightarrow M_m^k, \exists \mu: M_j^k \rightarrow M_m^k$ & $\exists \eta 1: M_i^k \rightarrow M_{m1}^k, \exists \mu 1: M_j^k \rightarrow M_{m1}^k$ & $\exists (\chi \in Hom(M_m^k, M_{im1}^k)) \rightarrow (M_m^k, \beta)$.

Правило определения частного случая: $(M_i^k, \varphi); (M_j^k, \nu); \exists \eta: M_i^k \rightarrow M_m^k$ & $\exists \mu: M_j^k \rightarrow M_m^k \exists \eta 1: M_i^k \rightarrow M_{m1}^k$ & $\exists \mu 1: M_j^k \rightarrow M_{m1}^k$ & $\exists (\vartheta \in Hom(M_n^k, M_{n1}^k)) \rightarrow (M_n^k, \beta)$.

В результате формируется структура из множества классов моделей, каждая из которых представлена иерархической структурой, элементы которой связаны отношением «общее-частное».

При формировании многомодельных комплексов применяется следующая стратегия оперирования вычислительными моделями [9,10]: модель является объектом дедуктивного вывода и формирует более детальный результат, либо модель является объектом индуктивного вывода для структур более высокой иерархии.

Таким образом, задание цели анализа вызывает активизацию нескольких конкурирующих вычислительных схем $\Upsilon = (\Gamma_1, \Gamma_2, \dots, \Gamma_n)$, формирование которых осуществляется на основе $I(\Upsilon, \gamma_j^k)$ множества игровых ситуаций.

Заключение

Разработаны алгоритмы индуктивного вывода на основе двухэтапной процедуры структурной таксономии: на первом этапе осуществляется предварительная классификация в декартовом пространстве; на втором этапе структурная классификация в компактном пространстве.

Разработанный формально - математический аппарат для представления, анализа и обработки разнородных данных и моделей, включающий в себя методы автоматической структуризации данных и формирования категорных моделей баз знаний, методы анализа свойств и отношений на моделях данных и знаний, методы формальной декомпозиции и агрегирования на элементах структур данных, показал свою эффективность в процессах автоматизации формирования модели городской транспортной системы. Анализ существующих транспортных систем явился основой формирования алгоритмов адаптации целевой системы, что позволило существенно снизить объём ручной доработки данного программного проекта, по сравнению с аналогичными.

Список литературы

- [1] Вагин В. Н. Достоверный и правдоподобный вывод в интеллектуальных системах / В. Н. Вагин, Е. Ю. Головина, А.А. Загорянская, М. В. Фомина. - М.: Физматлит, 2004. – 704 с.
- [2] Батищев В.И. Аппроксимационные методы и системы промышленных измерений, контроля, испытаний, диагностики / В.И. Батищев, В.С. Мелентьев. - М.: Машиностроение-1, 2007. - 393 с.
- [3] Охтилев М. Ю. Интеллектуальные технологии мониторинга и управления структурной динамикой сложных технических объектов / М. Ю. Охтилев, Б. В. Соколов, Р. М. Юсупов. - М.: Наука, 2006. - 410 с.
- [4] Батищев В.И. Категорное представление сложных технических объектов в индуктивных системах логического вывода / В.И. Батищев, Н.Г. Губанов // Проблемы управления и моделирования в сложных системах: Тр. IX Международ. конф. – Самара: СНЦ РАН, 2008. – С. 185-191.
- [5] Стефанюк В.Л. Локальная организация интеллектуальных систем / В.Л. Стефанюк. - М.: Физматлит, 2004. – 328 с.
- [6] Соколов Б.В. Концептуальные основы оценивания и анализа качества моделей и полимодельных комплексов / Б.В. Соколов, Р. М. Юсупов // Изв. РАН Теория и системы управления. 2004. №6. С. 5-16
- [7] Загоруйко Н.Г. Прикладные методы анализа данных и знаний / Н.Г. Загоруйко. - Новосибирск: Изд-во Ин-та математики, 1999. - 270с.
- [8] Потапов А.С. – Распознавание образов и машинное восприятие: Общий подход на основе принципа минимальной длины описания / А.С. Потапов. – СПб.: Политехника, 2007. – 548 с.: ил.
- [9] Губанов Н.Г. Категорный подход при формировании полимодельных комплексов сложных систем / Н.Г. Губанов // Вестник Самар. гос. техн. ун-та. Сер. Технические науки. 2008. Вып. 1 (21) С. 183-185.
- [10] Батищев В.И. Методология оперативной реструктуризации информационных систем анализа состояния сложных технических объектов / В.И. Батищев, Н.Г. Губанов // Проблемы управления и моделирования в сложных системах: Тр. IX Международ. конф. – Самара: СНЦ РАН, 2008. – С. 187-193.

ПРИМЕНЕНИЕ МАРКОВСКИХ МЕТРИК «НАКОПЛЕННОГО ДОХОДА» В ЗАДАЧАХ АНАЛИЗА ЖИВУЧЕСТИ КОРПОРАТИВНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

П.А. Зиновьев, А.В. Мейко

Казанский государственный технический университет им. А.Н. Туполева
420111, Казань, ул. Б.Красная, 55, Татарстан, Россия
razin77@mail.ru
тел./факс: +7 (843) 231-00-56

Ключевые слова: анализ живучести, качество обслуживания, корпоративные ИТ-системы, критически важные приложения, марковские модели, метрики «накопленных доходов», подсистемы «back-end», показатели качества функционирования

Abstract

An approach to evaluation of a corporative IT-systems survivability metrics based on Markov accumulated reward modeling technique is discussed. An application of this approach to “back-end” IT-subsystems dependability and Quality-of-Service (QoS) analysis is presented.

Введение

Обеспечение высокой степени надежности и живучести корпоративных информационных систем (КИС) в течение всего жизненного цикла (ЖЦ) является одной из наиболее сложных и ответственных задач современной системотехники. Вместе с тем, даже простая оценка этих ключевых характеристик КИС представляет собой чрезвычайно серьезную проблему из-за высокой сложности подобных объектов [1, 2].

Функциональная живучесть характеризует эффективность работы КИС с позиций обеспечения заданных значений показателей качества функционирования (ПКФ) в процессе выполнения системой своих основных функций и задач. Именно живучесть представляется наиболее адекватной характеристикой, позволяющей в наилучшей степени оценить все технические и технологические аспекты функциональной надежности (dependability) КИС, которая эксплуатируется в постоянно изменяющейся внешней среде и подвергается в течение всего ЖЦ модернизациям, производимым с целью поддержания ее ПКФ на должном уровне.

Об оценке живучести КИС на основе марковских моделей с доходами

В работах [3, 4] для количественной оценки живучести было предложено использовать ПКФ, основанные на указанных в общем техническом задании (ТЗ) требованиях к выходным характеристикам проектируемой системы. В соответствии с этими требованиями качество функционирования КИС определяется множеством показателей $Q = \{q_1 \dots q_S\}$, которые должны находиться в пределах установленных значений (или диапазонов) для эксплуатационных характеристик. Как правило, конструктивно-технологические ограничения задаются в форме технических требований (ТТ) следующего вида:

$$(1) \quad q_j \geq q_j^{TT} \quad \text{или} \quad q_j \leq q_j^{TT}, \quad j = \overline{1, S},$$

которые формализуют условия соответствия ПКФ данного объекта (т.е. КИС) определенным в ТЗ требованиям к его выходным характеристикам. Параметры q_j^{TT} , $j = \overline{1, S}$ представляют собой предельно допустимые граничные значения соответствующих ПКФ системы.

Количественная оценка живучести сложной системы может быть сделана на основе метрик, характеризующих потерю функциональности на протяжении определенного периода времени. Для анализа живучести и оценки степени деградации КИС предложено использовать нормированные показатели $z_j(k)$, $j = \overline{1, S}$, которые вычисляются как:

$$(2) \quad z_j(k) = a_j \frac{q_j(k) - q_j^{TT}}{q_j^{TT}}, \quad j = \overline{1, l} \quad \text{для ТТ вида} \quad q_j \geq q_j^{TT} \quad \text{или}$$

$$(3) \quad z_j(k) = a_j \frac{q_j^{TT} - q_j(k)}{q_j^{TT}}, \quad j = \overline{l+1, S} \quad \text{для ТТ вида} \quad q_j \leq q_j^{TT}.$$

Здесь a_j - весовой коэффициент, характеризующий степень значимости j -го ПКФ для интегральной оценки качества функционирования системы в целом;

k - количество отказов в системе за рассматриваемый период времени (в том числе с учетом восстановлений).

Если для всех заданных ПКФ в течение рассматриваемого периода имеет место выполнение требований ТТ вида (1), то

$$(4) \quad \min_j z_j(k) \geq 0, \quad j = \overline{1, S} \quad \text{и, следовательно,}$$

степень деградации системы остается в рамках нормы, а значит, на данном временном отрезке она может считаться жизнеспособной.

На практике интерес представляют модели для оценки живучести подсистем КИС с восстановлением вычислительных мощностей, в которых возможен возврат системных кондиций до значений ПКФ, отвечающих ТТ. Для формализации задач построения и анализа подобных моделей, как правило, используется аппарат теории массового обслуживания и конечных цепей Маркова. В [4, 5] нами была рассмотрена модель web-узла КИС как подсистемы front-end, в которой наряду с потоками отказов-восстановлений были одновременно учтены и модельные представления для потоков входных заявок.

В этих условиях подсистема front-end, построенная на базе стойки-шасси web-серверов из N «лезвий», представлялась в виде СМО, содержащей N параллельно включенных обслуживающих устройств, обрабатывающих поступающие от пользователей заявки на обслуживание и в тоже время подверженных процессам отказов и восстановлений. Если считать, что каждый из серверов в системе подвергается пуассоновскому потоку отказов с интенсивностью λ , а восстановление каждого из них осуществляется с интенсивностью μ , то имеет место марковский процесс переходов состояний $\{X(t), t \geq 0\}$ на множестве состояний $E = \{E_0, E_1, \dots, E_N\}$. Граф переходов состояний системы соответствует конечной регулярной марковской цепи. Вероятности нахождения системы в каждом из состояний в переходном режиме обозначим как $P(t) = \{P_0, P_1, \dots, P_N\}$, а в установившемся (стационарном) режиме - как $p = \{p_0, p_1, \dots, p_N\}$ соответственно.

Каждому из состояний системы E_i , $i = \overline{0, N}$ может быть поставлено в соответствие значение некоторого «дохода» или «доходов», характеризующих функциональные возможности исследуемого объекта, т.е. в данном случае – модели подсистемы КИС класса front-end. Тогда общие размеры (объем) доходов, получаемых за определенный период времени эксплуатации системы, определяются вектором ставок «доходов» $R = \{r_0, r_1, \dots, r_N\}$, т.е. величиной удельных доходов в единицу времени, и средним временем пребывания системы в каждом из возможных состояний за указанный период.

Для построения марковских моделей с доходами (ММД) предлагается использовать в качестве «ставок» соответствующие значения ПКФ вида $q_j, j=\overline{1, S}$, вычисленные для каждого состояния из множества E . Применение моделей такого типа, где независимо рассматриваются структурная составляющая модели (цепь состояний) и приписанные к ней качественные характеристики (т.е. ПКФ), позволяет анализировать различные сложные проектные и эксплуатационные ситуации с целью принятия адекватных решений.

Технологические приемы построения ММД, а также подходы к исследованию показателей живучести для таких моделей базируются на предварительном вычислении вектора распределения вероятностей состояний $p = \{p_0, p_1, \dots, p_N\}$. В целях повышения эффективности вычислительных алгоритмов имеет смысл рассматривать в ММД только те p_i , значения которых не меньше заданной величины, например, $\varepsilon = 1 \times 10^{-7}$. Таким образом, из всего множества состояний системы E достаточно исследовать только практически значимое подмножество E_V , подмножество индексов для которого определяется как $I_V = \{i = \overline{0, N} \mid p_i \geq \varepsilon\}$.

Тогда для анализа характеристик функциональной живучести деградирующих сложных систем можно применять усредненные оценки значений ПКФ (мат. ожидания уровней доходов) в следующем виде:

$$\tilde{q}_j = \sum_{i \in I_V} q_i^{(j)} p_i,$$

где $q_i^{(j)}$ – значение j -го показателя качества, определенное для i -го состояния системы.

Подставляя полученные значения \tilde{q}_j в (2) и (3) вместо $q_j(k)$, получим величины средних оценок $\tilde{z}_j, j = \overline{1, S}$, которые характеризуют «запасы живучести» системы по каждому из заданных в ТЗ ПКФ. В этом случае, исходя из (4), можно формально определить в качестве основного критерия пребывания КИС в пределах «области живучести» выполнение условия:

$$(5) \quad \min_j \tilde{z}_j \geq 0, \quad j = \overline{1, S}.$$

Полезными и наглядными характеристиками ММД представляются также функции распределения уровня доходов, которые фактически определяют вероятности выполнения требований ТЗ вида (1). Исходя из возможных видов указанных ограничений, можно построить два типа таких функций:

$$(6) \quad P(q_j \geq x) = \left\{ \sum_{i \in I_V} p_i \mid q_i^{(j)} \geq x \right\}, \quad j = \overline{1, l} \quad \text{для ТТ вида } q_j \geq q_j^{TT} \quad \text{или}$$

$$(7) \quad P(q_j \leq x) = \left\{ \sum_{i \in I_V} p_i \mid q_i^{(j)} \leq x \right\}, \quad j = \overline{l+1, S} \quad \text{для ТТ вида } q_j \leq q_j^{TT}.$$

Функции вида (6) и (7) могут использоваться для анализа распределения мгновенных значений ПКФ и оценки вероятности выполнения условий живучести. Исследование данных функций позволяет ответить на ключевой вопрос с точки зрения живучести: какова вероятность того, что на рассматриваемом отрезке жизненного цикла (месяц, квартал, год) система будет функционировать на определенном качественном уровне с показателями эффективности не хуже заданных в ТЗ.

Применение метрик «накопленного дохода» для анализа живучести КИС

Дальнейшим развитием предложенного подхода к анализу живучести является оценка численных значений метрик, опирающихся на расчеты интегральных характеристик, которые

определяются в процессе их накопления в течение некоторого периода эксплуатации Тпл. В теории марковских процессов для этих целей используется понятие «накопленного дохода» (accumulated reward) [6, 7]. Модели данного типа могут применяться для анализа живучести подсистем КИС класса back-end, т.е. крупных Центров обработки данных, построенных на основе кластеров серверов приложений, БД-серверов и сетей хранения.

В свете вопроса о живучести подсистем класса back-end целесообразно выделить следующие категории (виды) приложений в зависимости от степени их ответственности:

- критически важные;
- ответственные (приоритетные);
- рутинные (неприоритетные);
- фоновые.

Поскольку для таких подсистем оценку живучести целесообразно осуществлять с точки зрения их способности выполнять критически важные приложения, то существенное значение имеет как раз фактор потенциальной возможности системы выполнить заданный объем работ (иначе говоря, приложений) в течение определенного планового периода.

Очевидно, что в системах с высокой степенью живучести даже при выходе из строя в результате отказов части оборудования должна сохраняться безусловная гарантированная реализуемость приложений первой группы и высокая вероятность успешного выполнения как можно большего числа приложений из второй группы в соответствии с заданными значениями показателей качества обслуживания QoS. Иначе говоря, постепенная деградация технических характеристик аппаратных средств системы не должна приводить к функциональным последствиям неотвратимого характера с точки зрения выполнения наиболее ответственного прикладного программного обеспечения. Система, которая не может выполнять критически важные бизнес-приложения, не отвечает требованиям ТЗ и не может считаться жизнеспособной.

Пусть в некотором i -м состоянии ММД ставка «дохода» составляет r_i (ед/час). При этом в качестве такого показателя могут выступать, например, удельные затраты на восстановление системы после отказов или стоимость единицы компьютерного ресурса. Обозначим через \tilde{t}_i среднее время пребывания системы в i -м состоянии в течение планового периода $[0, t)$, отсюда для переходного режима имеем:

$$\tilde{t}_i = \int_0^t P_i(t) dt.$$

Тогда накопленный за период $[0, t)$ «доход» будет представлять собой интегральную функцию от изменяющихся во времени ставок доходов для марковского процесса $\{X(t), t \geq 0\}$:

$$(8) \quad D(t) = \int_0^t r[X(t)] dt.$$

Математическое ожидание накопленного за эксплуатационный период $[0, t)$ «дохода» для переходного режима вычисляется следующим образом:

$$(9) \quad M[D(t)] = \sum_{i \in I_V} \int_0^t r_i P_i(t) dt.$$

Для стационарного режима, чем выше вероятность p_i пребывания системы в некотором i -м состоянии, тем выше ожидаемая доля накопленного в данном состоянии «дохода» d_i . Среднее время пребывания системы в i -м состоянии в установившемся режиме в течение планового периода $T_{пл}$ определяется как:

$$\tilde{t}_i = T_{пл} \cdot p_i.$$

Тогда доля накопленного в данном состоянии «дохода» в условиях стационарного режима за период $T_{пл}$ должна составить:

$$d_i = r_i \cdot \tilde{\tau}_i = r_i \cdot T_{nl} \cdot p_i .$$

А ожидаемая суммарная величина общего «дохода», накопленного в течение планового эксплуатационного периода, может быть получена из выражения:

$$(10) D_R = \sum_{i \in I_V} d_i = T_{nl} \sum_{i \in I_V} r_i p_i .$$

Последнее выражение может быть использовано для определения мат.ожиданий таких полезных характеристик исследуемой системы как общие затраты на ремонт и восстановление отказавших устройств в течение заданного периода эксплуатации, объем полезных работ, выполненных в течение некоторого планового периода, объем прибыли и т.д.

Предложенный подход, в основе которого лежат интегральные оценки вида (8)-(10), можно также модифицировать таким образом, чтобы оценки запасов живучести (5) рассчитывались через математические ожидания количества выполненных в полном объеме критически важных и ответственных приложений.

Пример вычисления метрик «накопленного дохода» для подсистем КИС

Рассмотрим подсистему КИС класса back-end, состоящую из 2048 серверных «лезвий», которую обслуживают l ремонтных бригад. Пусть среднее время наработки на отказ каждого сервера составляет от 3000 до 7000 часов, а среднее время восстановления – от трех до пяти часов. Тариф на восстановление одного сервера $r = 100$ руб/час. В качестве примера анализа показателей живучести приведем расчетные графики функций распределения вероятностей не превышения некоторого уровня «накопленного дохода» (затрат на ремонт) при выполнении фиксированного количества приложений в течение планового периода $T_{nl} = 720$ час (рисунок 1). С помощью графиков можно оценить вероятность того, что системные эксплуатационные расходы за указанный период не превысят установленный лимит.

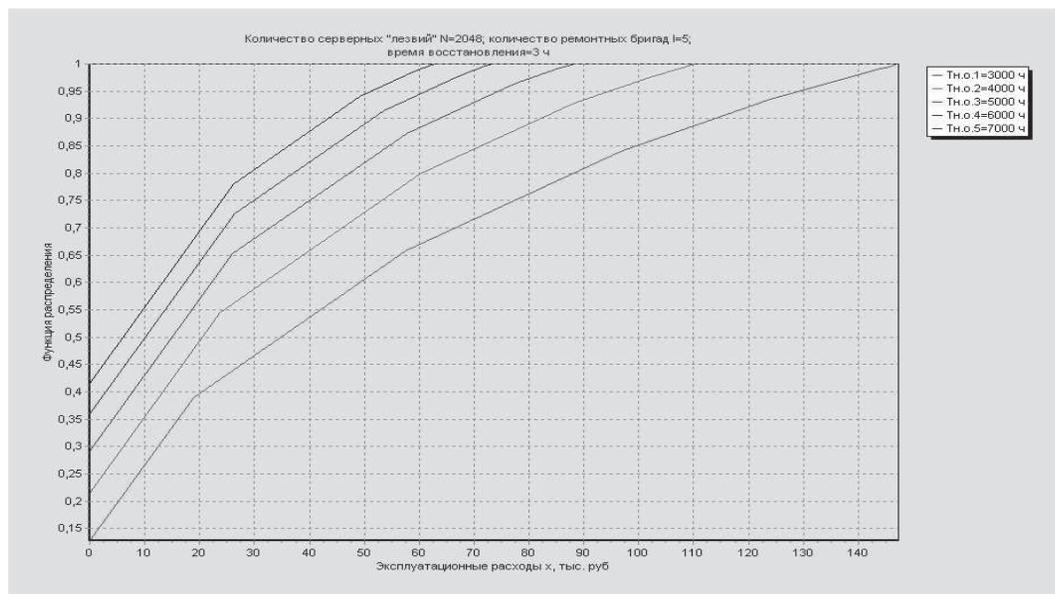


Рисунок 1 – Функция распределения «накопленного дохода» - затрат на ремонт ($t_{восст} = 3$ час)

Заключение

Создание крупномасштабных КИС, сохраняющих необходимый уровень качества функционирования в условиях неизбежной деградации, остается одной из актуальных проблем для разработчиков ИТ-систем. Для ее решения необходимо не только создание эффективных вычислительных моделей таких систем, но и методов их анализа. Представленный подход является одним из возможных вариантов решения данной проблемы. Численные эксперименты осуществлялись для ММД с числом состояний от 2^5 до 2^{16} с помощью специализированной программы SSMP (System Survivability Modeling Program) v.1.2. При этом мощность подмножества эффективных состояний E_V в рамках ММД не превышала 3×10^2 .

Список литературы

- [1] Додонов А.Г., Кузнецова М.Г., Горбачик Е.С. Введение в теорию живучести вычислительных систем. – Киев: Наукова думка, 1990.
- [2] Nicola V.F., Shahabuddin P., Nakayama M. Techniques for the Fast Simulation of Highly Dependable Systems // IEEE Trans. on Reliability. – 2001. – Vol.50. – №3. – P.246-264.
- [3] Зиновьев П.А. Анализ факторов и механизмов живучести в корпоративных информационных системах // В сб.: Исследования по информатике. Вып.12. - Казань: Отечество, 2007. - С.3-30.
- [4] Зиновьев П.А., Мейко А.В. Анализ живучести корпоративных информационных систем на основе модельных представлений // Проблемы управления и моделирования в сложных системах: Труды X Международной конференции. - Самара: Самарский НЦ РАН, 2008. - С.200-207.
- [5] Зиновьев П.А., Мейко А.В. Применение Марковских моделей с доходами для оценки живучести корпоративных информационных систем // Проблемы управления и моделирования в сложных системах: Труды XI Международной конференции. - Самара: Самарский НЦ РАН, 2009. - С.161-166.
- [6] Reibman A., Trivedi K. Transient Analysis of Cumulative Measures of Markov Model Behavior // Journal of Communications Statistic-Stochastic Models. – 1989. – Vol.4. – №5. – P.683-710.
- [7] Trivedi K.S., Malhotra M., Frics R.M. Markov Reward Approach to Performability and Reliability Analysis // In: Proc. MASCOTS-1994, the Second International Workshop. - P.7-11.

МЕТОДИКА УПРАВЛЕНИЯ ВНУТРЕННИМИ РИСКАМИ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В КОРПОРАТИВНЫХ ИНФОРМАЦИОННЫХ СЕТЯХ

И.В. Аникин

Казанский государственный технический университет им. А.Н. Туполева
420111, Казань, ул. К.Маркса, 10, Россия, Республика Татарстан
anikinigor@evm.kstu-kai.ru
тел: +7 (843) 231-00-56

Ключевые слова: корпоративные информационные сети, информационная безопасность, управление информационной безопасностью

Abstract

This article describes methodic of insiders risk management in corporative information systems. Suggested some models for risk management objects. Suggested methods for insiders threat possibility and impact assessment.

Введение

Высочайшая степень автоматизации современных предприятий ставит в зависимость надежность и эффективность их функционирования от степени защищенности информационных систем, обеспечивающих автоматизацию бизнес-процессов. При этом исследования показывают, что большинство угроз информационной безопасности (ИБ) реализуется внутренними сотрудниками, их объем составляет более 2/3 процентов общего числа. Особо остро проблема внутренних угроз стоит в крупных организациях с распределенной архитектурой, обработка информации в которых осуществляется в рамках корпоративных информационных сетей (КИС). В настоящее время обеспечение ИБ КИС часто осуществляется с позиции управления рисками ИБ. Такой подход позволяет формировать наиболее эффективные систем защиты с позиции ожидаемого ущерба.

Риск ИБ определяется как функция (1), зависящая от возможности $P(T_i, V_i)$ реализации угрозы $T_i \in T$ через конкретные уязвимости $V_i \in V$, и от ущерба $IMPACT_i$, понесенного в результате реализации данной угрозы.

$$(1) \text{ Risk}_i = R(P(T_i, V_i), IMPACT_i).$$

Такой подход к оценке рисков ИБ делает актуальным разработку методов достоверной оценки факторов риска ИБ – возможности реализации угрозы P и ущерба $IMPACT$.

Проведенный анализ показывает, что определение точного уровня риска ИБ часто не представляется возможным. В связи с этим, на практике широкое распространение получили методы приближенной оценки рисков, основанные на экспертных оценках. К сожалению, большинство из них не учитывает специфику внутренних рисков ИБ, источником которых являются внутренние нарушители. Данные риски представляют наибольшую опасность для современных КИС, в связи с чем, задача управления ими становится особенно актуальной.

Постановка задачи управления риска заключается в разработке методов достоверного оценивания факторов внутренних рисков ИБ (вероятности реализации угроз и объема наносимого ущерба), а также стратегии снижения внутренних рисков до приемлемого уровня. При решении данной задачи необходимо учесть следующие особенности.

1) Источником реализации внутренних угроз ИБ является человек. В связи с этим возникает проблема учета человеческого фактора при оценке вероятности их реализации.

- 2) Наносимый угрозой ущерб часто имеет качественный характер, который требуется оценить в количественном виде. Кроме этого величина ущерба во многом определяется характером взаимодействия активов КИС, который должен быть учтен методом оценки. В данной статье предлагается методика управления внутренними рисками ИБ, учитывающая данные особенности.

1 Разработка модели КИС

Разработанная модель КИС включает в себя информационную модель $M_{инф}$, множество аппаратных активов $A_{АО}$, логическую структуру $G_{ЛС}$, модель ИТ-сервисов $M_{ИТС}$.

Множество аппаратных активов КИС включает в себя множества АРМ $A_{АРМ}$, серверов $A_{серв}$, телекоммуникационного оборудования $A_{ТК}$.

Логическая структура КИС представляется в виде графа $G_{ЛС} = (V, E)$ с двухцветной раскраской вершин и ребер. Вершинами графа являются взаимодействующие сегменты КИС и телекоммуникационное оборудование, а ребрами – контролируемые и неконтролируемые каналы связи. Множество каналов связи $A_{КСвязи}$ определяется ребрами графа $G_{ЛС}$ с петлями в вершинах, соответствующих сегментам сети.

Для формализации модели ИТ-сервисов введены множества $A_{ИТС}$ реализуемых ИТ-сервисов и ориентированный граф (лес) $G_{ИТС}$, представляющий деревья зависимостей ИТ-сервисов. Каждому из элементов $a_{ИТС} \in A_{ИТС}$ поставлен в соответствие граф $G_{пред_ИТС}$ представления ИТ-сервиса, являющийся ориентированным подграфом $G_{ЛС}$ с возможными петлями. Определены два бинарных отношения: $R_{ИТС}^{АО} \subseteq A_{ИТС} \times A_{АО}$ – между ИТ-сервисами и реализующим их оборудованием, и $R_{ИТС}^{инф} \subseteq A_{ИТС} \times A_{инф}$ – между ИТ-сервисами и предоставляемыми информационными активами.

Информационная модель КИС базируется на введенной модели IP информационных потоков и тернарном отношении $R_{инф}^{АО}$, определяющем характер работы оборудования с информацией (хранение или обработка). Каждый из информационных потоков модели IP представляет собой тройку элементов $\{a_{инф}, a_{ИТС}, Path\}$, где $a_{ИТС}$ – ИТ-сервис, предоставляющий ресурс $a_{инф}$, $Path = \{p_j\}_{j=1}^{N_p}$ – множество цепей в графе $G_{ЛС}$, представляющих собой информационные потоки для $a_{инф}$ от отправителя к получателю..

Модель угроз КИС определена в виде трехдольного графа $G_{угр}$, вершинами которого являются элементы множества уязвимостей $A_{уязв}$, угроз $A_{угр}$, объектов защиты $A_{ОЗ} = A_{инф} \cup A_{АО} \cup A_{КСвязи}$. Ребра графа $G_{угр}$ между угрозами и активами помечены элементами множества $\{cc, ci, cd\}$, где cc соответствует хищению информации, ci - модификации или отрицанию подлинности, cd - уничтожению или блокированию актива.

2 Оценка уровня ущерба от реализации угроз ИБ

Уровень ущерба от реализации угроз ИБ определяется критичностью активов КИС, на которые воздействуют угрозы. Предлагается определять следующие категории активов:

- конфиденциальность $cc_{инф}$, целостность $ci_{инф}$, доступность $cd_{инф}$ для информационных ресурсов;
- критичность $cc_{АРМ}$ для АРМ;
- критичность $cc_{серв}$ и доступность $cd_{серв}$ для серверов;
- доступность $cd_{ТК}$ для телекоммуникационного оборудования;
- доступность $cd_{ИТС}$ для ИТ-сервисов.

Для определения категорий критичности информационных активов введены частные показатели ущерба $ЧП_j^K, j=1, \dots, p^K$ – для конфиденциальности; $ЧП_j^U, j=1, \dots, p^U$ – для целостности; $ЧП_j^D, j=1, \dots, p^D$ – для доступности. Данные показатели характеризуют различные виды ущербов при нарушении соответствующих свойств информационных активов.

Одной из основных проблем количественной оценки категорий критичности является качественный характер большинства частных показателей ущерба. В этих условиях для формирования оценок предложено использовать метод анализа иерархий. Сформирована четырехуровневая иерархия, используемая для определения весовых коэффициентов критичности информационных активов и разработан алгоритм «Экспертная оценка категорий» для количественной оценки их уровней критичности.

Критичность АРМ определяется согласно выражению $cc_{АРМ} = \sum cc + \sum ci$, где cc, ci – категории конфиденциальности и целостности информационных активов, обрабатываемых на АРМ.

Доступность ИТ-сервиса определяется согласно (2).

$$(2) \quad cd_{ИТС} = \max(cd_{ИТС \text{ эксп}}, \sum cd_{инф}, \sum cd_{ИТС \text{ tree}}),$$

где $cd_{ИТС \text{ эксп}}$ – формируется путем оценки частных показателей ущерба, связанных с нарушением доступности ИТ-сервиса, по алгоритму «Экспертная оценка категорий»;

$\sum cd_{инф}$ – сумма уровней доступности информационных активов, предоставляемых ИТ-сервисом;

$\sum cd_{ИТС \text{ tree}}$ – сумма уровней доступности зависимых ИТ-сервисов, определяемая согласно $G_{ИТС}$.

Определение уровней доступности ИТ-сервисов осуществляется в строго определенном порядке. В первую очередь определяется доступность тех, которые являются висячими вершинами дерева зависимостей ИТ-сервисов $G_{ИТС}$. В последнюю очередь определяется доступность ИТ-сервисов, влияющих на другие.

Критичность сервера определяется в виде $cc_{серв} = cc + ci$. Доступность сервера определяется в виде

$$cd_{серв} = \sum_{j | (a_{ИТС}^j, a_{серв}) \in R_{ИТС}^{AO}} (cd_{ИТС}^j).$$

Доступность телекоммуникационного оборудования определяется согласно выражению

$$cd_{ТК} = \sum_{j | a_{ТК} \in G_{пред_ИТС}^j} (cd_{ИТС}^j)$$

На основе уровней критичности активов, можно определить уровень ущерба, наносимого реализацией угрозы. Множество активов КИС, на которые воздействует угроза $threat$, определяется согласно (3):

$$(3) \quad A_{threat} = \{a^j | \exists e_{угр}, e_{угр}(threat, a^j)\}$$

где $e_{угр}$ – ребро в графе $G_{угр}$, связывающее угрозу $threat$ и актив a^j .

Множество A_{threat} является объединением трех подмножеств $A_{threat} = A_{threat}^{CC} \cup A_{threat}^{CI} \cup A_{threat}^{CD}$, включающих в себя множество активов КИС, для которых угроза $threat$ нарушает конфиденциальность, целостность, доступность соответственно. Один и тот же актив a^j может принадлежать одновременно нескольким подмножествам $A_{threat}^{CC}, A_{threat}^{CI}, A_{threat}^{CD}$.

Обозначим через $c.R_{CЗИ}^{ytp}(threat, cm)$ уровень снижения ущерба от реализации угрозы $threat$ путем реализации защитной меры cm , представляющей собой контур системы защиты информации (СЗИ). Обозначим через $cc_{threat}, ci_{threat}, cd_{threat}$ частные степени влияния угрозы на КИС при нарушении конфиденциальности, целостности и доступности активов соответственно. Тогда степень влияния угрозы $threat$ с учетом реализованных защитных мер и многоконтурной СЗИ определяется согласно (4).

$$(4) \quad IMPACT_{threat} = cc_{threat} + ci_{threat} + cd_{threat},$$

$$\text{где } cc_{threat} = \sum_{a^j \in A_{threat}^{CC}} cc_{a^j} \cdot \prod_{cm} (1 - c.R_{CЗИ}^{ytp}(threat, cm)),$$

$$ci_{threat} = \sum_{a^j \in A_{threat}^{CI}} ci_{a^j} \cdot \prod_{cm} (1 - c.R_{CЗИ}^{ytp}(threat, s))$$

$$cd_{threat} = \sum_{a^j \in A_{threat}^{CD}} cd_{a^j} \cdot \prod_{cm} (1 - c.R_{CЗИ}^{ytp}(threat, cm))$$

3 Оценка возможности реализации угроз

Обозначим через $P_{реал}$ возможность реализации угрозы $threat$ внутренним нарушителем КИС. Ее оценку предлагается осуществлять в виде $P_{реал} = P_{возн} \cdot P_{усп}$, где $P_{возн}$ - возможность возникновения условий реализации угрозы (принятия решения внутренним нарушителем о ее реализации); $P_{усп}$ - вероятность успешной реализации угрозы после возникновения данных условий. При оценке показателей $P_{возн}$ и $P_{усп}$ требуется учесть человеческий фактор в качестве источника реализации угроз.

Значение $P_{возн}$ предлагается определять согласно (5).

$$(5) \quad P_{возн} = L_M \cdot L_V \cdot L_K,$$

где $L_M \in [0;1]$ – уровень мотивации нарушителя при реализации угрозы; $L_V \in [0;1]$ – уровень враждебности нарушителя; L_K - уровень осведомленности нарушителя о КИС и средствах защиты. Показатели L_M, L_V, L_K определяются для каждого из пользователей КИС $u \in U$.

Уровень враждебности L_V пользователя определяется через уровень его соответствия нечеткой модели «Потенциальный нарушитель». Модель предложена автором в [8] и представляется формально в виде (6).

$$(6) \quad L(x_1, \dots, x_N) = \langle X, W, \beta, \rho \rangle,$$

где: $X = \{x_1, \dots, x_N\}$ – лингвистические переменные, определяющие свойства потенциального нарушителя;

β – синтаксис лингвистического описания, представляемый в виде грамматики $G = (V_N, V_T, P, S)$ с множествами

- терминальных символов $V_T = \{x^i \text{ есть } L_j^i\}$, где L_j^i – значение лингвистической переменной x^i ;
- нетерминальных символов $V_N = \{S\}$;
- правил вывода $S \rightarrow (x^i \text{ есть } L_j^i)$; $S \rightarrow ((x^i \text{ есть } L_j^i) \text{ И } S)$; $S \rightarrow ((x^i \text{ есть } L_j^i) \text{ ИЛИ } S)$; $S \rightarrow mS$ (где m – модификаторы ОЧЕНЬ, ДОВОЛЬНО, НЕ, БОЛЕЕ ЧЕМ)
- $\rho = \{\rho_i\}_{i=1}^N$ – процедуры оценки значений признаков x^i ;
- $W = (w_1, \dots, w_N)$ – весовые коэффициенты признаков.

В качестве свойств потенциального нарушителя использованы лингвистические переменные «Возраст», «Объем атакующих средств на узле», «Частота использования атакующих средств», «Частота посещения хакерских сайтов», «Уровень любознательности в сети», «Количество регистраций на хакерских форумах», «Количество подписок на хакерские новости и новости по ИБ», «Количество «нехороших» поисковых запросов в единицу времени», «Уровень владения ПК».

Оценка степени соответствия пользователя КИС нечеткой модели осуществляется согласно алгоритму, предложенному в [8].

Мотивация нарушителя L_M определяется в виде отношения $L_M = \frac{Benefits}{Cost}$, где

$Benefits$ – выгоды нарушителя от реализации угрозы, $Cost$ – затраты на реализацию угрозы нарушителем. В качестве факторов, определяющих выгоды нарушителя, предлагается рассматривать моральное удовлетворение и материальные выгоды. К факторам, определяющим затраты нарушителя, были отнесены временные ресурсы, риск наказания, трудоемкость технической реализации угрозы. Определение приоритетов выгод $w_{threat}^{Benefits}$ и затрат w_{threat}^{Cost} нарушителей при реализации угрозы $threat$ осуществляется с помощью метода анализа иерархий [9]. Далее экспертным путем определяется значение $Benefits \in [0;1]$ для группы сотрудников, обладающих максимальной категорией выгод, и $Cost^i \in [0;1]$ для группы сотрудников, обладающих максимальной категорией затрат. Мотивация остальных сотрудников рассчитывается следующим образом:

$$Benefits = \frac{w_{threat}^{Benefits} \cdot Benefits}{w_{threat}^{Benefits}}, \quad Cost = \frac{w_{threat}^{Cost} \cdot Cost}{w_{threat}^{Cost}}, \quad L_M = \frac{Benefits}{Cost}.$$

Уровень осведомленности $L_K^j \in [0;1]$ сотрудников о КИС и средствах защиты также предлагается определять с помощью метода анализа иерархий.

Вероятность $P_{\acute{o}íáá}$ успешной реализации угрозы $threat_i$ от сотрудника u определяется для многоконтурных СЗИ согласно (7).

$$(7) \quad P_{ycn} = \begin{cases} \max_{\substack{j, connect_{G_{yep}}(i, j), \\ Qual_u > diff^j}} \left\{ P^{ij} \cdot \prod_k (1 - R_{C3I}^{y336}[cm_k, j]) \cdot \prod_k (1 - p \cdot R_{C3I}^{y3p}[cm_k, i]) \right\}, \\ 0, \text{ если } Qual_u \leq diff^j \end{cases}$$

где $Qual_u \in [0;1]$ – уровень квалификации пользователя u характеризующий его возможность использовать уязвимости различного уровня сложности. Для оценки уровня квалификации используется подход, основанный на методе анализа иерархий;

$diff_j \in [0;1]$ – сложность использования уязвимости $vuln_j$, определяемая экспертным путем;

cm_k – защитная мера, представляющая собой контур СЗИ;

$R_{СЗИ}^{уязв}[cm_k, j] \in [0;1]$ – уровень снижения вероятности использования уязвимости $vuln_j$ в результате внедрения защитной меры cm_k ;

$p \cdot R_{СЗИ}^{уязв}[cm_k, j] \in [0;1]$ – уровень снижения вероятности реализации угрозы $threat_i$ в результате внедрения защитной меры cm_k ;

$connect_{G_{уязв}}(i, j) \in \{true, false\}$ определяет связность угрозы $threat_i$ и уязвимости $vuln_j$ в трехдольном графе $G_{уязв}$.

Возможность реализации угрозы $threat_i$ в КИС рассчитывается согласно (8).

$$(8) P_{реал} = \max_{u \in U} (P_{реал})_u, \text{ где}$$

$$(P_{реал})_u = (P_{возм})_u \cdot (P_{усл})_u = (L_M)_u \cdot (L_V)_u \cdot (L_K)_u \cdot$$

$$\left\{ \begin{array}{l} \max_{\substack{j, connect_{G_{уязв}}(i,j), \\ Qual_u > diff^j}} \left\{ P^{ij} \cdot \prod_k (1 - R_{СЗИ}^{уязв}[cm_k, j]) \cdot \prod_k (1 - p \cdot R_{СЗИ}^{уязв}[cm_k, i]) \right\} \\ 0, \text{ если } Qual_u \leq diff^j \end{array} \right\}$$

4 Снижение внутренних рисков ИБ

Общий уровень риска по КИС R определяется в виде $R = \sum_{threat} R_{threat}$, где

$R_{threat} = IMPACT_{threat} \cdot P_{реал}$. Эффективность группы защитных мероприятий CM будем рассчитывать согласно (9).

$$(9) F(CM) = \sum_i \frac{\Delta R_{cm_i} - Cost_P_i}{Cost_V_i},$$

где $cm_i \in CM$ – защитная мера, $\Delta R_{cm_i} = R_{cm_i}^{old} - R_{cm_i}^{new}$ – уровень снижения риска ИБ при внедрении защитной меры cm_i , $Cost_P_i$ – стоимость поддержки защитной меры, $Cost_V_i$ – стоимость внедрения защитной меры.

Будем решать на этапе снижения рисков задачу выбора варианта подсистемы информационной безопасности, оптимизированной по критерию «стоимость-эффективность» при заданном уровне остаточных рисков. Формальная постановка данной задачи представлена выражением (10).

$$(10) \begin{cases} F(CM) \rightarrow \max \\ R^{new} \leq R^{порог} \end{cases}$$

где $R^{порог}$ – пороговое значение остаточных рисков в КИС, приемлемое для организации.

Для решения поставленной задачи на множестве защитных мер использовался генетический алгоритм. Решение задачи кодировалось в виде строк-особей вида (1 0 1). Длина строки определяется возможным количеством защитных мер M . Бит «1» означает, что соответствующая мера безопасности используется в КИС, «0» - не используется. В случае, когда

для КИС по требованиям безопасности необходима реализация базового набора защитных мер, работа алгоритма осуществляется на шаблонах с маской вида $1*1***1$, где символы «*» соответствуют битам с переменными значениями. Остальные биты, установленные в 1, должны оставаться неизменными. К ним не применяются операторы мутации на этапе работы генетического алгоритма.

Заключение

Разработанная методика управления внутренними рисками ИБ КИС обладает следующими преимуществами по сравнению с известными.

- 1) Учитывает специфику внутренних рисков ИБ, источником которых является внутренний нарушитель. Человеческий фактор учитывается при оценке вероятности реализации угроз.
- 2) Оценка рисков осуществляется в количественном виде, что позволяет эффективно использовать методику для оценки эффективности СЗИ и решения оптимизационных задач на этапе снижения рисков ИБ.
- 3) Использование формальной модели КИС позволяет учесть взаимодействие активов при оценке ущерба.

Список литературы

- [1] Петренко С.А., Симонов С.В. Управление информационными рисками. Экономически оправданная безопасность. М.: Компания АйТи: ДМК Пресс, 2004. – 384 с.
- [2] Черешкин Д.С., Кононов А.А., Новицкий Е.Г., Цыгичко В.Н. Методика оценки рисков нарушения информационной безопасности в автоматизированных системах. - М.: Институт системного анализа РАН, 1999.
- [3] Кононов А.А., Бурдин О.А. Аксиоматика оценки рисков нарушения информационной безопасности компьютеризированных организационных систем // Проблемы информационной безопасности. Компьютерные системы / Санкт-Петербургский государственный технический университет. 2002. № 1. С. 27-30.
- [4] Аникин И.В. Подход к оценке и анализу рисков информационной безопасности, основанный на использовании метода анализа иерархий // Сборник трудов 3-й ежегодной научно-практической конференции «Инфокоммуникационные технологии глобального информационного общества», Казань, 8-9 сентября 2005. - Казань, 2005, с. 321-332.
- [5] Модели технических разведок и угроз безопасности информации. Коллективная монография / Под ред. Сухарева Е.М. Кн. 3. – М.: Радиотехника, 2003 – 144 с.
- [6] Домарев В.В. Безопасность информационных технологий. Системный подход. – К.: ООО ТИД Диа Софт, 2004. –992 с.
- [7] Хоффман Л.Дж. Современные методы защиты информации: Пер. с англ. // Под ред. В.А. Герасименко. – М.: Сов. Радио, 1980, 264 с.
- [8] Glova V.I., Anikin I.V. Method for Recognition of Fuzzy 2D Primitives via a Technology of Soft Computing // Pattern Recognition and Image Analysis. Vol. 11, No 1, 2001.
- [9] Саати Т. Принятие решение решений. Метод анализа иерархий. – М.: Радио и связь, 1993. – 278 с.

МЕТОДОЛОГИЯ ВЫЯВЛЕНИЯ ВНУТРЕННИХ НАРУШИТЕЛЕЙ В КОРПОРАТИВНЫХ ИНФОРМАЦИОННЫХ СЕТЯХ

И.В. Аникин, В.С. Лукоянов

Казанский государственный технический университет им. А.Н. Туполева
420111, Казань, ул. К.Маркса, 10, Россия, Республика Татарстан
anikinigor@evm.kstu-kai.ru
тел: +7 (843) 231-00-56

Ключевые слова: *корпоративные информационные сети, информационная безопасность, обнаружение нарушителей, Data Mining*

Abstract

This article describes methodology of insiders detection in corporative information systems. Suggested data mining algorithms for forming user profile. Suggested method and intelligent software for detection critical incidences from user account.

Введение

Своевременное выявление попыток нарушения информационной безопасности (ИБ) является одной из важнейших задач для корпоративных информационных сетей (КИС). При этом наибольшую актуальность приобретает выявление нарушений ИБ, реализуемых внутренними сотрудниками организации. Количество данных нарушений в современных КИС составляет более 2/3 общего их числа.

Для выявления нарушений ИБ в КИС используются системы мониторинга событий и обнаружения вторжений. К сожалению, подобные системы не способны качественно выявлять все инциденты, исходящие от внутренних нарушителей. Это в частности связано с тем, что такие нарушения могут не включать стадию, предвещающую атаку. Сложно отличить действия злоумышленника от действий обычного пользователя, если злоумышленник не нарушает установленные правила поведения в системе. В качестве примера подобных нарушений можно привести кражу злоумышленником пароля легального пользователя и получение от его имени доступа к системе. Для обнаружения подобных событий предлагается использовать системы выявления аномального поведения пользователя в КИС.

Системы анализа аномалий (САА) нашли широкое применение при выявлении сетевых атак, в частности отказов в обслуживании [1], однако возможности их использования для выявления внутренних нарушителей исследованы недостаточно хорошо. В последнее время усиливается внимание [2] к обучению САА используя технологии KDD и Data Mining [3]. Однако большинство из существующих работ посвящено использованию подобных САА для выявления сетевых атак.

В статье предлагается метод анализа аномалий для выявления событий, связанных с получением нарушителя доступа к системе под учетной записью другого пользователя. Данный подход построен на технологиях KDD и Data Mining.

1 Общая характеристика используемых алгоритмов

Для выявления указанных событий формируются эталонные профили поведения пользователей КИС и выявляются отклонения от них в ходе работы пользователя в системе. Профиль поведения пользователя формируется путем анализа его работы с прикладными программами

и выявления скрытых закономерностей, характеризующих данную работу. Для формирования профиля используется следующая основная информация:

- 1) время входа/выхода пользователя из системы;
- 2) дата, время запуска/завершения работы различных программ;
- 3) последовательность запуска программ пользователем.

В качестве источника данных о поведении пользователя используется журнал событий Security EventLog ОС Windows.

На рисунке 1 представлена структура разработанной САА. Данная система функционирует в двух режимах.

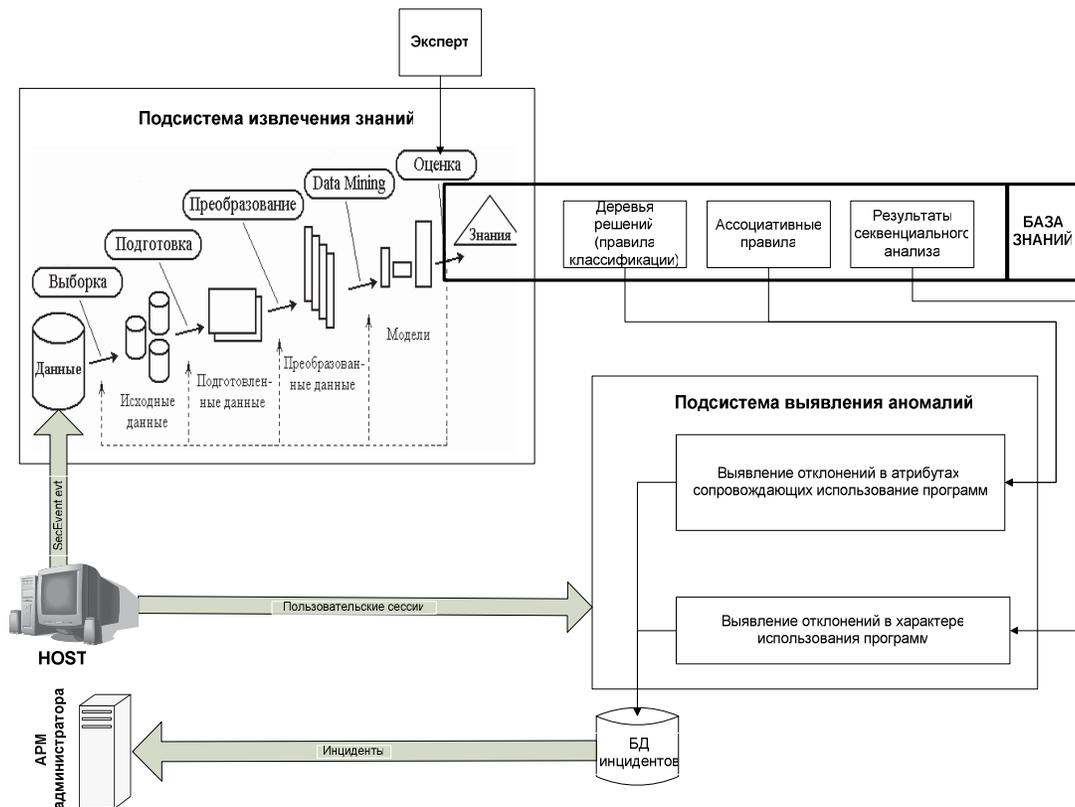


Рисунок 1 – Структура системы анализа аномалий

1. Обучение, в рамках которого формируются эталонные профили поведения пользователей и формирование базы знаний САА. В качестве способов представления знаний используются деревья решений, ассоциативные правила, а также совместно используемые наборы программ, полученные методами секвенциального анализа данных [3].

Деревья решений используются для представления полного набора правил классификации, характеризующих временные сроки использования критичных программ. Данные правила могут свидетельствовать, например, о том, что процедура резервного копирования на узле компьютерной сети запускается в четко установленные сроки.

Ассоциативные правила представляют собой продукционные правила вида $X \rightarrow Y, C, Sup, I$, где X и Y - наборы элементов, $X \cap Y = \emptyset$, C - достоверность, Sup - поддержка, I - улучшение правила. Данные правила используются для представления «полезных» знаний о временных сроках использования программ.

Результаты секвенциального анализа представляют собой частые наборы совместно используемых пользователем программ. Эти знания характеризуют то, в какой последовательности и в каком сочетании запускаются программы пользователем на узле компьютерной сети.

2. Выявление аномального поведения пользователей на узле компьютерной сети. Осуществляется выявление аномалий двух видов:

- отклонения в атрибутах, сопровождающих использование программ. Например, процедура резервного копирования запускается не в установленные сроки;
- отклонения в характере использования программ. Например, после входа в систему не соблюдается последовательность запускаемых программ.

Выявленным инцидентам назначается вес, в зависимости от степени их критичности. Информация о выявленных инцидентах вносится в БД инцидентов. О наиболее критичных инцидентах немедленно извещается администратор безопасности с целью своевременного реагирования на них.

Рассмотрим работу САА в режимах обучения и выявления аномалий.

Формирование эталонных профилей поведения пользователей

Обозначим через $S = \{s_1, \dots, s_q\}$ множество записей в журнале регистрации событий Security EventLog. Каждую из записей s_i охарактеризуем тройкой элементов $\langle d_i, t_i, type_i \rangle$, где $type_i$ – тип события (например, вход в систему, запуск определенной программы и т.д.), d_i – дата фиксации события, t_i – время фиксации события.

Формирование базы знаний САА, осуществляется подсистемой извлечения знаний, при участии эксперта (рисунок 1). Извлечение знаний предполагает реализацию шагов выборки данных, их подготовки, преобразования, получения знаний методами Data Mining, а также оценку полученных знаний экспертом.

Выборка данных. На данном шаге из множества S выделяется подмножество событий $\bar{S} \subseteq S$, $\bar{S} = \{s_1, \dots, s_n\}$, относящихся к запуску/завершению процессов или программ, а также к входу/выходу пользователя из системы.

Подготовка данных для анализа осуществляется путем решения следующих подзадач.

Очистка данных. Реализуется при построении дерева решений путем фильтрации множества \bar{S} и удаления из него событий, не критичных относительно дня недели или времени запуска. Данная задача решается экспертом, а также путем расчета корреляции между фактом возникновения события и данными параметрами.

Удаление избыточных данных. Появление одного события в множестве \bar{S} может породить появление других. В данном случае в \bar{S} оставляется только одно головное событие, которое порождается активностью пользователя.

Удаление нерелевантных ко времени событий, порожденных ошибками пользователя. Критерием удаления события является условие $t_a - t_b < 10$, где t_a – время запуска программы пользователем, t_b – время завершения ее работы.

Склеивание повторяющихся событий. Реализуется при проведении секвенциального анализа. Часто в множестве \bar{S} одинаковые события следуют друг за другом. Для секвенциального анализа данный вариант не представляет интереса, так как работа осуществляется с одной программой. Такие события склеиваются в одно, например, последовательность событий АААВВС преобразуется в последовательность АВС.

Преобразование данных осуществляется путем приведения событий множества \bar{S} к требуемой для анализа форме.

Атрибуты d_i и t_i преобразуются в дискретный вид с множествами значений \aleph_{date} и \aleph_{time} соответственно.

- 1) Множество значений \aleph_{date} атрибута d_i имеет следующие значения: начало недели, середина недели, конец недели, выходные.
- 2) Множество значений \aleph_{time} атрибута «Время» имеет следующие значения: раннее утро, утро, середина дня, конец дня, вечер, ночь, глубокая ночь.

При проведении секвенциального анализа на данном шаге дополнительно выделяются пользовательские сессии (анализируемые транзакции). Точками, разделяющими сеансы работы пользователя в множестве \bar{S} , считаются:

- события входа в систему под исследуемой учетной записью;
- запуск хранителя экрана;
- отсутствие активности в системе более 10 минут.

Data Mining

Для построения *дерева решений* использован алгоритм C4.5 [4]. Формируемое дерево решений представляет полную систему правил классификации, определяющих день и время использования программ на узле компьютерной сети. Если запуск программы не соответствует сформированным правилам классификации, то можно вести речь о возникшем инциденте, связанным с использованием программы.

К недостаткам данного способа представления знаний для САА относится необходимость жесткого отбора на этапе фильтрации программ, критичных относительно дня недели или времени запуска. В ином случае может произойти необоснованное усложнение построенного дерева решений.

Поиск ассоциативных правил

Будем рассматривать в качестве транзакций элементы $s_i \in \bar{S}$. Данные транзакции представляют собой тройки элементов $\langle d_i, t_i, type_i \rangle$, где $d_i \in \aleph_{date}, t_i \in \aleph_{time}, type_i \in \aleph_{type}$

В отличие от построения полной системы правил классификации, в данном случае осуществляется поиск в множестве транзакций небольшого количества продукционных правил вида (1), являющихся «полезными» для выявления рассматриваемых инцидентов.

$$(1) X \rightarrow Y, C, I,$$

где X и Y - набор элементов множеств $\aleph_{date}, \aleph_{time}, \aleph_{type}$, встречающихся в транзакциях совместно, причем совместно встречаются и элементы $X \cup Y$, а $X \cap Y = \emptyset$.

C – достоверность правила, показывающая вероятность того, что из наличия в транзакции набора X следует наличие в ней набора Y . I – улучшение, показывающее, полезнее ли правило случайного угадывания

$$C = \frac{Support(X \cup Y)}{Support(X)}, \quad I = \frac{Support(X \cup Y)}{Support(X) \cdot Support(Y)},$$

где $Support(F) = \frac{|\bar{S}_F|}{|\bar{S}|}$ - поддержка набора F , а \bar{S}_F - количество транзакций, в которые входит набор F .

Ассоциативное правило (1) считается полезным, если выполняются условия:

$$(2) \dot{N} \geq Supp_{\min} \geq Supp_Y, \quad I > 1.$$

Для построения ассоциативных правил в разработанной САА использован алгоритм построения частотных деревьев [5].

Секвенциальный анализ

В данном случае в качестве транзакций рассматриваются пользовательские сессии (3), состоящие из последовательно запускаемых пользователем программ. Формирование пользовательских сессий осуществляется на этапе преобразования данных.

$$(3) \text{ Session} = \{type_1, \dots, type_n\}.$$

Выполняется анализ транзакций (3) путем построения ассоциативных правил и выделения в них частых совместно используемых наборов программ. Конкретный набор из k программ, назовем k -последовательностью. Для их построения в разработанной САА использован алгоритм построения частотных деревьев.

Оценка знаний

На данном этапе осуществляется интерпретация и анализ полученных знаний человеком-экспертом. Для сформированных ассоциативных правил и k -последовательностей осуществляется оценка полезности в соответствии с условиями (2). Для k -последовательностей особое внимание следует уделить тем, которые начинаются с событий {Вход в систему} или {Хранитель экрана}.

2 Оценка критичности выявленных инцидентов

Разобьем выявленные инциденты i_j на три категории.

Инциденты первого рода. Временные сроки и день запуска программы p не соответствуют правилам построенного дерева решений.

Для данных инцидентов определяются расстояния $\rho(p_s)$ до ближайших листьев построенного дерева решений, помеченных p . Данные расстояния определяются на порядковой шкале значениями от 1 до 5. Степень критичности инцидентов первого рода определяется согласно (4)

$$(4) \text{ critical}(i_j) = \text{critical}(p) \circ \min_s \rho(p_s),$$

где $\text{critical}(p) \in \{1, \dots, 5\}$ – степень критичности программы p , определяемая экспертным путем, \circ – операция на порядковых шкалах.

Инциденты второго рода. Не выполняется условие $A \rightarrow B$ построенного ассоциативного правила при удовлетворении условия A .

Степень критичности инцидентов второго рода определяется согласно выражению (5):

$$(5) \text{ critical}(i_j) = \text{critical}(A \rightarrow B) \cdot C,$$

где $\text{critical}(A \rightarrow B)$ – степень критичности ассоциативного правила, определяемая экспертным путем, $C \in [0; 1]$ – достоверность правила.

Инциденты третьего рода. Нарушается последовательность использования программ в k -последовательностях.

Степень критичности инцидентов третьего рода определяется согласно выражению (6):

$$(6) \text{ critical}(i_j) = \text{critical}(I_k) \cdot \min_k \frac{\text{Support}(X \cup Y'_k)}{\text{Support}(X \cup Y_k)}, \text{ где}$$

I_k – k -последовательность базы знаний, начало которой совпадает с событием X ;

Y'_k – теоретическое продолжение k -последовательности в базе знаний;

Y_k – реальное продолжение k -последовательности в пользовательской сессии.

Обозначим через $I_A = \{i_1, \dots, i_k\}$ – множество накопленных инцидентов КИС, порожденных конкретной учетной записью u за конкретный промежуток времени T . Общий уровень инцидентов $C(u)$, исходящих от учетной записи A , предлагается определять согласно (7).

$$(7) \quad C(u) = \sum_{i=1}^k \text{critical}(i_i).$$

Исходя из уровня инцидентов $C(u)$, принимается решение об обнаружении аномалии. Для администратора безопасности уровень инцидентов представляется в цвете путем определения значения показателя $\check{O}\check{C}$ (цвет инцидентов):

$$ЦИ = \begin{cases} \text{зеленый, если } C(u) \leq C^{\text{желтый}}(u) \\ \text{желтый, если } C^{\text{желтый}}(u) < C(u) < C^{\text{красный}}(u) \\ \text{красный, если } C(u) \geq C^{\text{красный}}(u) \end{cases}$$

зеленый – уровень инцидентов, исходящих от учетной записи является приемлемым и не свидетельствует о компрометации учетной записи;

красный – уровень инцидентов, исходящих от учетной записи, является неприемлемым, свидетельствует о компрометации учетной записи. Данный факт требует немедленного реагирования от администратора;

желтый – уровень инцидентов, исходящих от учетной записи, не является приемлемым, однако не является и критичным. Возможно, учетная запись скомпрометирована. Администратору следует обратить внимание на данный факт.

Формирование пороговых уровней инцидентов $C^{\text{желтый}}(u)$ и $C^{\text{красный}}(u)$ осуществляется следующим образом.

Обозначим через $U = \{u_i\}_{i=1}^{N_{\text{польз}}}$ – множество учетных записей КИС.

$C(u_i) \geq 0$ – уровень исходящих инцидентов от учетной записи u_i , вычисляемый согласно (7) на этапе обучения САА. Обучение происходит в условиях, когда учетная запись не скомпрометирована, поэтому значение $C(u_i)$ является допустимым для этой учетной записи.

$K(u_i)$ – множество единиц знаний для учетной записи u_i , сформированное на этапе обучения САА, и определяющее профиль поведения пользователя. Эти единицы знаний представляют собой деревья решений, ассоциативные правила или наборы совместно используемых программ.

После обучения САА производится настройка параметров процедуры принятия решений путем формирования пороговых уровней $C^{\text{желтый}}(u_i)$ и $C^{\text{красный}}(u_i)$ для каждого из пользователей $u_i \in U$. Для этого, набирается статистика работы пользователей в тестовом режиме работы САА. Должно выполняться условие $T_{\text{test}} \approx T_e$, где T_{test} – время сбора тестовых данных, T_e – время обучения САА.

Обозначим через $E(u_j, K(u_i))$ – расстояние между профилем поведения пользователя u_i (согласно $K(u_i)$) и поведением пользователя u_j на этапе тестирования. Данное расстояние вычисляется согласно (4)-(6).

Используя информацию о поведении пользователя u_i в дискретные моменты времени $0 < T_i \leq T_{\text{test}}$, на этапе тестирования формируются и нормируются распределения расстояний

$P_i^t(r)$ от собственного профиля поведения пользователя, и $P_i^f(r)$ от профилей поведения других пользователей.

Тогда на этапе выявления аномалий $P_i^t(r)$ является вероятностью того, что степень критичности инцидента – зеленая, а $P_i^f(r)$ – вероятностью того, что степень критичности инцидента – красная.

Вычисление пороговых уровней инцидентов $C^{\text{желтый}}(u)$ и $C^{\text{красный}}(u)$ осуществляется согласно (8)

$$(8) \quad C^{\text{желтый}} = r, \text{ при котором } \frac{P_i^t(r)}{P_i^f(r)} = 2, \quad C^{\text{красный}} = r, \text{ при котором } \frac{P_i^t(r)}{P_i^f(r)} = \frac{1}{2}$$

Для администратора безопасности палитра и цвет инцидентов определяется согласно (9)

$$(9) \quad \text{Color} = \left(\overline{P_i^f(r)} \cdot 255, \overline{P_i^t(r)} \cdot 255, 0 \right), \text{ где } \overline{P_i^t(r)} = P_i^t(r) \cdot k, \quad \overline{P_i^f(r)} = P_i^f(r) \cdot k,$$

$$k = \max \left\{ \frac{1}{P_i^t(r)}, \frac{1}{P_i^f(r)} \right\},$$

где $\overline{P_i^f(r)} \cdot 255$ – уровень красного цвета в палитре, $\overline{P_i^t(r)} \cdot 255$ – уровень зеленого цвета в палитре.

3 Результаты экспериментов

Разработанный метод выявления аномалий и модели интеллектуального анализа данных реализованы в программном комплексе обнаружения внутренних нарушителей на узле компьютерной сети. Данный программный комплекс опробован для выявления аномального поведения пользователей на компьютерных системах с ОС Windows. Проведенные эксперименты показали эффективность предложенного метода для выявления инцидентов, связанных с несанкционированным получением доступа злоумышленника к системе. В ходе экспериментов уровень ошибок классификации первого рода получился равным 0,2, уровень ошибок классификации второго рода равным 0,1.

Заключение

Практическое значение предложенного подхода заключается в возможности эффективно выявления внутренних нарушений информационной безопасности, которые трудно выявляются классическими системами обнаружения вторжений. Проведенные эксперименты показывают приемлемый уровень ошибок классификации первого и второго рода.

Список литературы

- [1] Wang H., Zhang D., Shin K. Detection SYN flooding attacks // Proceedings of the IEEE Infocom Conference, 2002, pp. 1530-1539.
- [2] Brigger S.T. Data mining methods for network intrusion detection. University of California, Davis, 2004.
- [3] Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining. – СПб.: БХВ-Петербург, 2004.
- [4] Quinlan J.R. C4.5: Programs for Machine learning. Morgan Kaufmann Publishers 1993.
- [5] Han J, Pei J., Mortazavi-Asl B., Chen Q., Dayal U., Hsu M. FreeSpan: frequent pattern-projected sequential pattern mining. Proceedings of the 2000 international conference on knowledge discovery and data mining (KDD'00), Boston, MA, August 2000.



Control and Measurement in Complex Technical Systems
**Управление и измерения в сложных
технических системах**

УПРАВЛЕНИЕ ПЕРЕВЁРНУТЫМ МАЯТНИКОМ С ПОДВИЖНОЙ ТОЧКОЙ ПОДВЕСА

Ю.Н. Золотухин, А.А. Нестеров

Институт автоматки и электрометрии СО РАН
630090, Новосибирск, пр. ак. Коптюга, 1, Россия
zol@idisys.iae.nsk.su
тел: +7 (383) 333-26-25, факс: +7 (383) 333-38-63

Ключевые слова: перевёрнутый маятник, подвижная точка подвеса, управление положением, управление скоростью точки подвеса

Abstract

The new control algorithms of inverted pendulum position and platform velocity are presented.

Введение

В связи с разработкой нового вида транспортных средств, представляющих собой двухколёсную пассажирскую платформу (системы типа Segway) [1], возродился интерес к классической задаче стабилизации положения перевёрнутого маятника с подвижной точкой подвеса [3-7]. В докладе представлен алгоритм, решающий как задачу стабилизации маятника в заданном относительно вертикали положении, так и задачу стабилизации скорости платформы, несущей маятник.

Модель

Схема исследуемого устройства представлена на рисунке 1.

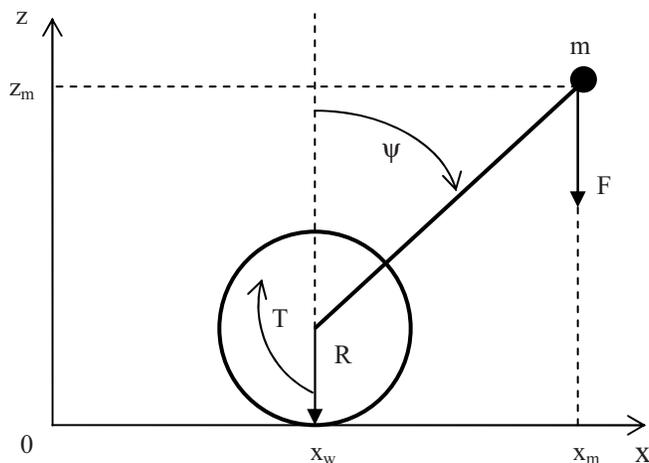


Рисунок 1 – Кинематическая схема (R – радиус колеса, L – расстояние от центра масс маятника до точки подвеса, (x_w, R) – координаты на плоскости (x, z) центра колеса, (x_m, z_m) – координаты центра масс маятника, M_w – момент вращения, приложенный к колесу, $F = -mg$ – сила тяжести, приложенная к маятнику)

Динамические параметры системы: J_w – момент инерции колеса, m - масса маятника, m_w – масса колеса. В качестве обобщённых координат механической системы удобно выбрать ψ – угол отклонения маятника от вертикали и ϑ – угол поворота колеса. Будем считать, что колесо движется вдоль оси x без проскальзывания, а силы трения в системе пренебрежимо малы. В этом случае координаты колеса и маятника определяются через обобщённые координаты:

$$\begin{aligned} x_w &= R \cdot \vartheta; \\ (1) \quad x_m &= x_w + L \cdot \sin \psi = R \cdot \vartheta + L \cdot \sin \psi; \\ z_m &= R + L \cdot \cos \psi. \end{aligned}$$

Движение системы в обобщённых координатах описывается уравнениями Лагранжа:

$$(2) \quad \frac{d}{dt} \frac{\partial T}{\partial \dot{\vartheta}} - \frac{\partial T}{\partial \vartheta} = \frac{\delta A_{\vartheta}}{\delta \vartheta};$$

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{\psi}} - \frac{\partial T}{\partial \psi} = \frac{\delta A_{\psi}}{\delta \psi},$$

где $T(\psi, \vartheta, \dot{\psi}, \dot{\vartheta})$ - кинетическая энергия системы, δA_{ψ} и δA_{ϑ} - элементарные работы внешних сил F и M_w по координатам ψ и ϑ соответственно. Здесь и далее производные по времени обозначаются точками над символами переменных. Кинетическая энергия системы является суммой энергии движения колеса $T_w = \frac{1}{2} [J_w \cdot \dot{\vartheta}^2 + m_w \cdot \dot{x}_w^2]$ и энергии движения маятника

$T_m = \frac{1}{2} m \cdot [\dot{x}_m^2 + \dot{z}_m^2]$. Суммируя эти слагаемые с учётом (1), получим

$$(3) \quad T = \frac{1}{2} [J_w \cdot \dot{\vartheta}^2 + m_w \cdot R^2 \cdot \dot{\vartheta}^2 + m \cdot R^2 \cdot \dot{\vartheta}^2 + m \cdot L^2 \cdot \dot{\psi}^2 + 2mLR \cdot \dot{\vartheta} \cdot \dot{\psi} \cdot \cos \psi].$$

Для элементарных работ внешних сил получим:

$$(4) \quad \begin{aligned} \delta A_{\vartheta} &= M_w \cdot \delta \vartheta; \\ \delta A_{\psi} &= -mg \cdot \delta z_m = mgL \cdot \delta \psi \cdot \sin \psi. \end{aligned}$$

Подставляя (3) и (4) в (2) и разрешая полученные уравнения относительно вторых производных, получим уравнения движения системы:

$$(5) \quad \begin{aligned} \ddot{\psi} &= \frac{(J_w + (m_w + m) \cdot R^2) \cdot g \cdot \sin \psi - (M_w + mLR \cdot \dot{\psi}^2 \cdot \sin \psi) \cdot R \cdot \cos \psi}{L \cdot (J_w + m_w \cdot R^2 + mR^2 \cdot \sin^2 \psi)}; \\ \ddot{\vartheta} &= \frac{M_w + mLR \cdot \dot{\psi}^2 \cdot \sin \psi - mRg \cdot \sin \psi \cdot \cos \psi}{J_w + m_w \cdot R^2 + mR^2 \cdot \sin^2 \psi}. \end{aligned}$$

Для упрощения дальнейших преобразований введём обозначения:

$$(6) \quad \begin{aligned} f1(\psi, \dot{\psi}) &= \frac{(J_w + (m_w + m) \cdot R^2) \cdot g \cdot \sin \psi - mLR^2 \cdot \dot{\psi}^2 \cdot \sin \psi \cdot \cos \psi}{L \cdot (J_w + m_w \cdot R^2 + m \cdot R^2 \cdot \sin^2 \psi)}; \\ f2(\psi) &= -\frac{R \cdot \cos \psi}{L \cdot (J_w + m_w \cdot R^2 + m \cdot R^2 \cdot \sin^2 \psi)}; \\ f3(\psi, \dot{\psi}) &= \frac{mR \cdot (L \cdot \dot{\psi}^2 \cdot \sin \psi - g \cdot \sin \psi \cdot \cos \psi)}{J_w + m_w \cdot R^2 + m \cdot R^2 \cdot \sin^2 \psi}; \\ f4(\psi) &= \frac{1}{J_w + m_w \cdot R^2 + m \cdot R^2 \cdot \sin^2 \psi}. \end{aligned}$$

С учётом этих обозначений из (5) получаем компактную форму уравнений движения исследуемой системы:

$$(7) \quad \begin{aligned} \ddot{\psi} &= f1(\psi, \dot{\psi}) + f2(\psi) \cdot M_w; \\ \dot{\vartheta} &= f3(\psi, \dot{\psi}) + f4(\psi) \cdot M_w. \end{aligned}$$

Управление положением маятника

Под управлением положением маятника будем понимать перевод маятника из текущего состояния $\psi(t)$ в заданное состояние ψ_{ref} и, при необходимости, удержание его в этом состоянии. Для решения поставленной задачи в нашем распоряжении – единственный управляющий параметр M_w - момент вращения, приложенный к колесу.

Очевидно, можно различными способами выбрать закон изменения момента M_w , решающий нашу задачу, и, следовательно, различные траектории перемещения маятника в заданное положение. Выберем желаемую траекторию в соответствии с законом

$$(8) \quad \ddot{\psi} + K \cdot (\psi - \psi_{ref}) = 0.$$

Реализация движения по траектории (8) гарантирует аperiodическое движение маятника в заданное положение.

Однако в некоторый момент t угол $\psi(t)$ может не удовлетворять уравнению (8). Следовательно, необходимо с помощью изменения момента M_w вывести систему на траекторию (8) и удерживать её на этой траектории до перехода в точку $\psi = \psi_{ref}$.

Оценим близость системы к траектории (8) функцией

$$(9) \quad S(t) = \dot{\psi}(t) + K \cdot (\psi(t) - \psi_{ref})$$

и потребуем выполнения условия

$$(10) \quad \frac{d}{dt}(S^2(t)) \leq 0,$$

причём знак равенства в (10) допустим только при $S(t) = 0$. Выполнение условия (10) гарантирует перевод системы на траекторию (8) и удержание её на этой траектории.

Условие (10) допускает некоторый произвол при определении момента M_w . Поэтому ужесточим условие (10), потребовав выполнения равенства

$$(11) \quad \frac{d}{dt}(S^2(t)) = -2\alpha \cdot S^2(t).$$

Здесь параметр α определяет скорость сходимости решения уравнения (11) к нулю

$$(12) \quad S^2(t) = S^2(t_0) \cdot \exp(-2\alpha t),$$

где $t = t_0$ - некоторое начало процесса управления.

Следует заметить, что при выполнении условия (11) функция $S^2(t)$ удовлетворяет условиям Ляпунова [7] и гарантирует экспоненциально устойчивый переход системы на траекторию (8).

Остаётся определить момент $M_w(\psi, \dot{\psi})$, гарантирующий выполнение условия (11).

Из (11) и (9) получим

$$\ddot{\psi} = -\alpha \cdot S - K \cdot \dot{\psi}$$

и, подставив $\ddot{\psi}$ в первое уравнение (7), определим необходимый закон изменения момента $M_w = M_w(\psi, \dot{\psi})$:

$$(13) \quad M_w(\psi, \dot{\psi}) = -\frac{\alpha \cdot S(\psi, \dot{\psi}) + K \cdot \dot{\psi} + f1(\psi, \dot{\psi})}{f2(\psi)}.$$

Действительно, подстановка (13) в первое уравнение (7) приводит к линейному дифференциальному уравнению второго порядка

$$(14) \quad \ddot{\psi} + (\alpha + K) \cdot \dot{\psi} + \alpha K \cdot \psi = \alpha K \cdot \psi_{ref}$$

с собственными значениями

$$\lambda_1 = -\alpha, \quad \lambda_2 = -K.$$

Очевидно, момент $M_w(\psi, \dot{\psi})$ можно трактовать как нелинейную обратную связь, преобразующую исходную систему в устойчивую линейную систему второго порядка.

До сих пор предполагалось, что значения момента M_w ничем не ограничены. Однако из (6) и

(13) видно, что при $\psi \rightarrow \pm \frac{\pi}{2}$ $f2(\psi) \rightarrow 0$ и $M_w \rightarrow \pm \infty$. Отсюда следует, что при наличии ограничений на величину момента M_w необходимо ограничивать диапазон возможных значений угла ψ .

К точке равновесия $\psi = \psi_{ref}$ система в соответствии с (14) приближается с нулевой скоростью. Следовательно, для удержания маятника в положении $\psi = \psi_{ref}$ необходимо в соответствии с (13) значение момента

$$(15) \quad M_w = -\frac{f1(\psi_{ref}, 0)}{f2(\psi_{ref})} = \frac{g}{R} (J_w + m_w \cdot R^2 + m \cdot R^2) \cdot \text{tg} \psi_{ref}.$$

Если абсолютное значение момента ограничено величиной M_w^0 , т. е. $|M_w| \leq M_w^0$, то допустимое абсолютное значение угла ψ можно получить из (15):

$$(16) \quad |\psi| \leq \arctg \left(\frac{R \cdot M_w^0}{g \cdot (J_w + m_w \cdot R^2 + m \cdot R^2)} \right).$$

Для получения монотонного переходного процесса следует использовать биномиальное представление характеристического полинома уравнения (14) [8], т. е. положить $\alpha = K = \gamma$. В этом случае $\tau_n = \frac{\tau_{n1}}{\gamma}$, где τ_n - желаемое время переходного процесса, γ - коэффициент сжатия времени переходного процесса, $\tau_{n1} = 4,8 \text{ с}$ - время нормированного переходного процесса при $\gamma = 1$.

Стабилизация скорости поступательного движения

Из (15) видно, что при $\psi_{ref} \neq 0$ момент $M_w \neq 0$ и платформа движется ускоренно для поддержания условия $\psi = \psi_{ref}$.

Под задачей стабилизации скорости платформы будем понимать разгон или торможение платформы до достижения значения скорости вращения колеса $\dot{\vartheta} = \dot{\vartheta}_{ref}$ и поддержание заданной скорости.

Введём линейные скорость и ускорение платформы $v = R \cdot \dot{\vartheta}$ и $\dot{v} = R \cdot \ddot{\vartheta}$, а также желаемое значение линейной скорости $v_{ref} = R \cdot \dot{\vartheta}_{ref}$. Оценим значение $\dot{v}(\psi, \dot{\psi})$ при $\psi = \psi_{ref}$ и $\dot{\psi} = 0$. Для этого умножим второе уравнение (7) на R , подставим значения $\psi = \psi_{ref}$ и $\dot{\psi} = 0$ и значение момента M_w из (15). После простых преобразований получим

$$(17) \quad \dot{v}(\psi_{ref}, 0) = g \cdot \operatorname{tg} \psi_{ref}.$$

Единственный управляющий параметр M_w уже использован для стабилизации положения маятника. Однако из (17) видно, что скоростью платформы можно управлять, изменяя значение ψ_{ref} . Отметим, что $\operatorname{sgn}(\dot{v}(\psi_{ref}, 0)) = \operatorname{sgn}(\psi_{ref})$, и введём обратную связь по скорости в виде

$$(18) \quad \psi_{ref} = \beta \cdot (v_{ref} - v).$$

В итоге из (14) и (18) получим систему уравнений

$$(19) \quad \begin{cases} \ddot{\psi} + (\alpha + K) \cdot \dot{\psi} + \alpha \cdot K \cdot \psi + \alpha \cdot \beta \cdot K \cdot v = \alpha \cdot \beta \cdot K \cdot v_{ref}; \\ \dot{v} = g \cdot \operatorname{tg}(\beta \cdot (v_{ref} - v)). \end{cases}$$

Система (19) имеет очевидную точку равновесия

$$(20) \quad v = v_{ref}, \psi = \dot{\psi} = \ddot{\psi} = 0.$$

Очевидно также, что $\operatorname{sgn}(\dot{v}) = \operatorname{sgn}(v_{ref} - v)$, т. е. скорость v стремится к точке равновесия.

Однако устойчивость всей системы (19) требует проверки.

Линеаризуем второе уравнение (19) при $\beta(v_{ref} - v) \rightarrow 0$ и приведём (19) к системе уравнений первого порядка:

$$(21) \quad \begin{cases} \dot{\psi} = \omega; \\ \dot{\omega} = -\alpha \cdot K \cdot \psi - (\alpha + K) \cdot \omega - \alpha \cdot \beta \cdot K \cdot v + \alpha \cdot \beta \cdot K \cdot v_{ref}; \\ \dot{v} = -g \cdot \beta \cdot v + g \cdot \beta \cdot v_{ref} \end{cases}$$

Матрица коэффициентов системы (21) имеет собственные значения $p_1 = g \cdot \beta$, $p_2 = -\alpha$, $p_3 = -K$, т. е. система (21) экспоненциально устойчива, и её характеристический полином имеет вид:

$$(22) \quad H(p) = p^3 + (\alpha + K + g \cdot \beta) \cdot p^2 + (\alpha \cdot K + \alpha \cdot \beta \cdot g + K \cdot \beta \cdot g) \cdot p + \alpha \cdot \beta \cdot K \cdot g.$$

Для получения монотонного переходного процесса коэффициенты полинома следует подчинить условию

$$(23) \quad \alpha = K = \beta \cdot g = \gamma.$$

В этом случае $\tau = \frac{6,3}{\gamma}$.

Моделирование

Эффективность предложенных в работе алгоритмов управления продемонстрирована путём моделирования системы в среде MATLAB/Simulink.

На рисунке 2 представлен пример управления *положением маятника*; изображены графики переходных процессов угла наклона маятника ψ и вращающего момента колеса T . В исходном состоянии маятник установлен вертикально, затем он должен занять и поддерживать положение -0.1 рад., на второй секунде даётся команда занять и удерживать положение $+0.1$ рад. и, наконец, на четвёртой секунде даётся команда занять и удерживать положение $+0$ рад.

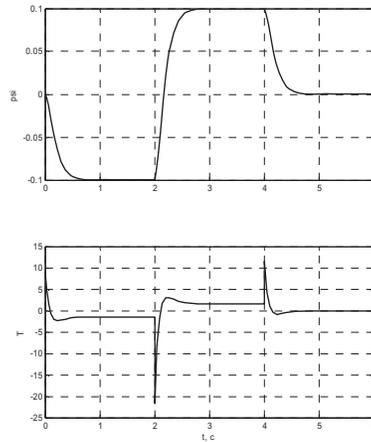


Рисунок 2 – Переходные процессы при управлении только положением маятника

На рисунке 3 представлены пример одновременного управления *положением маятника и скоростью движения платформы*; изображены графики переходных процессов угла наклона маятника ψ и угловой скорости колеса $\dot{\varphi}$. В исходном состоянии маятник расположен практически вертикально, аппарат из состояния покоя должен достичь скорости вращения колеса в 1 рад./с, сохраняя вертикальное положение маятника, на тридцатой секунде даётся команда увеличить скорость до 1.2 рад./с, на шестидесятой секунде даётся команда снизить скорость до 0.8 рад./с.

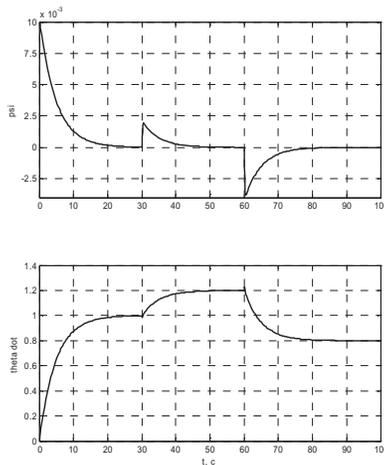


Рисунок 3 – Переходные процессы при одновременном управлении положением маятника и скоростью движения платформы

Заклучение

Предложены и обоснованы алгоритмы управления положением маятника на подвижном основании и скоростью платформы.

Приведены рекомендации по выбору параметров системы управления и даны оценки времени переходных процессов. Моделирование подтвердило эффективность предложенных алгоритмов.

Список литературы

- [1] <http://www.segway.com/>
- [2] Baloh M., Parent M.: Modeling and Model Verification of an Intelligent Self-Balancing Two-Wheeled Vehicle for an Autonomous Urban Transportation System, The Conference on Computational Intelligence, Robotics, and Autonomous Systems, 2003.
- [3] Ryo Watanabe. Motion Control of NXTway (LEGO Segway). (http://web.mac.com/ryo_watanabe/iWeb/Ryo%27s%20Holiday/NXTway-G_files/nxtway-g-1.pdf)
- [4] Robert Grepl. BALANCING WHEELED ROBOT: EFFECTIVE MODELLING, SENSORY PROCESSING AND SIMPLIFIED CONTROL. Engineering MECHANICS, Vol. 16, 2009, No. 2, p. 141–154. (http://dlib.lib.cas.cz/5310/1/16_2_141.pdf)
- [5] Adam Cardi, Matt Wagner. Control of an RC Segway. (<http://www.prism.gatech.edu/~gth856d/buffalo/paper.pdf>)
- [6] Tsai-Jiun Ren, Tien-Chi Chen, Chun-Jung Chen. Motion control for a two-wheeled vehicle using a self-tuning PID controller. Control Engineering Practice 16 (2008) 365–375. (<http://www.nrec.ir/files/articles/Fuzzy%20PID%20Control%20with%20Accelerated%20Reasoning%20for.pdf>)
- [7] А.М. Формальский. О стабилизации перевёрнутого маятника с неподвижной или подвижной точкой подвеса. ДАН, 2006, том 406, № 2, с. 175-179.
- [8] Зубов В.И. Устойчивость движения. (Методы Ляпунова и их применение). М.: Высшая школа, 1973.
- [9] Мирошник И.В. Теория автоматического управления. Линейные системы. СПб.: Питер, 2005.

СИСТЕМА УПРАВЛЕНИЯ УСТАНОВКОЙ МАГНЕТРОННОГО НАПЫЛЕНИЯ

И.С. Бабер, А.Н. Гинзбург, Н.И. Карпов, С.П. Марьев, О.Г. Митрофанов

Институт прикладной физики РАН
603950, Нижний Новгород, ул. Ульянова, 46, Россия
om@appl.sci-nnov.ru
+7(831)-436-81-44

Ключевые слова: промышленная автоматизация, программирование, вакуумные установки, магнетронное напыление

Abstract

The automated control system for the magnetron sputtering is presented.

Введение

Для нанесения защитно-декоративных покрытий на различные изделия уже не один год в нашей стране и за рубежом используются вакуумные технологии. Эти технологии отличаются высокой эффективностью, гибкостью и экологичностью, однако требуют при этом применения сложного оборудования и детального, по большей мере, автоматического выполнения технологических процессов. Эти обстоятельства предопределяют оснащение вакуумных установок современными системами управления. Настоящий материал посвящен описанию автоматизированных систем управления вакуумными установками, производимых научно-производственной компанией «Элан-практик» (г. Дзержинск) - ведущей отечественной компанией, занимающейся разработкой, производством вакуумных установок и технологий к ним. Одно из наиболее серийных и универсальных изделий компании «Элан-практик» - вакуумные установки УНИП-900. Данные установки имеют объем рабочей камеры 900 куб.дм., в камере может быть установлено до 6 магнетронов [1].

Технологические возможности УНИП-900:

- 1) Твердые декоративные покрытия MeN, MeC, MeCN (Me=Ti, Zr, Cr, Nb) с нанокompозитным антикоррозионным подслоем на изделия массового применения (замковая фурнитура, сантехника, оправы для очков, корпуса часов и др.) без гальванического подслоя.
- 2) Сверхтвердые нанокompозитные покрытия 2D-типа для защиты от износа и коррозии взамен твердого гальванического хрома: TiN/CrN, TiN/NbN, TiN/SiN, TiN/CN, TiN/AlN, CrN/SiN, CrN/NbN, на золотниковые пары, автомобильные детали, запорную арматуру для нефте-газопроводов, эндопротезы и др.[2].

Какие основные требования, помимо функциональности, предъявляются к системам автоматизации таких установок?

Во-первых, это высокая надежность. Установки рассчитаны на круглосуточную эксплуатацию, а рабочая камера может вмещать до нескольких сотен изделий, так что даже единичный сбой в системе управления может обойтись очень дорого. Поэтому необходимо применение надежного промышленного электронного оборудования и программного обеспечения, имеющего возможность восстановления после сбоев и позволяющего продолжить выполнение технологического процесса.

Во-вторых, это простота использования. Для применения на промышленном предприятии важно, чтобы установку мог эксплуатировать персонал, не имеющий специального образования в области вакуумных технологий. Система должна иметь высокую степень автоматизации техпроцесса, самодиагностику, дружелюбный человеко-машинный интерфейс,

включающий систему предупреждений и подсказок оператору, должна позволять быстро выявлять неисправности и исключать влияние ошибок персонала на качество продукции [3].

1 Краткое описание установки и системы управления

Оборудование установки может быть условно разделено на вакуумное, технологическое и вспомогательное.

Вакуумная подсистема предназначена для проведения операции откачки рабочей камеры. Состоит из двух ступеней форвакуумного насоса, байпасного и форвакуумного клапанов, диффузионного насоса, вакуумного затвора. Все насосы управляются с помощью электромагнитных пускателей, клапаны и затвор имеют электропневматическое управление, их положение контролируется с помощью электромагнитных датчиков. Вода для охлаждения диффузионного насоса контролируется с помощью датчика протока. Имеются аварийные блокировки по току и температуре. Чтобы контролировать давления в вакуумных магистралях, имеются датчики, позволяющие проводить измерения вакуума до 0.001 Па.

Технологическая подсистема предназначена для проведения различных технологических операций. Состоит из магнетронов и заслонок, блоков питания, клапанов подачи газа, регуляторов расхода газа, привода вращения изделий, датчиков давления.

Вспомогательное оборудование для контроля давления и протока воды, давления воздуха для питания пневмооборудования.

Архитектура системы управления достаточно традиционна, система управления построена по двухуровневой схеме [4]. Нижний уровень построен на базе промышленного контроллера, он обеспечивает управление всеми устройствами вакуумной и технологической частей установки, управление блоками питания магнетронов, позволяет автономно проводить технологический процесс после загрузки его параметров с верхнего уровня. Верхний уровень построен на базе промышленного панельного компьютера, обеспечивает визуализацию работы оборудования, связь с нижним уровнем, конфигурирование, протоколирование работы установки, индикацию аварий, диспетчерское управление. Связь между уровнями осуществляется по интерфейсу RS485.



Рисунок 1 – Шкаф системы управления вакуумной установки УНИП-900

2 Оборудование автоматизированной системы управления

На нижнем уровне системы применено оборудование стандарта MicroPC фирмы Fastwel. В составе контроллера используются плата CPU188, две платы 16-канального аналогового ввода AI16, две платы 8-канального аналогового вывода AO16V8, плата 32-канального цифрового ввода с гальванической изоляцией DI32, плата 32-канального цифрового вывода с гальванической изоляцией DO32. В некоторых установках нижний уровень построен на основе промышленного оборудования фирмы Advantech серии ADAM-5000. В качестве рабочей станции оператора используется промышленный панельный компьютер PPC-103T производства фирмы Advantech. Все электронное оборудование автоматизированной системы управления размещено в одном шкафу, внешний вид которого приведен на рисунке 1.

3 Программное обеспечение системы управления

Программное обеспечение разрабатывалось таким образом, чтобы во-первых, быть единым для всей гаммы вакуумных установок, а во-вторых, иметь возможность работать на нескольких аппаратных платформах (это актуально для нижнего уровня установки). Чтобы иметь возможность выполнения первого условия, в составе системы есть конфигуратор (рисунок 2), с помощью которого настраиваются параметры, индивидуальные для данного типа установки и проводимых на ней технологических процессов, а именно:

- Название установки.
- Количество, тип, интерфейс, параметры источников питания.
- Количество магнетронов, расположение в рабочей камере и наличие на них заслонок.
- Параметры, калибровки регуляторов расхода газа.
- Параметры ПИД-регуляторов.
- Настройка таблицы технологического процесса.
- Настройка таблиц газоподачи.
- Настройка формул газоподачи.
- Настройка паролей на изменение параметров технологического процесса и управления технологическим процессом.
- Сохранение параметров в файле настроек, который читается при запуске диспетчерской программы управления вакуумной установкой.

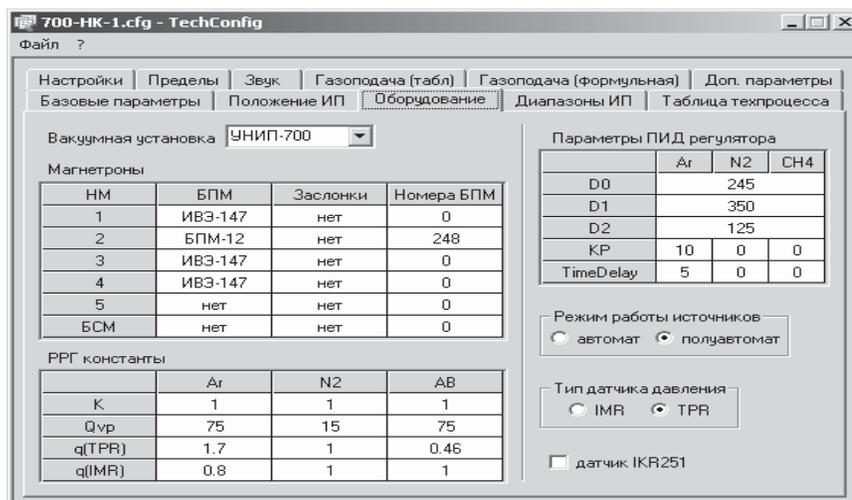


Рисунок 2 – Панель конфигуратора

Диспетчерская программа управления вакуумной установкой реализует следующие функции:

- Чтение параметров из файла конфигурации, если в рабочей директории программы имеются несколько файлов конфигурации, предлагается выбрать из списка.
- Визуальное представление работы вакуумного и технологического оборудования на соответствующих панелях, вывод значений основных рабочих параметров – давлений, токов и напряжений, расходов. Отображение текущего состояния и аварийных ситуаций (рисунок 3 и рисунок 4).
- Настройка порогов давлений для переключения насосов, клапанов и высоковакуумного затвора установки.
- Настройка параметров тестирования натекания и прокачки.
- Настройка рабочих параметров технологического процесса.
- Поддержка связи с контроллером установки, загрузка техпроцесса, параметров установки в контроллер, получение от контроллера текущих состояний, параметров работы оборудования, аварий.
- Выдача команд на запуск откачки, запуск техпроцесса.
- Выдача команд на остановку, завершение техпроцесса.
- Выдача команд на завершение работы установки.
- Проведение измерений параметров установки – натекания и прокачки.
- Ведение многостраничного журнала техпроцесса – страница таблицы проведенных операций с усредненными и заданными значениями параметров, страница сообщений, страница измеренных значений во времени, страница настроек табличной и формульной газоподачи, страница усредненных значений всех параметров установки во времени с 30-секундным разрешением. Журнал создается в начале техпроцесса и заканчивает формироваться в конце техпроцесса.



Рисунок 3 – Визуальное представление работы вакуумного оборудования

Программное обеспечение контроллера нижнего уровня обеспечивает:

- автоматическое проведение операций по подготовке, откачке рабочей камеры, в полуавтоматическом режиме - загрузка и выгрузка изделий.
- автоматическое проведение операций технологического процесса.
- отслеживание состояния оборудования, формирование и обработку сигналов аварии.
- управление источниками питания магнетронов, выдачу ими заданных токов и напряжений. Обрабатывает сигналы аварийных состояний источников питания магнетронов, контролирует выходные напряжения и токи. В случае отклонения тока или напряжения источника на 20% от заданных значений, формируется соответствующее аварийное сообщение, производится его обработка. Для некоторых технологических операций контролируется доза – количество отданной магнетрону энергии.



Рисунок 4 – Визуальное представление работы технологического оборудования

- автоматическое управление вакуумным и технологическим оборудованием установки.
- прием команд и двусторонний обмен информацией с рабочей станцией оператора, передачу текущих параметров состояния оборудования и хода технологического процесса.

Заключение

Преимущества оснащения вакуумных установок автоматизированными системами управления очевидны:

- Простота ввода установки в эксплуатацию. Система автоматизации позволяет облегчить процесс ввода в эксплуатацию, за счет встроенных средств диагностики имеется возможность определить ошибки монтажа и оценить качество вакуумной подсистемы.
- Простота эксплуатации и обслуживания. Простота обучения обслуживающего персонала и технологов. Установка имеет высокую степень автоматизации технологических процессов. Обслуживающий персонал и технологи проходят короткий базовый курс обучения. Возможность анализа аварийных ситуаций, усредненных параметров работы оборудования.

ния, прохождения техпроцессов позволяет отслеживать состояние оборудования, своевременно производить профилактику, замену и ремонт.

- Высокая повторяемость заданных качеств изделий. Система автоматизации позволяет минимизировать влияние человеческого фактора на качество изделий, есть возможность отслеживать такие факторы, как загрязнение камеры, отклонение параметров работы магнетронов, количественный состав газовых смесей.

Установки УНИП-900 работают на следующих предприятиях:

- "Ковровский Электромеханический Завод", производство деталей замков.
- Угличский часовой завод, производство часовых корпусов.
- Завод «LIW LEWANT», Польша, г.Вроцлав, производство оправ для очков.
- Предприятие «Зенит», г.Дмитровград

Список литературы

- [1] Ю.В.Агабеков, А.М.Сутырин «Несбалансированные магнетронные распылительные системы с усиленной ионизацией плазмы». Из кн. Труды постоянно действующего научно-технического семинара. "Электровacuумная техника и технология" (за 1997/98 гг.), Москва, 1999 г., сс. 102-108.
- [2] Ю.В. Агабеков «Защитно-декоративные керамические покрытия на изделиях массового потребления – вакуумные технологии и оборудование с использованием несбалансированных магнетронов». Из кн. Труды постоянно действующего научно-технического семинара. "Электровacuумная техника и технология" (за 1999-2002 гг.), Москва, 2003, сс. 24-31.
- [3] И.С. Бабер, В.Г. Безуглов, И.В. Галахов, А.Н. Гинзбург, Н.И. Карпов, М.Ю. Кирдяшкин, С.П. Марьев, О.Г. Митрофанов, В.А. Осин «Автоматизированная система управления емкостными накопителями энергии установки «Луч». Труды IX Международной конференции «Проблемы управления и моделирования в сложных системах», Самара, Россия, 22-28 июня 2007 г., стр. 273-278.
- [4] И.С., Бабер, А.Н. Гинзбург, Н.И.Карпов, С.П. Марьев, О.Г. Митрофанов "Автоматизированная система управления вакуумными установками ИМПЛАЗ-1000", Труды XI Международной конференции «Проблемы управления и моделирования в сложных системах», Самара, Россия, 22-24 июня 2009г., с. 371-378, Издательство Самарский научный центр РАН, ISBN 978-5-93424-423-2.

АЛГОРИТМЫ СЕЛЕКЦИИ СПЕКТРОГРАММ В СИСТЕМАХ АВТОМАТИЗАЦИИ АТМОСФЕРНЫХ МИКРОВОЛНОВЫХ НАБЛЮДЕНИЙ

А.Н. Гинзбург¹, В.М. Дёмкин²

¹Институт прикладной физики РАН
603005, Нижний Новгород, ул. Ульянова, 46, Россия
aginzburg@bk.ru
тел. +7 (831) 436-80-67

²Нижегородский филиал Государственного университета – Высшая школа экономики
603155, Нижний Новгород, ул. Б.Печерская, 25/12, Россия
valerdem@mail.ru
тел: +7 (831) 416-95-49

***Ключевые слова:** платформа Windows, язык программирования C++, графический интерфейс пользователя, приложение, класс, библиотека Qt (Trolltech, Nokia Qt), виджеты, потоки, система автоматизации, алгоритмы селекции спектрограмм, парная корреляция*

Abstract

This article presents some results of single-threaded applications use in the systems of automation of atmospheric investigations. Software development is oriented to the use Qt library from Trolltech and C++ for Windows platform. Algorithms of selection of spectrograms and project decisions of development of the systems of automation of atmospheric observations by ground-based microwave sounding are also presented in the article.

Введение

Одной из характерных особенностей процесса разработки программного обеспечения современных систем автоматизации микроволновых спектральных атмосферных исследований, проводимых с помощью созданных в ИПФ РАН уникальных радиометрических комплексов, является стремление к созданию этих систем в виде однопоточных или многопоточных Qt-приложений для платформы Windows [1-2].

Каждый из таких радиометрических комплексов “по-своему” ориентирован на современную вычислительную технику и универсальное или специализированное интерфейсное оборудование. Каждый из них так же “по-своему” ориентирован на проведение многочасовых спектральных измерений методом наземного микроволнового зондирования атмосферы: для одних комплексов это наблюдения стратосферного озона [3-5], для других – зондирование термической структуры стратосферы [6-7].

Функциональной основой для создания высокопроизводительных Windows-приложений является широко известная библиотека Qt (Trolltech – с 1996г., Nokia Qt – с 2008г.) [8].

С помощью Qt, используя один и тот же код на C++, можно создавать как однопоточные, так и многопоточные приложения для платформ Windows (от 3.11 до 7), Mac OS X, а также Linux, Solaris, IRIX и многих других вариантов UNIX.

Однопоточные Windows-приложения просты, поэтому они до сих пор востребованы во многих системах автоматизации, разрабатываемых в ИПФ РАН. Поскольку процесс вторичной обработки данных не всегда удается выполнить в рамках главного потока вместе с процессами сбора и первичной обработки данных, то разработчики зачастую вынуждены проектировать два независимых однопоточных приложения, которые могут работать как автономно, так и вместе: одно приложение по завершении очередного цикла сбора данных записывает файлы первичных данных в архив, а другое – извлекает их оттуда для последующей обработки.

В частности, именно такой особенностью обладают современные мобильные озонметры, где сбор данных, первичная обработка данных и управление режимами наблюдений осуществляются под управлением независимого однопоточного Windows-приложения [3-5]. Для большинства дистанционных методов микроволнового зондирования верхней атмосферы в экспериментах, связанных с наблюдением резонансных радиолиний молекул примесных газов, характерное время накопления сигнала радиометра составляет единицы минут, что позволяет в реальном масштабе времени осуществлять сбор и первичную обработку данных, а также поддержку архива файлов первичных данных. Вторичная обработка, как правило, требует тщательного анализа данных, поэтому в экспедиционных условиях проведения наблюдений она редко используется в фоновом режиме непосредственно во время измерений. Заметим, что в стационарных условиях вторичную обработку архивов файлов первичных данных правильной будет называть “первичной”, так как каждый этап такой обработки связан со “своей” задачей.

Для иллюстрации механизма взаимодействия между двумя независимыми приложениями в экспериментах, связанных с наблюдением резонансных радиолиний молекул примесных газов, приведем скриншот рабочего стола для Windows-приложения, осуществляющего обработку архива файлов первичных данных (рис. 1). Как видно из рисунка, главный виджет приложения всякий раз перекрывается одним из виджетов вызванного процесса. В свою очередь, детализация какого-либо процесса может привести к появлению на рабочем столе новых виджетов, которые здесь не показаны. Как правило, при обработке архива файлов первичных данных на рабочем столе отображаются только три основных виджета: главный виджет приложения “Первичная обработка данных микроволновых наблюдений атмосферного озона”, виджет “Спектральная линия” и виджет “Восстановление высотного распределения озона”. Для анализа поведения опорного канала спектроанализатора за весь период наблюдений, связанный с данным набором файлов первичных данных, на рабочий стол можно выложить виджет “Опорный канал”. По завершении процедуры восстановления высотного распределения озона на рабочем столе может быть отображен также виджет “Вертикальный профиль озона”.

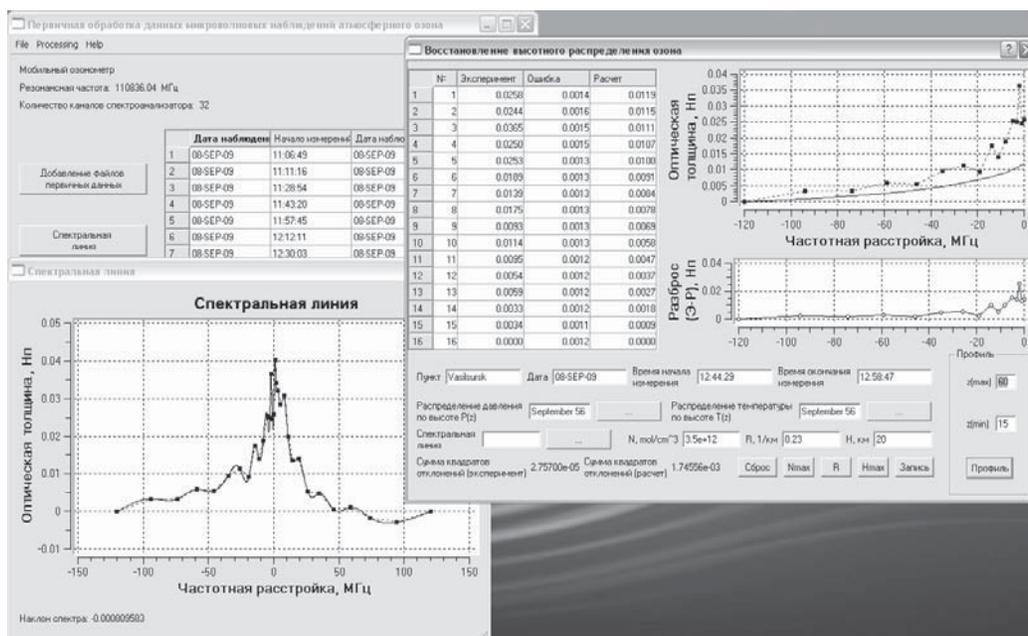


Рисунок 1 – Скриншот рабочего стола однопоточного Windows-приложения во время сеанса обработки архива файлов первичных данных

Отметим, что обработка архива файлов первичных данных к настоящему времени пока производится только в режиме “файл за файлом” (или “спектрограмма за спектрограммой”). Если спектрограмма по какой-либо причине отбраковывается, то такой файл обработке уже не подлежит. Очевидно, что в случае перехода от ручного режима обработки к автоматическому процедура визуального контроля “качества” спектрограмм должна быть заменена процедурой их селекции, поначалу ручной, а затем и автоматической. Очевидно также, что без эвристик в алгоритмах селекции никак не обойтись.

1 Алгоритмы селекции спектрограмм – первый шаг

Перед тем как перейти к обсуждению алгоритмов селекции, отметим здесь характерные особенности спектрограмм на основе их визуального представления, указав на “хорошую” спектрограмму (рис. 2), “не очень хорошую” (рис. 3) и “плохую” (рис. 4). Уровни градации “качества” спектрограмм весьма условны: здесь многое зависит как от условий проведения наблюдений, так и от “точки зрения” самого пользователя.

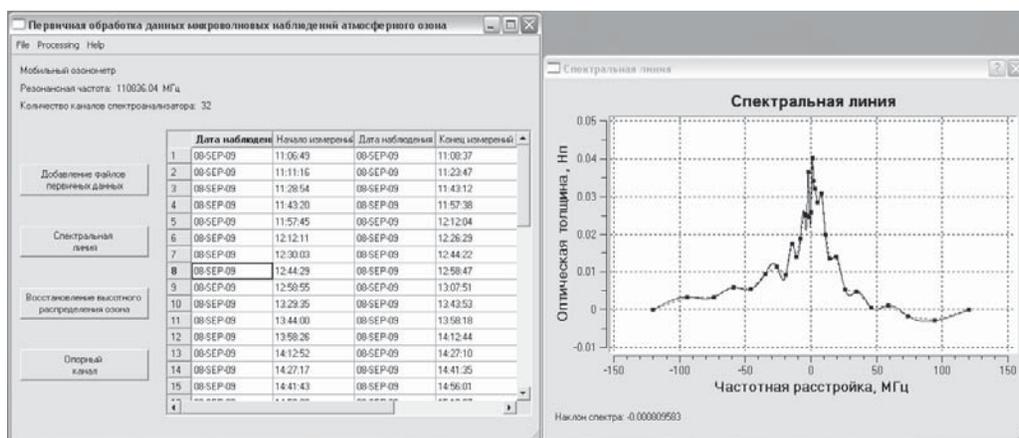


Рисунок 2 – Скриншот рабочего стола однопоточного Windows-приложения во время сеанса визуализации архива файлов первичных данных (“хорошая” спектрограмма)

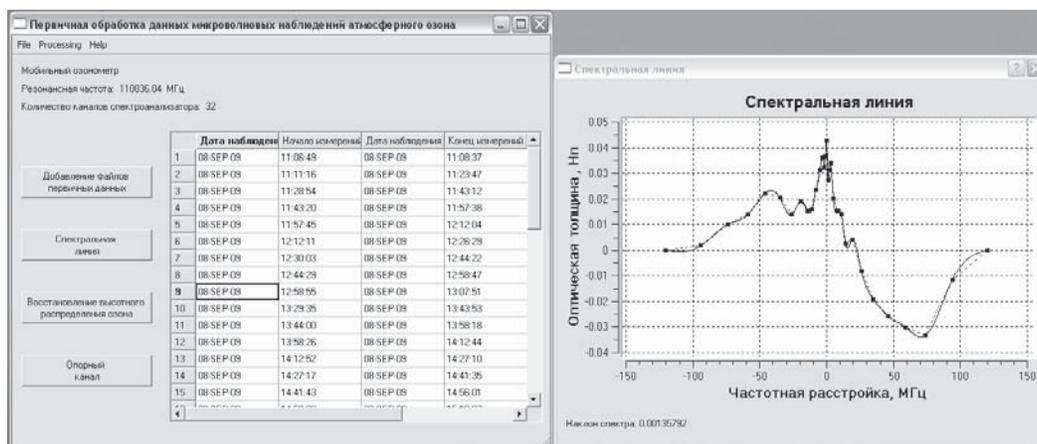


Рисунок 3 – Скриншот рабочего стола однопоточного Windows-приложения во время сеанса визуализации архива файлов первичных данных (“не очень хорошая” спектрограмма)

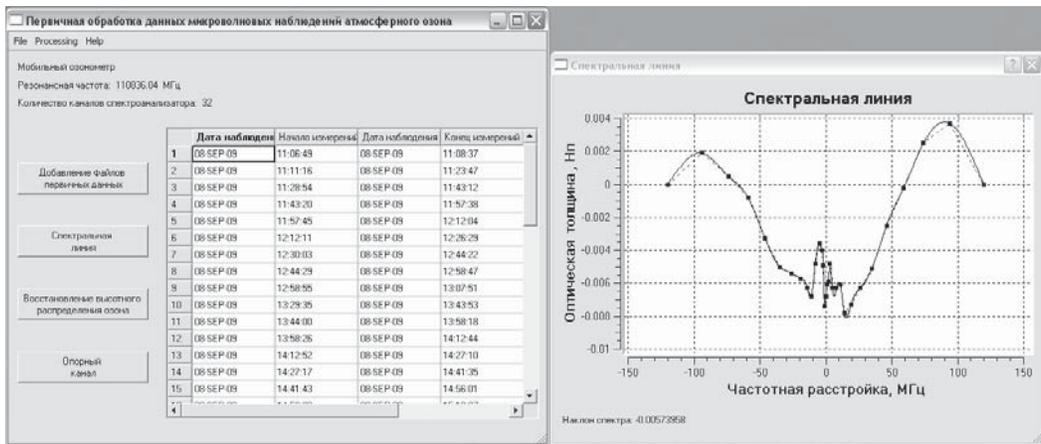


Рисунок 4 – Скриншот рабочего стола однопоточного Windows-приложения во время сеанса визуализации архива файлов первичных данных (“плохая” спектрограмма)

Для реализации ручного режима селекции спектрограмм можно воспользоваться поддержкой механизма обработки правого клика мыши для таблицы QTable, записи которой соответствуют файлам первичных данных, хранящихся в контейнере после выполнения процедуры “Добавление файлов первичных данных”. В настоящее время контейнер предназначен для хранения 200 файлов, что позволяет обеспечить обработку данных, зарегистрированных в течение суточных наблюдений.

Алгоритм ручного режима селекции спектрограмм:

- 1) переключение режима селекции на “ручной”;
- 2) активация элемента таблицы с помощью левого клика мыши;
- 3) визуальный контроль “качества” спектрограммы;
- 4) селекция “не очень хороших” и “плохих” спектрограмм с помощью правого клика мыши с изменением цвета фона строки таблицы.

Дальнейшие действия пользователя могут быть связаны с удалением “помеченных” файлов из контейнера благодаря выполненной селекции спектрограмм.

2 Алгоритмы селекции спектрограмм – второй шаг

Для реализации автоматизированного режима селекции спектрограмм можно воспользоваться процедурой расчета коэффициентов парной корреляции (по Пирсону) для заданной базовой спектрограммы из категории “хороших” по отношению ко всем остальным и последующего удаления “помеченных” файлов из контейнера. Эвристика, основанная на сопоставлении коэффициентов парной корреляции, в последующем позволит задать порог и для автоматической селекции всех “не очень хороших” и “плохих” спектрограмм.

Эвристический алгоритм автоматизированного режима селекции спектрограмм:

- 1) переключение режима селекции на “автоматизированный”;
- 2) активация элемента таблицы с помощью правого клика мыши (указание базовой спектрограммы);
- 3) селекция “не очень хороших” и “плохих” спектрограмм на основе расчета коэффициентов парной корреляции с изменением цвета фона строки таблицы;
- 4) визуальный контроль “качества” спектрограмм по результатам их селекции.

Дальнейшие действия пользователя могут быть связаны либо с удалением “помеченных” файлов из контейнера благодаря выполненной селекции спектрограмм, либо с повтором процедуры селекции для новых условий.

Эвристический алгоритм автоматического режима селекции спектрограмм:

- 1) переключение режима селекции на “автоматический”;
- 2) задание порогового значения коэффициента парной корреляции;
- 3) селекция “не очень хороших” и “плохих” спектрограмм на основе расчета коэффициентов парной корреляции с удалением “помеченных” файлов из контейнера.

3 Апробация алгоритмов селекции спектрограмм

В качестве инструментального средства для апробации алгоритмов селекции на основе эвристики, основанной на сопоставлении коэффициентов парной корреляции, была выбрана среда SPSS (в современной редакции PASW) [9].

Для иллюстрации механизма, положенного в основу будущих алгоритмов селекции, приведем скриншот окна SPSS Viewer версии 13.0 (рис. 5). Из рисунка видно, что коэффициенты парной корреляции рассчитывались для выборки из 10 спектрограмм. Сами данные (называемые наблюдениями в терминологии SPSS) для переменных анализа должны располагаться во вкладке Data View окна SPSS Data Editor, а описания этих переменных – во вкладке Variable View. Видно также, что для каждой переменной анализа, связанной со своей спектрограммой, представлены 16 наблюдений (отсчеты для левого крыла спектральной линии).

		12:30:03-12:44:22	12:44:29-12:58:47	12:58:55-13:07:51	13:29:35-13:43:53	13:44:00-13:58:18	13:58:26-14:12:44	14:12:52-14:27:10	14:27:17-14:41:35	14:41:43-14:56:01	14:56:08-15:10:27
12:30:03-12:44:22	Pearson Correlation	1	.831**	.811**	.760**	.759**	.830**	.677**	.715**	.849**	.839**
	Sig. (2-tailed)		.000	.000	.001	.001	.000	.004	.002	.000	.000
	N	16	16	16	16	16	16	16	16	16	16
12:44:29-12:58:47	Pearson Correlation	.831**	1	.848**	.859**	.807**	.896**	.843**	.732**	.925**	.810**
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.001	.000	.000
	N	16	16	16	16	16	16	16	16	16	16
12:58:55-13:07:51	Pearson Correlation	.811**	.848**	1	.787**	.822**	.822**	.765**	.832**	.809**	.684**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.001	.000	.000	.003
	N	16	16	16	16	16	16	16	16	16	16
13:29:35-13:43:53	Pearson Correlation	.760**	.859**	.787**	1	.771**	.830**	.870**	.775**	.826**	.669**
	Sig. (2-tailed)	.001	.000	.000		.000	.000	.005	.000	.000	.003
	N	16	16	16	16	16	16	16	16	16	16
13:44:00-13:58:18	Pearson Correlation	.759**	.807**	.822**	.771**	1	.822**	.804**	.691**	.747**	.780**
	Sig. (2-tailed)	.001	.000	.000	.000		.000	.000	.004	.001	.000
	N	16	16	16	16	16	16	16	16	16	16
13:58:26-14:12:44	Pearson Correlation	.830**	.896**	.822**	.830**	.792**	1	.685**	.792**	.927**	.781**
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.003	.000	.000	.000
	N	16	16	16	16	16	16	16	16	16	16
14:12:52-14:27:10	Pearson Correlation	.677**	.843**	.765**	.870**	.804**	.685**	1	.647**	.690**	.668**
	Sig. (2-tailed)	.004	.000	.001	.005	.000	.003		.007	.003	.005
	N	16	16	16	16	16	16	16	16	16	16
14:27:17-14:41:35	Pearson Correlation	.715**	.732**	.832**	.775**	.681**	.792**	.647**	1	.815**	.719**
	Sig. (2-tailed)	.002	.001	.000	.000	.004	.000	.007		.000	.002
	N	16	16	16	16	16	16	16	16	16	16
14:41:43-14:56:01	Pearson Correlation	.849**	.925**	.809**	.838**	.747**	.837**	.690**	.815**	1	.809**
	Sig. (2-tailed)	.000	.000	.000	.000	.001	.000	.002	.000		.000
	N	16	16	16	16	16	16	16	16	16	16
14:56:08-15:10:27	Pearson Correlation	.839**	.810**	.894**	.809**	.780**	.701**	.868**	.719**	.809**	1
	Sig. (2-tailed)	.000	.000	.003	.003	.000	.000	.005	.002	.000	
	N	16	16	16	16	16	16	16	16	16	16

Рисунок 5 – Скриншот окна SPSS Viewer

Анализ отчета позволил сформулировать эвристику для заданной выборки и с помощью визуального контроля “качества” убедиться в применимости предлагаемого подхода к будущей реализации автоматизированных и автоматических алгоритмов селекции спектрограмм. Например, для 2-ой спектрограммы (12:44:29–12:58:47) нижний порог коэффициента парной корреляции составляет 0.732 (максимальное значение в выборке), а верхний порог – 0.925 (при максимальном значении в выборке 0.937). По критерию “нижний порог” эта спектрограмма может получить статус базовой из категории “хороших”, а по критерию “верхний порог” 4-ая (13:29:35–13:43:53), 6-ая (13:58:26–14:12:44) и 9-ая (14:41:43–14:56:01) спектрограммы могут получить статус “хороших”, что и подтверждается при визуальном контроле их “качества”.

Основанием для такого утверждения является превышение верхнего порога коэффициента парной корреляции базовой спектрограммы: для 4-ой спектрограммы он составляет 0.936, а для 6-ой и 9-ой – 0.937. Оставшиеся 6 спектрограмм получают статус “не очень хороших”, так как их верхний порог коэффициента парной корреляции существенно ниже базового: для 1-ой (12:30:03–12:44:22) он составляет 0.849, для 3-ей (12:58:55–13:07:51) – 0.848, для 5-ой (13:44:00–13:58:18) и 8-ой (14:27:17–14:41:35) – 0.832, а для 7-ой (14:12:52–14:27:10) – 0.843.

Заключение

Сформулируем основные результаты работы:

- 1) Выделены особенности спектрограмм как объектов алгоритмов селекции.
- 2) Представлены проектные решения по разработке алгоритмов селекции спектрограмм (ручной, автоматизированный и автоматический режимы).
- 3) Даны рекомендации по внедрению алгоритмов селекции спектрограмм в однопоточные Windows-приложения разрабатываемых систем автоматизации микроволновых атмосферных исследований.

Благодарности

Работа выполнена при поддержке грантов РФФИ (2009-11гг., грант 09-05-97014-р_поволжье_a; 2009-10гг., грант 09-05-97015-р_поволжье_a).

Список литературы

- [1] Гинзбург А.Н., Дёмкин В.М. Использование многопоточных приложений в системах автоматизации физического эксперимента // Проблемы управления и моделирования в сложных системах: Труды X Международной конференции (Самара, 23-25 июня 2008г.). – Самара: Самарский научный центр РАН, 2008, с. 518–521.
- [2] Гинзбург А.Н., Дёмкин В.М. Принципы проектирования многопоточных Qt-приложений в системах автоматизации атмосферных исследований // Проблемы управления и моделирования в сложных системах: Труды XI Международной конференции (Самара, 22-24 июня 2009г.). – Самара: Самарский научный центр РАН, 2009, с. 652–657.
- [3] Куликов Ю.Ю., Красильников А.А., Дёмкин В.М., Рыскин В.Г. Вариации концентрации мезосферного озона во время полного солнечного затмения 29 марта 2006 года по данным микроволновой радиометрии // Изв. РАН. Физика атмосферы и океана, 2008, т.44, №4, с.522–526.
- [4] Yu.Yu.Kulikov, A.A.Krasilnikov, V.M.Demkin Fast variations of thermal emission of middle atmosphere in a line of ozone at frequency 110.8 GHz on plateau Shatshatmas-Kislovodsk // Abstr. of the 31st Annual Apatity seminar “Physics of Auroral Phenomena”. – Apatity, 2008, p.72–73.
- [5] Yu.Yu.Kulikov, V.L. Frolov, V.M.Demkin, G.P. Komrakov, A.A.Krasilnikov, V.G. Ryskin Influence of HF powerful waves on the ozone content in the Earth’s atmosphere // Abstr. of the 33rd Annual Apatity seminar “Physics of Auroral Phenomena”. – Apatity, 2010, p.76.
- [6] Федосеев Л.И., Швецов А.А., Шкаев А.П., Дёмкин В.М., Караштин Д.А., Кукин Л.М., Божков В.Г., Геннеберг В.А., Петров И.В., Щитов А.М. Радиометры миллиметрового диапазона длин волн с твердотельными модуляторами-калибраторами // Материалы 18-й Международной Крымской конф. “СВЧ-техника и телекоммуникационные технологии” (КрыМиКо’ 2008). Тез. докл. – Севастополь, 2008, с.878–879.
- [7] Швецов А.А., Дёмкин В.М., Караштин Д.А., Скалыга Н.К., Федосеев Л.И. Микроволновый спектро-радиометрический комплекс для дистанционного зондирования термической структуры стратосферы // Известия ВУЗов. Радиофизика, 2009, т.52, №8, с.671–677.
- [8] <http://www.qt.nokia.com>
- [9] <http://www.spss.ru>

СИСТЕМА ИЗМЕРЕНИЙ КООРДИНАТ СМЕЩЕНИЙ ПРИВОДНОГО ВАЛА ТОПЛИВНОГО НАСОСА СИЛОВОЙ УСТАНОВКИ

В.А. Белослудцев, Ю.Н. Секисов

Институт проблем управления сложными системами РАН

443020, Самара, ул. Садовая, 61, Россия

bva_69@nm.ru

тел: +7 (846) 333-26-77, факс: +7 (846) 333-27-70

***Ключевые слова:** силовая установка, датчик частоты вращения, вихретоковый датчик, осевое смещение, радиальный зазор*

Abstract

The original approach for obtaining supplementary information by the means of sensors with advanced functional capabilities is offered. The example of approach realization by the use of single - coil eddy - current sensors is given. The measuring system description as well as an error's estimation are given too.

Введение

Важнейшим компонентом современных силовых установок являются средства противоаварийной защиты. Одним из этапов совершенствования средств противоаварийной защиты является необходимость увеличения общего объема полезной информации о рабочих процессах, протекающих в силовых установках, и параметрах, обеспечивающих безопасный режим работы. Требуемый объем информации может быть получен за счет установки на объект соответствующего количества датчиков.

Применительно к уже находящимся в эксплуатации силовым установкам увеличение объема информации связано с установкой дополнительных датчиков. В большинстве случаев, когда установка дополнительных датчиков требует нарушения целостности элементов конструкции, например, сверления дополнительных установочных отверстий, увеличение объема информации за счет установки дополнительных датчиков практически невыполнимо, так как, прежде всего, ухудшаются прочностные характеристики этого элемента конструкции силовой установки.

Предлагается оригинальный подход получения дополнительной информации без «вмешательства» в конструкцию силовой установки. Он заключается в замене штатных датчиков, установленных на силовой установке, датчиками оригинальной конструкции, обладающих повышенными информационными возможностями, что позволяет сохранить каналы прежней информации и получить каналы дополнительной информации.

Рассматривается решение конкретной технической задачи, связанной с износом подшипника в турбинном приводе топливного насоса силовой установки. Дается описание системы измерения и ее технические характеристики.

Пример реализации предлагаемого подхода

Физическим параметром, характеризующим износ подшипника в турбинном приводе топливного насоса, является осевое перемещение вала турбины. На реальном агрегате – насосе с турбинным приводом установлены два датчика частоты вращения (ДЧВ), которые взаимодействуют с измерительным диском, расположенным на вале турбины. На диске имеется выступы

в радиальном направлении в форме «зуба» с прямоугольным сечением. Датчики установлены на корпусе агрегата над диском в радиальном направлении и сдвинуты по направлению вращения на некоторый угол. Каждый датчик выдает независимую информацию о частоте вращения турбины по числу импульсов, возникающих при прохождении зуба в зоне чувствительного элемента ДЧВ.

В соответствии с предлагаемым подходом штатные датчики частоты вращения заменяются на оригинальные одновитковые вихретоковые датчики (ОВТД) [1], габаритные и установочные размеры которых полностью соответствуют штатным ДЧВ. Функциональные возможности предложенных датчиков, оригинальная структура средств измерения и используемые методы преобразования информации позволяют сформировать два независимых канала измерения частоты вращения турбины, а также дополнительный информационный канал измерения осевого смещения вала турбины. В общем случае информационный сигнал с ОВТД связан не только с оборотами и осевым смещением измерительного диска, но и зависит от радиального зазора между чувствительным элементом ОВТД и торцом зуба диска. Для снижения погрешности от изменений зазора на работающем агрегате алгоритм вычисления осевого смещения вала определяет и учитывает величину радиального зазора.

Описание системы измерения

Разработан и создан стендовый вариант системы измерения оборотов и осевых смещений вала турбины агрегата. Структурная схема системы содержит два оригинальных одновитковых вихретоковых датчиков, два измерительных преобразователя, модуль АЦП E14-440 фирмы L-CARD и вычислитель, в качестве которого в стендовом варианте системы используется персональный компьютер (Рисунок 1).

Измерительный преобразователь содержит измерительную цепь в виде моста Блумлейна и средства нормализации сигнала в напряжение постоянного тока с уровнем 0...5В. Кроме того, измерительный преобразователь формирует аналоговый сигнал «п» для системы управления силовой установкой, заменяющий сигнал ДЧВ. Особенностью мостовой схемы является отсутствие компенсационного датчика, роль которого выполняет имитатор, представляющий собой катушку индуктивности, расположенную на электронной плате измерительного преобразователя.

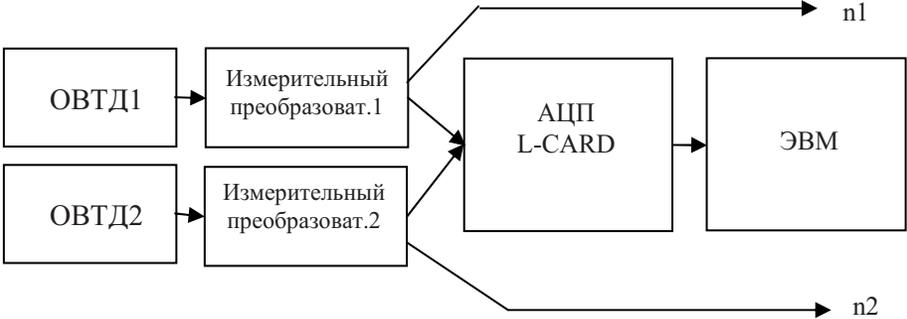


Рисунок 1 – Структурная схема системы

Конструкция ОВТД1 и ОВТД2 идентична и характерна тем, что чувствительный элемент в торцевой части датчика смещен относительно центра на половину интервала предполагаемого перемещения вала турбины. В установочные отверстия агрегата датчики устанавливаются так, чтобы их чувствительные элементы по отношению к плоскости вращения измерительного диска были смещены в разные стороны. В результате перемещение «зуба» измерительного диска в осевом направлении будет увеличивать перекрытие чувствительного элемента первого датчика и уменьшать перекрытие чувствительного элемента второго датчика. Таким образом, при вращении диска измерительный преобразователь будет вырабатывать импульсный сигнал в момент прохождения «зуба» под чувствительным элементом, а амплитуда импульса будет зависеть от смещения «зуба» в осевом направлении и радиального зазора. Временной интервал прохождения одноименного зуба под чувствительным элементом определит период и частоту вращения вала.

Программное обеспечение системы реализует сбор информации с измерительных каналов. Обработка информации с двух измерительных каналов позволяет вычислить ОС и РЗ, а также частоту вращения вала по двум каналам. В программе определения ОС и РЗ используется численный метод вычисления координатных смещений [2]. Информация о РЗ используется только для повышения точности измерения осевого смещения вала.

Результаты измерений оборотов и осевого перемещения непрерывно фиксируются памяти ЭВМ, а также выводятся на экран монитора в виде таблицы со сдвигающейся строкой с периодичностью 1 секунда. При достижении заданного (опасного) осевого смещения турбины система вырабатывает соответствующий сигнал для системы управления силовой установкой.

Результаты эксперимента

Функционирование системы измерения выполнено с использованием стенда, имитирующего работу турбинного привода в части взаимодействия датчиков с объектом. Стенд содержит имитатор корпуса агрегата, с установочными ОВТД1 и ОВТД2, а также измерительный диск с приводом от электродвигателя. Материал корпуса и «зубьев» диска, их геометрические размеры полностью соответствуют объекту. Стенд позволяет изменять осевое смещение и радиальный зазор как в статике, так при вращении измерительного диска. Величина перемещений контролируется средствами измерения с разрешающей способностью 0,01мм.

Сняты градуировочные характеристики измерительных каналов.

$$(1) \begin{cases} C_1 = F_1(X, Y) \\ C_2 = F_2(X, Y) \end{cases},$$

где C_1, C_2 – цифровые коды измерительных каналов с датчиками ОВТД1 и ОВТД2,
 F_1, F_2 – функции преобразования измерительных каналов.

Семейства градуировочных характеристик в графическом виде представлены на рисунке 2 и рисунке 3.

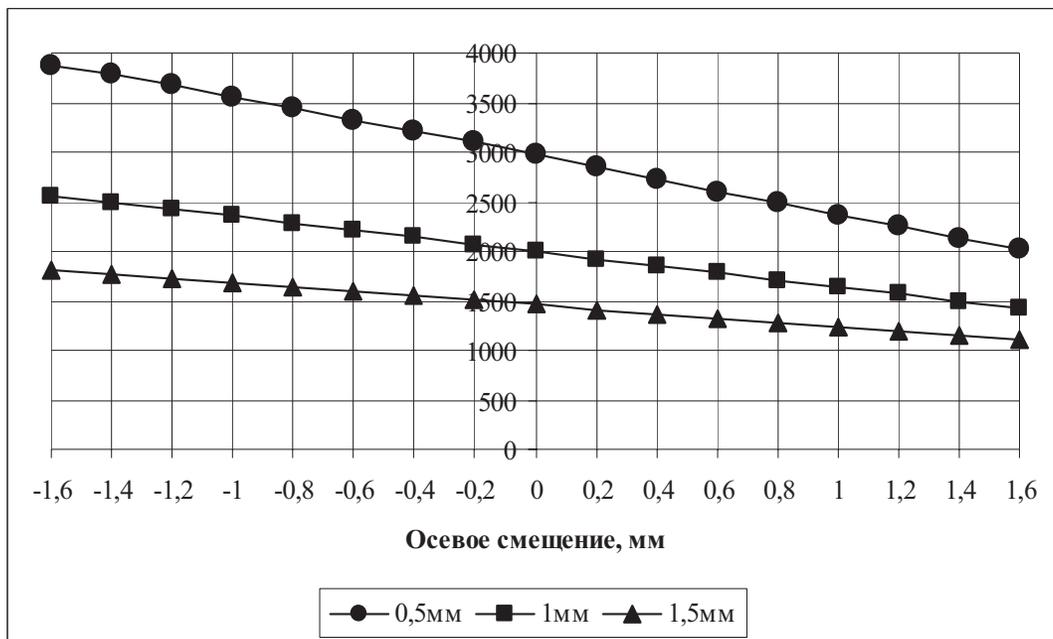


Рисунок 2 – Градуировочные характеристики измерительного канала с ОВТД 1

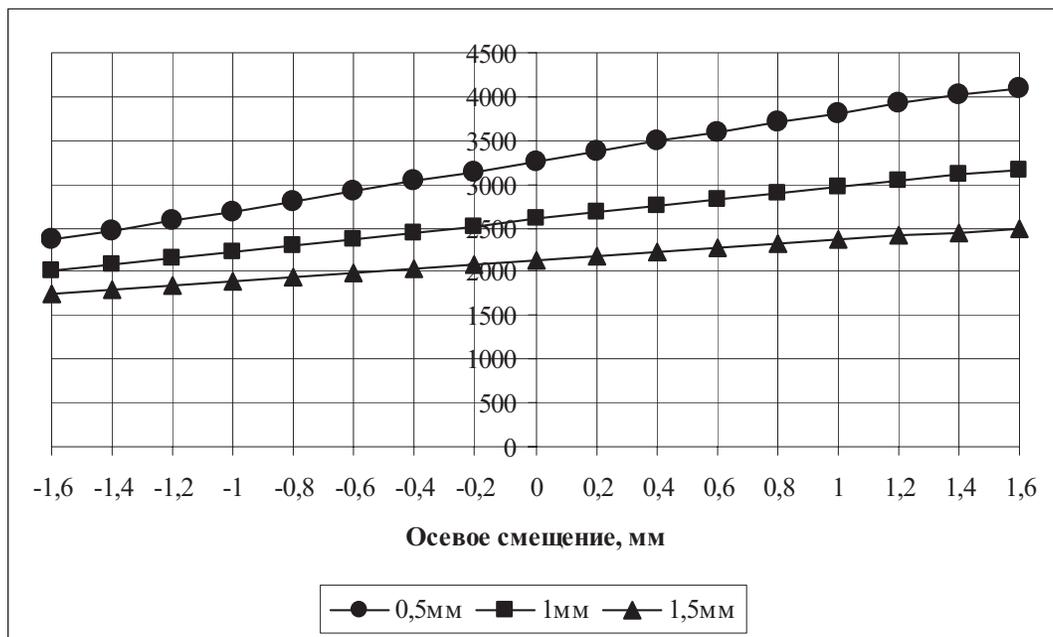


Рисунок 3 – Градуировочные характеристики измерительного канала с ОВТД 2

Оценка основной погрешности определения ОС выполнена для «сквозного» канала в статическом режиме. Для этого с помощью средств перемещения задавались осевые перемещения X_3 и зазоры Y_3 , значения которых сдвинуты на полшага относительно координат, принятых в градуировочных таблицах. В заданных точках X_3 , Y_3 фиксировались вычисленные системой осевые перемещения X_6 (зазор является мешающим фактором, его влияние учитывается градуировочными характеристиками). В заданных точках X_3 , Y_3 определены значения абсолютной погрешности Δx , как отклонения вычисленных от «эталонных» значений, и приведенная относительная погрешность для осевого смещения δx .

$$(2) \Delta x = X_6 - X_3;$$

$$(3) \delta_x = \frac{\Delta x}{X_{\max} - X_{\min}}$$

Относительная приведенная погрешность представлена в графическом виде на рисунке 4. Величина погрешности измерения ОС во всем диапазоне перемещений не превышает 0,17 %.

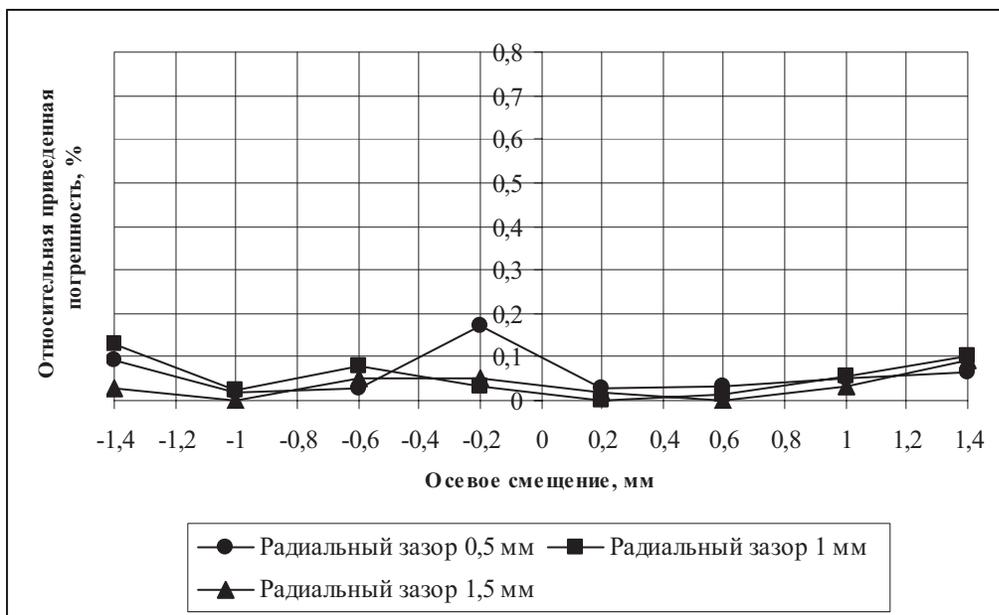


Рисунок 4 – Относительная приведенная погрешность вычисления осевого перемещения

Для оценки случайной составляющей погрешности выполнена регистрация 250 последовательных результатов измерений осевого смещения. Вычисленная величина среднеквадратического отклонения не превышает 0,02 мм.

Оценка измерения оборотов выполнена на частоте вращения электродвигателя путем сопоставления результатов замера с помощью системы измерения и с помощью штатного программного обеспечения фирмы L-CARD программой LGRAF.

Относительная приведенная погрешность измерительных каналов частоты вращения приведена в таблице 1 и не превышает 0,14%.

Таблица 1 – Относительная приведенная погрешность измерительных каналов частоты вращения

Каналы частоты вращения	Частота вращения, измеренная с помощью LGRAF, об/мин	Частота вращения по каналам, об/мин	Относительная приведенная погрешность измерительных каналов частоты вращения, %
Канал 1	1445,7	1447,3	0,10492
Канал 2	1445,7	1447,1	0,13932

Заключение

На практическом примере подтверждена эффективность предложенного подхода. Замена стандартных датчиков на ОБТД позволила получить дополнительно важную информацию для диагностики состояния агрегата без доработки его конструкции.

Список литературы

- [1] Методы и средства измерения многомерных перемещений элементов конструкций силовых установок // под ред. Секисова Ю.Н., Скобелева О.П. – Самара, 2001 – 186 с
- [2] Бахвалов Н.С. Численные методы. - М.: Наука. 1975.

ЭКСПЕРИМЕНТАЛЬНАЯ ПРОВЕРКА РАБОТОСПОСОБНОСТИ КЛАСТЕРНЫХ МЕТОДОВ ИЗМЕРЕНИЯ КООРДИНАТ СМЕЩЕНИЯ ТОРЦОВ ЛОПАТОК И ДЕФОРМАЦИЙ СТАТОРА В ЛАБОРАТОРНЫХ УСЛОВИЯХ

В.А. Белослудцев, С.Ю. Боровик, М.М. Кутейникова, Б.К. Райков,
Ю.Н. Секисов, О.П. Скобелев, В.В. Тулунова

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
m.kuteynikova@mail.ru
тел: +7 (846) 332-26-77, факс: +7 (846) 333-27-70

Ключевые слова: имитатор объекта, проверка работоспособности, кластерные методы, действующий макет системы

Abstract

The functioning of rotor wheel simulator in laboratory conditions is given. The description of the workable prototype of the measuring system as well as the metrological experiments' results and the results of simulator behavior analysis in work conditions are considered too.

Введение

В коллективной монографии [1], изданной около десяти лет назад, впервые в систематизированном виде были изложены основы кластерных методов измерения координат смещения торцов лопаток с помощью одновитковых вихретоковых датчиков (ОВТД) и принципы построения систем, реализующих эти методы. Там же приведены описания конкретных систем, которые использовались на предприятиях, занимающихся созданием газотурбинных двигателей (ГТД). Однако, практически во всех системах, предназначенных для стендовых испытаний ГТД, число ОВТД в составе кластера по тем или иным причинам ограничивалось, и решаемая задача сводилась к измерению лишь одной координаты - радиальных смещений лопаток, определяющих радиальные зазоры (РЗ).

Все последующие годы характеризуются интенсивным развитием и совершенствованием кластерных методов измерения многокоординатных смещений торцов лопаток, а также методов измерения деформаций статора на их основе, что нашло отражение в публикациях [2-6]. В то же время в отечественном двигателестроении затормозилась, а на многих предприятиях практически остановилась разработка ГТД, и, как следствие, прекратились стендовые и летные испытания новых изделий. В этой связи оказалось невозможной существовавшая до последнего времени непрерывная экспериментальная проверка работоспособности и эффективности новых методов и средств измерения в «огневых условиях» производственных испытаний (несмотря на актуальность и востребованность тематики, подтверждаемую отечественными специалистами на совещаниях и конференциях, посвященных развитию ГТД, а также зарубежными публикациями [7, 8]).

Поэтому экспериментальную проверку новых разновидностей рассматриваемых методов и принципов построения систем, реализующих эти методы, было решено проводить в лабораторных условиях на имитаторе объекта и с помощью действующего макета технических и программных средств, обеспечивающих получение информации о координатах смещения торцов лопаток (в том числе РЗ) и деформациях статора.

В настоящей работе приводится описание специально разработанных имитатора объекта, технических и программных средств макета системы, а также результатов метрологических экспериментов и результатов анализа поведения имитатора в рабочем состоянии.

1 Имитатор объекта

Имитатор объекта состоит из двух основных узлов – имитатора статора и имитатора ротора (рисунок 1).

Имитатор статора представляет собой сварную несущую конструкцию с основанием, выполненную из стальных уголков и обеспечивающую крепеж высокотемпературных ОВТД в четырех точках (1, 2, 3, 4) с угловым шагом 90 град. При этом каждая точка оборудована механизмом смещений чувствительных элементов (ЧЭ) датчиков ОВТД в радиальном направлении, контролируемым индикаторами часового типа (I_1, I_2, I_3, I_4) промышленного изготовления с разрешающей способностью 0,01 мм.

Имитатор ротора – колесо с лопаточным венцом (элемент конструкции реального компрессора ГТД) и электропривод (двигатель постоянного тока с регулируемой скоростью вращения). С помощью специального механизма обеспечивается осевое смещение колеса, а, следовательно, торцов лопаток относительно ЧЭ. При этом осевые смещения одинаковы во всех четырех точках крепления ОВТД, и они контролируются дополнительным индикатором (I_0).

Для получения информации о двух координатах смещений (осевом и радиальном)¹ в точке контроля 1 размещен кластер из двух датчиков – ОВТД₁₁ и ОВТД₁₂, которые закреплены под углом 90 град и таким образом, чтобы их ЧЭ (ЧЭ₁₁ и ЧЭ₁₂) в зоне контроля торцов лопаток были параллельны².

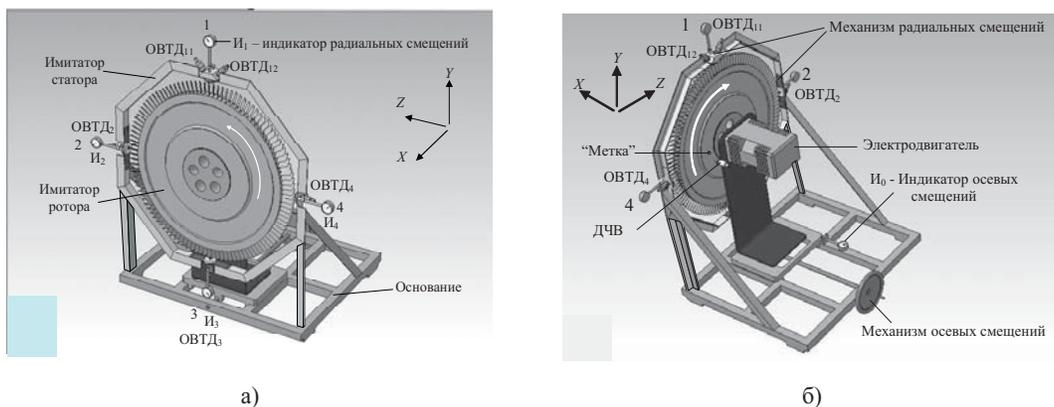


Рисунок 1 – Имитатор объекта (а), вид со стороны электропривода (б)

Перемещение кластера датчиков в радиальном направлении осуществляется с помощью винта, который вызывает прогиб упругого элемента (УЭ) в виде плоской пластины, закрепленной по торцам на уголке несущей конструкции имитатора статора. При этом прогиб УЭ увеличивается, а РЗ между ЧЭ и торцами лопаток уменьшается. При вращении винта в обратном направлении под действием упругой силы кластер датчиков (и его ЧЭ) возвращаются в исходное состояние.

¹ Получение информации о радиальных и осевых смещениях – это наиболее распространенная задача в многокоординатных измерениях смещений торцов лопаток.

² Чтобы обеспечить параллельность ЧЭ и расстояние между ними около 3 мм тоководы обоих датчиков в торцевой части приходится стачивать

В непосредственной близости к ОВТД₁₁ и ОВТД₁₂ расположена аналогичная пара датчиков, выполняющих компенсационные функции (на рисунке 1 они не показаны).

Подобный механизм смещений датчиков применяется и в остальных контрольных точках 2, 3, 4. Однако, в каждой из указанных точек нет необходимости в использовании кластера из двух датчиков, так как перемещения колеса с лопаточным венцом в осевом направлении в этих точках одинаковы, а потому достаточно одного датчика в каждой из них (ОВТД₂, ОВТД₃, ОВТД₄)³.

Питание электродвигателя производится от сети через лабораторный автотрансформатор (ЛАТР) с выпрямителем. Напряжением на выходе ЛАТР регулируется скорость вращения лопаточного колеса.

Изменениями положения лопаточного колеса в осевом направлении и положения ЧЭ ОВТД в точках 1, 2, 3, 4 при соответствующем контроле (И₀, И₁, И₂, И₃, И₄) обеспечивается градуировка измерительных каналов датчиков, их метрологическая аттестация, а также проверка работоспособности методов и средств получения информации о координатах смещений торцов лопаток и деформациях статора.

2 Технические и программные средства действующего макета системы.

Упрощенная структурная схема технических средств, представленная на рисунке 2, содержит высокотемпературные ОВТД (ОВТД₁₁, ОВТД₁₂, ОВТД₂, ОВТД₃, ОВТД₄), индивидуальные преобразователи индуктивностей ОВТД в напряжение, встроенные в линии связи (П₁₁, П₁₂, П₂, П₃, П₄), датчик частоты вращения (ДЧВ) с формирователем (Ф), выносную плату L-Card (E-440) для ввода аналоговых и дискретных сигналов с датчиков, а также ПЭВМ.

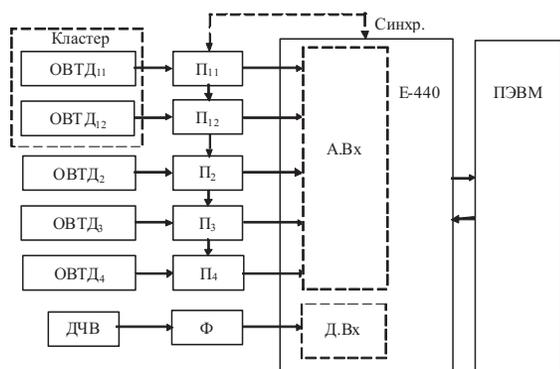


Рисунок 2 – Структурные схемы технических средств

Плата E-440 обеспечивает коммутацию и аналого-цифровое преобразование сигналов ОВТД⁴, а также преобразование в код периодов импульсов с ДЧВ, которые соответствуют скорости вращения лопаточного колеса имитатора объекта. При этом управление перечисленными операциями и первичная обработка полученной с датчиков информации производится с помощью встроенного в плату сигнального процессора.

Программное обеспечение макета содержит, в частности, программы преобразования в код периода вращения лопаточного колеса, коммутации каналов ОВТД и преобразования в код выходных параметров ОВТД, вычисление моментов прохождения лопатками геометрического центра (г.ц.) кластера ОВТД₁₁ и ОВТД₁₂ и выборку кодов соответствующих этим момен-

Индивидуальные преобразователи миниатюризированы. Помимо рабочего ОВТД в смежное плечо мостовой схемы на входе каждого преобразователя включен компенсационный ОВТД. Преобразователи П₁÷П₄ работают с единой частотой импульсного питания (период 6,5 мкс), причем импульсы питания синхронизированы между собой и обеспечивают синхронный запуск устройств в составе платы E-440. Выходные сигналы преобразователей подаются на аналоговые входы (АВх) платы E-440.

Сигналы ДЧВ после формирования поступают на дискретные входы (ДВх) платы E-440.

³ Соответственно уменьшается и число компенсационных датчиков

⁴ Запуск коммутатора и АЦП платы синхронизирован импульсами питания

там времени, программы поиска решения системы уравнения на основе семейства градуировочных характеристик (ГХ) кластера и полученных кодов в каналах ОВТД₁₁ и ОВТД₁₂ (т.е. вычисление координат x и y (РЗ) смещений торцов лопаток) и другие программы.

3 Результаты экспериментальных исследований

Экспериментальные исследования призваны подтвердить работоспособность разработанных методов и средств, причем работоспособность рассматривается в первую очередь, как метрологическая состоятельность измерительных каналов кластера ОВТД действующего макета, оценка которой осуществляется по результатам исследований, проводимых на имитаторе объекта в статических условиях.

Кроме того, приводятся результаты получения информации о РЗ (y) и осевых смещениях (x) торцов лопаток, а также об имитируемых деформациях статора в динамике, т.е. в процессе вращения лопаточного колеса. При этом имитатор рассматривается как объект с неизвестными параметрами, а для подтверждения работоспособности макета полученные результаты сравниваются (сопоставляются) с теми, что найдены в статике.

Метрологическим исследованиям в каналах ОВТД предшествовали оценки погрешностей канала измерения периода вращения лопаточного колеса, представляющего особую значимость для точности последующих вычислений искомых координат, а также его стабильности (допущение постоянства примыкающих периодов вращения лопаточного колеса является необходимым условием получения достоверной информации о координатах смещения торцов лопатки).

Эксперименты показали, что погрешность измерения периода вращения не превышает 0,0025 % в диапазоне 100÷600 об/мин. Нестабильность примыкающих периодов вращения не более 0,05 % при скоростях более 200 об/мин.

Следует отметить, что проведение метрологических исследований потребовало разработки специальных методик получения семейств ГХ в виде зависимостей кодов от координат смещения торцов лопаток лопаточного колеса имитатора и ГХ ОВТД₂÷ОВТД₄. При этом в процессе анализа семейств ГХ ОВТД₁₁ и ОВТД₁₂ была обнаружена потеря чувствительности по координате y из-за взаимного электромагнитного влияния ЧЭ ОВТД, что привело к необходимости сужения диапазона изменений координаты x до $\pm 0,6$ мм (диапазон изменений координаты y – от 0,4 до 1,4 мм).

Систематическая составляющая основной погрешности определяется погрешностью аппроксимации полиномиальной функции двух переменных (x, y) и составляет примерно 0,3 %: для канала с ОВТД₁₁ и 1 % для канала ОВТД₁₂. Случайная составляющая оценивается по СКО и в обоих каналах кластера не превышает 0,6 %.

Кроме того, оценивались алгоритмические погрешности поиска решения системы уравнений на основе семейств ГХ и кодов, соответствующих моменту прохождения лопатками г.ц. кластера (ЧЭ обоих датчиков). Искомые погрешности определяются отклонениями вычисленных при решении системы уравнений координат (x и y) от задаваемых в процессе градуировки. При этом максимальные значения погрешностей для координаты x составляют 1,4 % и координаты y – 2,4 %.

Если учитывать так называемые погрешности «входных данных», определяемые случайной составляющей основной погрешности [9], то максимальные значения алгоритмических погрешностей возрастают до 3,2 % и 5,5 % соответственно.

Систематическая составляющая погрешностей в каналах ОВТД₂÷ОВТД₄ не превышает 1,5 %, а случайная 0,1 %

Далее рассматриваются примеры экспериментов, которые объединяет общая методика их проведения, предусматривающая два этапа. На первом – экспериментальные исследования осуществляются в статических состояниях имитатора объекта с использованием соответст-

вующего контрольно-измерительного инструмента – индикаторов I_0, I_1 - I_4 (рисунок 1) и дополнительных аналогичных индикаторов. На втором – лопаточное колесо приводится во вращение электроприводом, а измерительная информация поступает с действующего макета системы.

В первом эксперименте с помощью кластера ОВТД₁₁ и ОВТД₁₂ исследуются координаты смещений торцов лопаток x и y (РЗ), причем в процессе вращения лопаточного колеса ожидаются изменения обеих координат. При этом одной из возможных причин изменения координаты x являются биения лопаточного колеса в направлении оси X .

На первом этапе, используя статические состояния лопаточного колеса и изменяя его угловое положение, определяются перемещения обода колеса в осевом направлении. Они фиксируются поблизости от места закрепления кластера ОВТД₁₁ и ОВТД₁₂ на имитаторе статора. При этом за начало отсчета принимается положение обода, когда торец лопатки №1 находится в г.ц. кластера. Последующие угловые положения колеса определяются набегающими номерами лопаток, торцы которых оказываются в г.ц.

Полученные результаты приведены на рисунке 3 в виде точек и пунктирной линии 1. Они показывают увеличение x с последующим уменьшением до нуля и аналогичные изменения отрицательных значений x с возвращением к исходному нулевому значению. Представляется

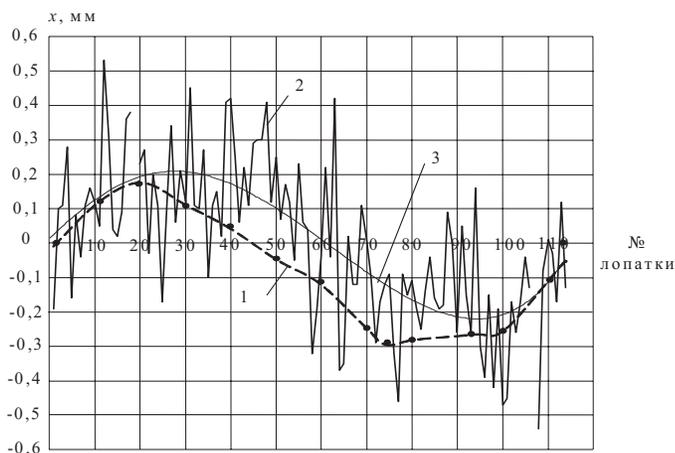


Рисунок 3 – Зависимость координаты x от номера лопатки

очевидным, что в динамике, т.е. при вращении лопаточного колеса от электропривода, обнаруженные в статике изменения x будут вызывать биения в направлении оси X , ожидаемый размах которых составит около 0,45 мм.

Продолжение эксперимента происходит в динамическом режиме – в процессе вращения лопаточного колеса со скоростью 600 об/мин (осевое положение колеса, установленного с помощью индикатора I_0 $x=0$ мм; радиальное смещение, установленное с помощью индикатора I_1 $y=1,0$ мм).

Результаты, полученные в каналах кластера ОВТД₁₁ и ОВТД₁₂ после преобразования в код и вычисления координат в моменты прохождения г.ц. лопатками, представлены на рисунке 3 («пульсирующая» линия 2). С помощью встроенной функции аппроксимации программы MS Excel на том же графике построена «линия тренда» (3). Сравнение «линии тренда» с x -смещениями обода в статике подтверждает наличие биений с ожидаемой амплитудой около 0,45 мм.

Последующий анализ связан с «пульсациями» (2) координаты x (рисунок 3). Чтобы уменьшить влияние биений, «пульсации» рассматриваются на участке от лопатки 26 до 35, где изменения x от биений (3) незначительны и не превышают 0,01 мм. На этом участке вычисляются приращения «пульсации», т.е. изменения осевых смещений Δx торцов лопаток относительно друг друга в парах 26-27, 27-28, ..., 34-35 (результаты вычислений представлены на рисунке 4).

Анализ причин наблюдаемых пульсаций показывает, что они в основном определяются неравномерностью шага торцов лопаток по углу.

И, наконец, в заключение, на рисунке 5 приведен график изменения координаты y (PЗ) по всем лопаткам (соответствующие графики координаты x представлены на рисунке 3). Из графика следует, что за полный оборот колеса координата y изменяется в незначительных пределах от лопатки к лопатке (это связано с пластической вытяжкой лопаток в процессе эксплуатации двигателя). Поскольку характеристики газотурбинного тракта оцениваются по средним значениям PЗ, то оно вычислено и составляет $0,544 \text{ мм}^5$, причем отклонения PЗ по лопаткам не превышают $0,11 \text{ мм}$.

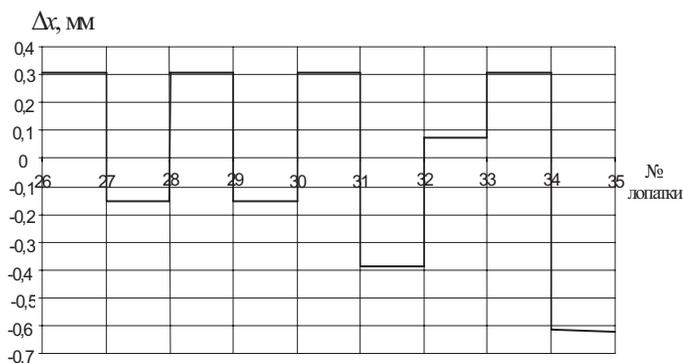


Рисунок 4 – Приращения (отклонения) осевых смещений

данных об осевых биениях лопаточного колеса, а также о неравномерных расстояниях между торцами лопаток, которые недоступны при использовании одного датчика. Это свидетельствует об эффективности рассматриваемого метода и реализующих его средств, и одновременно об их работоспособности.

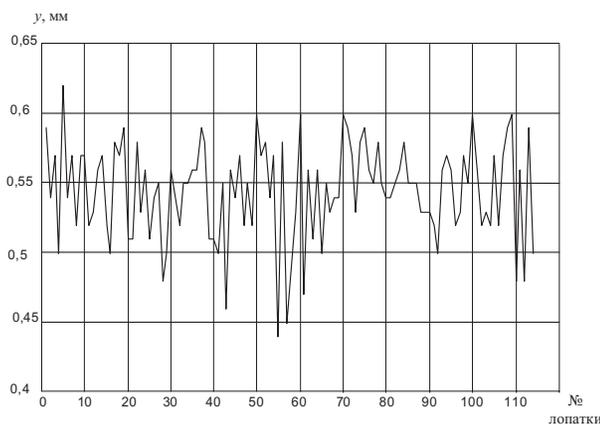


Рисунок 5 – Изменения координаты y (PЗ)

Предполагается, что в исходном состоянии деформация статора отсутствует, а профиль его сечения близок к окружности и PЗ между ЧЭ датчиков и торцами лопаток в точках 1, 2, 3, 4 (рисунок 1) равны и составляют, например, $0,8 \text{ мм}$.

На втором этапе лопаточное колесо приводится во вращение с помощью электропривода. Полученная в измерительных каналах ОВТД₁₁ и ОВТД₁₂, ОВТД₂, ОВТД₃ ОВТД₄ информация о PЗ (в точках 1 – 4 по лопаткам 1, 28, 57, 86 после соответствующих преобразований и вычислений) представлена в таблице:

Кроме того, необходимо подчеркнуть, что рассматриваемый кластерный метод двухкоординатного измерения, реализуемый в действующем макете системы, не только обеспечивает более достоверную информацию о PЗ, но и получение новых

Второй эксперимент проводится с целью проверки работоспособности действующего макета системы для получения информации о деформациях статора. Деформация имитируется перемещением ОВТД относительно крепежной конструкции (рисунок 1). На первом этапе в статическом состоянии лопаточного колеса с помощью индикатора И₀ устанавливается координата $x=0$. При этом лопатка №1 находится в области кластера ОВТД₁₁ и ОВТД₁₂, а лопатки 28, 57, 86 – вблизи ОВТД₂, ОВТД₃, ОВТД₄ соответственно (в таком положении лопаточного колеса снимались ГХ измерительных каналов указанных датчиков).

⁵ Причина отклонения среднего значения PЗ от заданного (1 мм) связана, в первую очередь, с угловым смещением вала в подшипнике электродвигателя под действием возникающей при вращении лопаточного колеса аэродинамической силы и деформациями опорной стойки.

Состояние статора	Номера точек			
	1	2	3	4
Деформации нет	0,78 мм	0,82 мм	0,85 мм	0,805 мм
Деформация есть	0,79 мм	1,18 мм	0,84 мм	1,22 мм

Как следует из данных, приведенных в таблице (строка 1), вычисленные РЗ отличаются от заданных в статике значений не более, чем на 0,05 мм.

Далее имитируется деформация статора: РЗ в точках 2 и 4 увеличиваются на 0,4 мм и профиль его сечения приобретает форму эллипса. Результаты, полученные в процессе вращения лопаточного колеса, приведены в той же таблице (строка 2). Изменения РЗ по сравнению с ожидаемыми отличаются не более чем на 0,02 мм, что также подтверждает работоспособность рассматриваемых методов и реализующих их средств.

Заключение

Разработаны и изготовлены имитатор объекта и действующий макет системы, в котором реализованы кластерные методы измерения координат смещения торцов лопаток деформации статора. Экспериментальные исследования погрешностей измерительных каналов (в частности, каналов кластера датчика при статических состояниях лопаточного колеса имитатора) подтвердили их метрологическую состоятельность, результаты измерения радиальных и осевых смещений торцов лопаток, а также имитируемых деформаций статора, полученные в процессе вращения лопаточного колеса, показали работоспособность и эффективность указанных методов.

Список литературы

- [1] Методы и средства измерения многомерных перемещений элементов конструкций силовых установок /под ред. Секисова Ю.Н., Скобелева О.П. - Самара: Самарский НЦ РАН, 2001, 188 с.
- [2] Боровик С.Ю., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Измерение и вычисление координатных составляющих многомерных перемещений торцов лопаток в процессе вращения ротора // Автометрия, №2, 2001, с. 103-111.
- [3] Секисов Ю.Н., Скобелев О.П. Измерение составляющих многокоординатных смещений элементов конструкций силовых установок с моделированием неизмеряемых составляющих (концепция и реализующие ее методы измерения) // Мехатроника, автоматизация, управление, 2004. -№7. – с. 29-35.
- [4] Боровик С.Ю., Секисов Ю.Н., Скобелев О.П. Обобщенное представление методов получения измерительной информации о координатах смещений торцов лопаток и лопастей // Мехатроника, автоматизация, управление / Управление и информатика в авиакосмических системах / Информационно-измерительные и управляющие системы летательных аппаратов / приложение к журналу №3, 2007, - с. 19-24.
- [5] Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П. Метод получения информации о деформациях статорной оболочки винтовентилятора, радиальных зазорах, смещениях ее геометрического центра и оси винта // Проблемы управления и моделирования в сложных системах: Труды X Международной конференции. Самара:, 2008 с.290-297.
- [6] Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П. Метод получения информации о деформациях статора, радиальных зазорах, смещениях его центра и биениях ротора // Проблемы управления и моделирования в сложных системах: Труды X Международной конференции. Самара:, 2008 с.298-304.
- [7] S. Lattime, B. Steinetz Test Rig for Evaluating Active Turbine Blade Tip Clearance Control Concepts // 39th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Huntsville, Alabama, July 20-23, 2003.
- [8] D.Simon, Sanjay Gang, G. Hunter, Ten-Huei Guo, K. Semega Sensor Needs for Control and Health Management of Intelligent Aircraft Engines // ASME Turbo Expo 2004, Vienna, Austria, 14-17 June, 2004
- [9] А.А. Инюцин, В.В. Тулупова Анализ погрешностей алгоритмов вычисления координат смещения торцов лопаток и лопастей // Проблемы управления и моделирования в сложных системах: Труды VII Международной конференции. Самара: СамНЦ РАН, 2005. – С.181-188.

АЛГОРИТМ ВЫЧИСЛЕНИЯ РАДИАЛЬНЫХ И ОСЕВЫХ СМЕЩЕНИЙ ТОРЦОВ ЛОПАТОК

М.М. Кутейникова, Ю.Н. Секисов

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
m.kuteynikova@mail.ru
тел: +7 (846) 332-26-77, факс: +7 (846) 333-27-70

Ключевые слова: система измерения, алгоритм, координатные составляющие, одновитковый вихретоковый датчик

Abstract

The article describes an algorithm for computing the two coordinate components of displacement in relation to measurements using cluster of single-coil eddy-current sensors. The algorithm provides expanding the range of measurements desired coordinates due to using no monotonic plot for one of the coordinates of the calibration characteristics.

Введение

Измерения радиальных зазоров (РЗ) между торцами рабочих лопаток компрессора и турбины, а также осевых смещений (ОС) лопаточных колес выполняется с использованием кластера, состоящего из двух одновитковых вихретоковых датчиков (ОВТД1, ОВТД2). Система измерений РЗ и ОС содержит измерительные каналы (ИК), обеспечивающие преобразование сигналов датчиков в цифровые коды соответственно $C1$ и $C2$. Определение значений РЗ и ОС выполняется на основе градуировочные характеристики (ГХ) измерительных каналов и значения текущих кодов с ИК. В общем случае для поиска координат смещения (КС) по имеющимся градуировочным характеристикам используются численные методы, которые, в частности, ориентированы на использование полиномиальных функций нескольких переменных, кусочно-линейную интерполяцию, метод Ньютона, метод простых итераций, метод секущих и др.[1-4]. В [4] приведена геометрическая интерпретация принципа определения двух координатных составляющих перемещения. В системе координат XYZ можно принять, что изменение РЗ выполняется по координатной оси Y, ОС – по оси X, а по оси Z откладывается величина цифрового кода. В этом случае градуировочные характеристики измерительных каналов будут иметь следующий вид

$$(1) \quad \left. \begin{aligned} C_1 &= F_1(X, Y), \\ C_2 &= F_2(X, Y) \end{aligned} \right\}$$

где $C1$, $C2$ – цифровые коды измерительных каналов с датчиками ОВТД1 и ОВТД2, $F1$, $F2$ – функции преобразования измерительных каналов.

Семейства градуировочных характеристик (1) в координатной системе XYZ представляют собой поверхности, а кодам с измерительных каналов в момент измерения $C_{изм1}$ и $C_{изм2}$ соответствуют поверхности, параллельные плоскости XY. Поверхности $C1=F1(X, Y)$ и $C2=F2(X, Y)$ образует с соответственно с плоскостями $C_{изм1}$, $C_{изм2}$ линии пересечения, которые на плоскости XY имеют проекции $Q1$ и $Q2$. Искомые РЗ и ОС определяются координатами точки пересечения не плоскости XY проекций $Q1$ и $Q2$.

Применительно к задачам определения координат смещений для поиска решения используется метод секущих Ньютона [1]. Метод представляет собой итерационный процесс, кото-

рый характеризуется повторяющимся построением «секущих» линий на проекциях $Q1$, $Q2$. Координаты точки пересечения «секущих» на последней итерации является искомыми значениями РЗ и ОС. Недостатком метода является условие «монотонности» семейства ГХ, то есть отсутствие локальных экстремумов на проекциях $Q1$ и $Q2$. В области локального максимума или минимума метод не применим, так как секущая должна принять практически горизонтальное положение [5]. Реальные ОВТД имеют участок с минимумом (максимумом) на градуировочной характеристике по координате X, в связи с этим необходимо смещать рабочий участок ГХ от области минимума и, следовательно, существенно ограничивать рабочий диапазон датчика.

Предлагаемый алгоритм

Расширение рабочего диапазона ОВТД в направлении осевых перемещений в предлагаемом алгоритме достигается за счет использования немонотонных участков градуировочных характеристик кластера ОВТД.

Основой вычислительных операций являются также семейство градуировочных характеристик ОВТД (1), представленных в виде таблиц, и коды измерительных каналов $C_{изм1}$ и $C_{изм2}$.

Геометрическая интерпретация предлагаемого алгоритма. На плоскости XY создается координатная сетка, для этого через координатные точки, указанные в градуировочных таблицах, проводятся линии, параллельные координатным осям. В результате будет создано множество ячеек, ограниченных координатными линиями, в узлах которых размещены коды градуировочных таблиц (Рисунок 1). Гипотетически линии $Q1$, $Q2$ накладываются на сетку плоскости XY (на рисунке показаны пунктирной линией). Фактически алгоритм не требует нахождения поверхностей $C1=F1(X,Y)$ и $C2=F2(X,Y)$, линий пересечения поверхностей с плоскостями $C_{изм1}$ и $C_{изм2}$, а также проекций $Q1$, $Q2$.

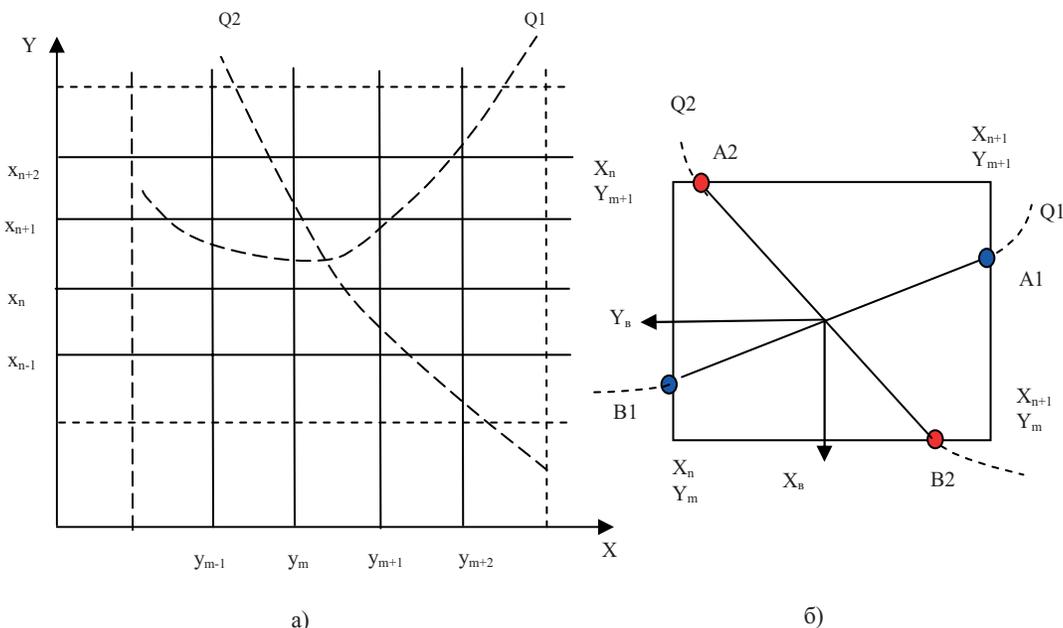


Рисунок 1 – Геометрическая интерпретация алгоритма
 а) Сетка градуировочной таблицы б) ячейка с пересечением $Q1$ и $Q2$

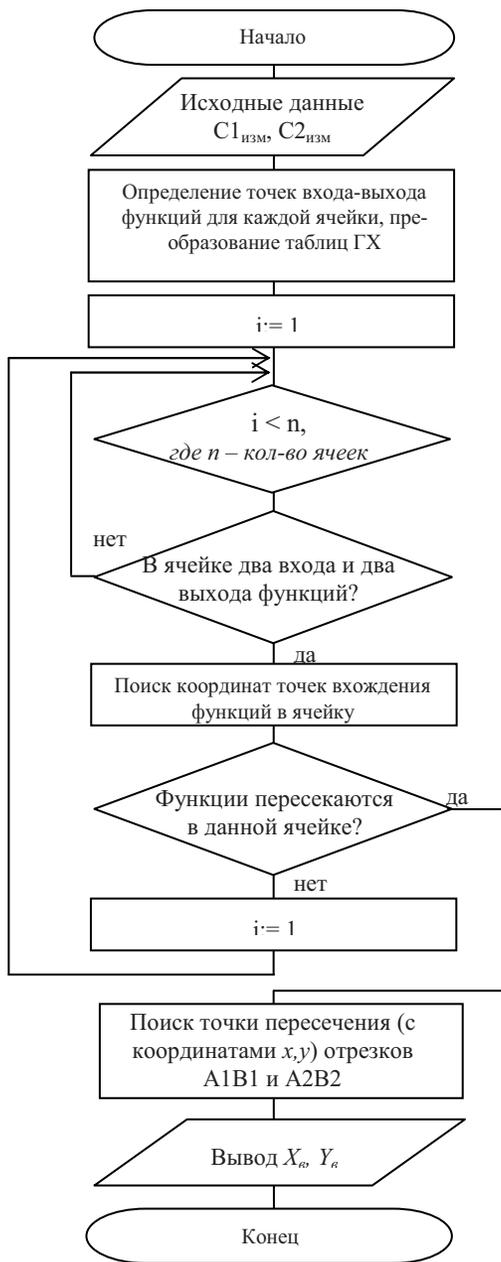


Рисунок 2 – Блок-схема алгоритма

их плотностью). Полученные характеристики в области $X=2$ мм и $X=-2$ мм.

Для оценки погрешности определения РЗ и ОС сформирована таблица «эталонных» координатных смещений X_s, Y_s . Значения X_s и Y_s , сдвинутые на полшага относительно координат, принятых в градуировочных таблицах. Выбраны следующие значения перемещений в радиальном направлении от 0,15 мм до 1,95 мм с шагом 0,3 мм и осевом направлении от -2,4 мм

На первом этапе алгоритм выполняет поиск ячейки, в поле которой должны пересекаться линии $Q1, Q2$. На втором этапе, используя линейную интерполяцию, вычисляются координаты точек входа – выхода линий $Q1, Q2$ в ячейку по величине кодов в прилегающих узлах. При этом точки входа $A1(X_{11}, Y_{11}), A2(X_{21}, Y_{21})$ и выхода $B1(X_{21}, Y_{21}), B2(X_{22}, Y_{22})$, могут располагаться на любой из четырех сторон ячейки. На третьем этапе находятся уравнения отрезков $A1B1$ и $A2B2$, координаты точки пересечения которых являются искомыми значениями РЗ и ОС. На рисунке 2 приведена блок-схема алгоритма, обеспечивающего нахождение измеряемых РЗ и ОС. Разработана реализующая приведенный алгоритм программа.

Оценка работоспособности алгоритма

Цель тестирования является проверка возможности вычисления РЗ и ОС во всем диапазоне перемещений по осям X и Y , а также определение погрешности вычисления РЗ и ОС.

Тестирование алгоритма должно проводиться с «эталонными» градуировочными характеристиками и набором «эталонных» кодов $C_{изм1}, C_{изм2}$. Градуировочные характеристики и коды $C_{изм1}, C_{изм2}$, полученные в натурном эксперименте, содержат погрешности технических средств и не могут быть использованы для объективной оценки характеристик алгоритма. Создана математическая модель, достаточно близко воспроизводящая характеристики $F1$ и $F2$ реальных датчиков ОВТД1 и ОВТД2 в составе кластера. Сформированы градуировочные таблицы, в которых значения $C1$ и $C2$ вычислены по оси X с шагом 0,2 мм в пределах -2,6...2,6 мм и по оси Y с шагом 0,1 мм в пределах 0,1 ... 1,9 мм. Семейства градуировочных характеристик представлены в графическом виде на рисунках 3 и 4 (показано часть графиков в связи с высокой

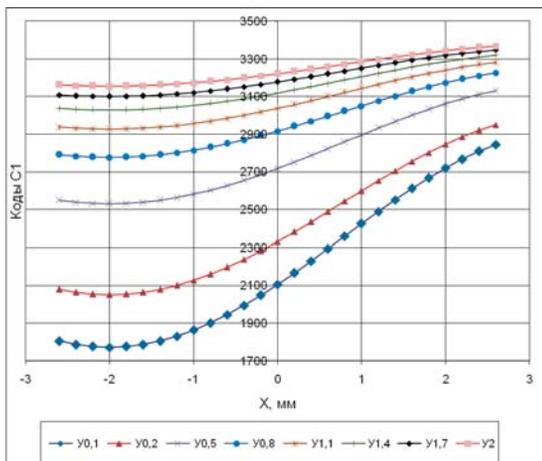


Рисунок 3 – Семейство ГХ для C1

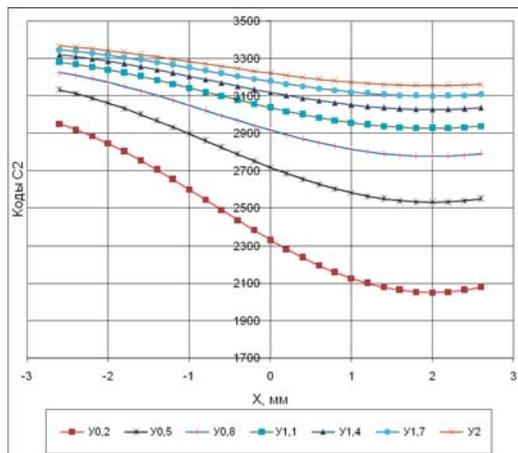


Рисунок 4 – Семейство ГХ для C2

до 2,4 мм с шагом 0,8 мм. С помощью функций $F1$ и $F2$ вычислены «эталонные» коды $C_{изм1}$, $C_{изм2}$, которые для разработанной программы являются исходными для вычисления искомым РЗ и ОС. Для заданных комбинаций $C_{изм1}$, $C_{изм2}$ вычислены значения радиального зазора Y_B и осевого перемещения X_B , а также значения абсолютной погрешности Δx , Δy в заданных точках X_3, Y_3 как отклонения вычисленных от «эталонных» значений

$$(2) \quad \Delta x = X_B - X_3; \quad \Delta y = Y_B - Y_3.$$

Зависимости абсолютных погрешностей в графическом виде представлены на рисунках 5 и 6 (показано часть графиков в связи с высокой их плотностью).

Погрешность вычисления РЗ практически не зависит от осевого смещения торца лопатки (Рисунок 5), включая зоны перегиба градуировочных характеристик в точках $X = -2$ мм и $X = 2$ мм. Величина погрешности возрастает с уменьшением зазора и при 0,15 мм не превышает значения 4,5 мкм.

Величина погрешности вычисления Y_B возрастает как с увеличением зазора, так и с увеличением смещения торца лопатки от центра кластера в осевом направлении (Рисунок 6). Во всем диапазоне осевых и радиальных перемещений величина абсолютной погрешности вычисления ОС составляет единицы нанометров и может не учитываться.

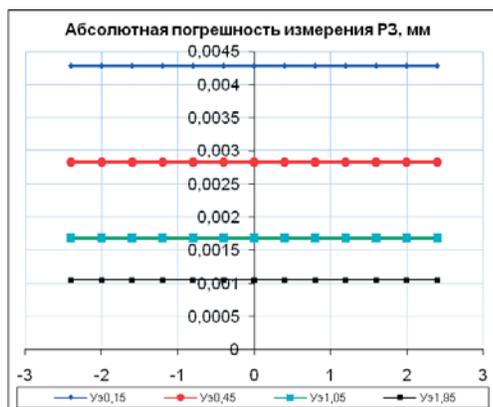


Рисунок 5 – Абсолютная погрешность по Y

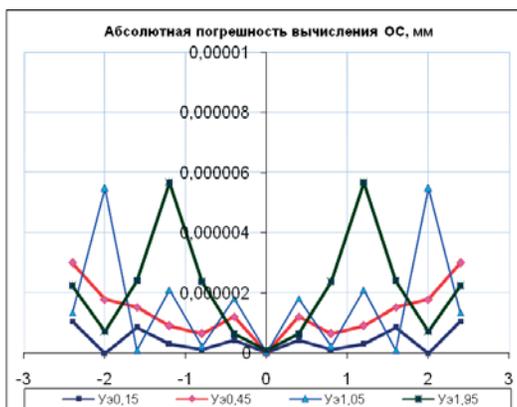


Рисунок 6 – Абсолютная погрешность по X

Заключение

Показано, что предложенный алгоритм для функций двух переменных обеспечивает устойчивое вычисление РЗ и ОС во всем диапазоне перемещений, включая области экстремумов (перегиба) градуировочных характеристик по координате X , при этом на погрешность вычисления как радиального зазора, так и осевого перемещения область перегиба не оказывает практического влияния. Использование области перегиба существенно расширяет рабочую область кластера датчиков в осевом направлении.

Благодарности

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, грант № 08-08-00422а.

Список литературы

- [1] Секисов Ю.Н. Методы и средства измерений многомерных перемещений элементов конструкций силовых установок: Дисс. Докт. техн. Наук. – Самара, 1999 – 267 с.
- [2] Тулупова В.В., Филимонов Р.В. Алгоритмические и программные средства аппроксимации градуировочных характеристик систем измерения полиномиальными функциями нескольких переменных. // Проблемы управления и моделирования в сложных системах: Труды III международной конференции. – Самара, 2001. – с. 573 – 577.
- [3] Тулупова В.В.. Системы измерения многокоординатных смещений торцов лопаток компрессора и лопастей винтовентилятора: Дисс. ... канд. техн. наук. – Самара, 2005. – 200 с.
- [4] Методы и средства измерения многомерных перемещений элементов конструкций силовых установок // под ред. Секисова Ю.Н., Скобелева О.П. – Самара, 2001 – 186 с.
- [5] Бахвалов Н.С. Численные методы. - М.: Наука. 1975.

СИСТЕМАТИЗАЦИЯ И КЛАССИФИКАЦИЯ ГРУППОВЫХ ИЗМЕРИТЕЛЬНЫХ ЦЕПЕЙ ОДНОВИТКОВЫХ ВИХРЕТОКОВЫХ ДАТЧИКОВ¹

Л.Б. Беленький, С.Ю. Боровик, Б.К. Райков, О.П. Скобелев

Институт проблем управления сложными системами РАН

443020, Самара, ул. Садовая, 61, Россия

borovik@iccs.ru

тел: +7 (846) 333-26-77, факс: +7 (846) 333-27-70

Ключевые слова: групповая измерительная цепь, сосредоточенный кластер, распределенный кластер, чувствительный элемент, структурные схемы, временные диаграммы

Abstract

The paper describes the classification of group measuring circuits. The signs of classification are variety of cluster (concentrated or allocated) and number of single-coil eddy-current sensors (sensing elements in the cluster sensor). Systematized descriptions of the block diagrams and theirs operations are given.

Введение

В процессе получения информации о многокоординатных смещениях торцов лопаток наибольшее применение нашли кластерные методы, реализуемые в индивидуальных измерительных цепях (ИИЦ) [1]. При этом с увеличением числа искомых координат соответственно возрастает и число ИИЦ, что приводит к аппаратурной избыточности и, как следствие, увеличению затрат на оборудование и его эксплуатацию.

Между тем, известен подход к построению групповых измерительных цепей (ГИЦ), в которых коммутация осуществляется на уровне датчиков, и не только устраняется аппаратурная избыточность и связанные с нею затраты, но и снижается энергопотребление при использовании параметрических датчиков [2]. В частности, в существующих ГИЦ, ориентированных на однородные индуктивные датчики промышленного изготовления, бесконтактная коммутация выполнялась в цепях питания датчиков, а для преобразования индуктивностей использовался метод первой производной [1, 2]. Аналогичные ГИЦ использовались при работе с так называемыми «экранными» многовитковыми вихретоковыми датчиками со сравнительно высоким уровнем выходной индуктивности и ее девиации [3, 4]. Однако неоднократные попытки построения ГИЦ, предназначенных для преобразования сигналов одновитковых вихретоковых датчиков (ОВТД) и их кластерных разновидностей (КОВТД) сдерживала элементная база и, в первую очередь, существовавшие в то время бесконтактные ключевые элементы и операционные усилители (ОУ).

Действительно, ныне существующие ОВТД (КОВТД) имеют выходные индуктивности не менее чем на порядок ниже индуктивностей «экранных» датчиков, и для преобразования их сигналов требуются малые остаточные параметры ключей в замкнутом состоянии и более высокочастотные ОУ с небольшим дрейфом нуля.

Современная элементная база (ключи на униполярных транзисторах и последние разработки ОУ), удовлетворяя предъявляемым требованиям, открывает путь к созданию ГИЦ, ориентированных на ОВТД (КОВТД), о чем свидетельствуют публикации, появившиеся в последние годы [5-8]. Вместе с тем, в указанных источниках основное внимание концентрирует-

¹ Работа выполнена при частичной поддержке РФФИ (грант №08-08-00422а).

ся на новых кластерных методах, а вопросы, связанные с измерительными цепями и их анализом разбросаны по отдельным публикациям и рассматриваются как вспомогательные и только с точки зрения реализуемости методов. В то же время представляется очевидным, что изучение ГИЦ имеет самостоятельную значимость, а потому в статье существующие ГИЦ, ориентированные на ОВТД (КОВТД), классифицируются в зависимости от вида кластера и числа ОВТД (чувствительных элементов (ЧЭ) КОВТД), а также приводится систематизированное описание структурных вариантов ГИЦ и их функционирования².

1 Структура групповых измерительных цепей

В таблице 1 представлены структурные схемы ГИЦ, в которых осуществляется преобразование сигналов с помощью сосредоточенного кластера (СК) ОВТД (КОВТД) (колонка 3) или распределенного кластера (РК) ОВТД (колонка 4) [5, 7]. Дополнительным классификационным признаком служит число ОВТД в составе кластера (ЧЭ в КОВТД): схемы, обслуживающие два рабочих ОВТД (ЧЭ), занимают строку 1, а схемы, обслуживающие три рабочих ОВТД (ЧЭ) — строку 2.

Структурные схемы ГИЦ (колонка 3) обеспечивают реализацию методов измерения координат смещений торцов лопаток (лопастей) при одновременных (поочередных) преобразованиях и фиксациях кодов [6]. От схем ИИЦ [1] схемы ГИЦ отличаются наличием ключевых элементов (К), которые на каждом из этапов преобразования обеспечивают образование пар рабочих и компенсационного ЧЭ (ЧЭ-Р, ЧЭ-К) в неравновесной мостовой схеме, плечами которой являются обмотки согласующих трансформаторов (СТ) в датчиках и образцовые резисторы R . Синхронное замыкание ключей в парах ЧЭ формирует также импульсное питание моста U_{Π} (амплитуда импульса соответствует напряжению E источника питания). Как и ИИЦ [1], так и ГИЦ работают по методу первой производной, а потому в схемах используются дифференцирующие и запоминающие устройства (ДУ и ЗУ) по одному на кластер ОВТД (КОВТД).

Если в составе СК ОВТД (КОВТД) используются два рабочих ОВТД (ЧЭ, предназначенных для измерения двух координат) и один компенсационный ОВТД («вырожденный» КОВТД с одним ЧЭ), то для функционирования схемы необходимы три ключа — два в плечах моста с рабочими датчиками (K_1, K_2) и один - в плече моста с компенсационным датчиком (K_K) (Таблица 1, колонка 3, строка 1).

Если в составе СК ОВТД (КОВТД) используются три рабочих ОВТД (ЧЭ), (предназначенных для трехкоординатных измерений), то общее число ключей, необходимых для работы схемы возрастает до четырех т.к. в рабочие плечи моста вводится дополнительный ключ (K_3) (Таблица 1, колонка 3, строка 2).

Рассматривая трехкоординатное измерение с помощью РК из трех датчиков (ОВТД₁, ОВТД₂, ОВТД₃) (колонка 4, строка 2), следует отметить наличие в мостовой схеме ГИЦ только трех ключей (K_1, K_2, K_3), а не четырех, как это было в схеме аналогичного назначения, обслуживающей СК датчиков (колонка 3, строка 2). Ключи обеспечивают образование пар рабочих и компенсационных датчиков в плечах мостовой схемы с первичными обмотками СТ и резисторами (R). Синхронное замыкание ключей в парах датчиков формирует также импульсное питание моста (U_{Π}) [8].

² Помещенная в настоящем сборнике статья [10] посвящена анализу основных характеристик и параметров ГИЦ. Рассматриваются обобщенные модели, построенные на основе эквивалентных схем ГИЦ, описания которых приведены в работах [6, 8, 9], причем сравнение и сопоставление полученных результатов моделирования проводится для одних и тех же исходных данных по элементам ГИЦ, в том числе для последних разработок применяемой электронной элементной базы.

Таблица 1 – Измерительные цепи для одновременного преобразования сигналов сосредоточенных и распределенных кластеров одновитковых вихревых датчиков и их кластерных разновидностей

		Вид кластера	
		Распределенный	
		4	
№ п.п.	1	Сосредоточенный	3
К-во ЧЭ	2		
К-во ЧЭ	2		
№ п.п.	2	Сосредоточенный	3
К-во ЧЭ	3		
К-во ЧЭ	3		

Та же ГИЦ может быть использована для двухкоординатного измерения с помощью РК из двух датчиков (ОВТД₁, ОВТД₂). При этом один из ключей (в цепи СТ ОВТД₃) постоянно находится в разомкнутом состоянии или исключается из схемы. Однако, возможен и такой вариант цепи, в котором число ключей минимизируется до одного, а его функции ограничиваются формированием импульсного питания моста (U_{Π}). Этот вариант ГИЦ, предназначенный для обслуживания РК из двух датчиков, размещен в Таблице 1, колонка 4, строка 1 [9].

Дополнительная катушка с индуктивностью $L_{\text{д}}$ позволяет получить на выходе моста и ДУ однополярный сигнал.

2 Функционирование групповых измерительных цепей

В Таблице 2 (колонка 3, строки 1 и 2) представлены временные диаграммы, поясняющие функционирование обеих структурных схем ГИЦ. Здесь и далее в описаниях преобразований, осуществляемых в ГИЦ, также предполагается предварительное измерение периода вращения ротора T_p и вычисление периода T_0 импульсов питания U_{Π} по формуле $T_0 = T_p/n$, где n – число лопаток. Эти импульсы необходимы и для управления элементами ГИЦ, в том числе ключами (эпюры 1). Каждый из этапов (их длительность соответствует периоду вращения) завершается получением искомой цифровой информации по всем лопаткам (лопастям).

При измерении двух координат (схема в Таблице 1, колонка 3, строка 1) на этапе 1 функционирует пара ЧЭ₁-Р и ЧЭ-К. Поэтому на ключи K_1 и K_K одновременно поступают импульсы управления, вызывающие их синхронное замыкание и соответственно подачу питания U_{Π} на мостовую схему (Таблица 2, колонка 3, строка 1, эпюры 2, 4). В результате преобразований на выходе мостовой схемы появляются напряжения U_M , а на выходе ГИЦ формируются напряжения U_{3y} и далее цифровые коды $C_{1-1}, C_{1-2}, \dots, C_{1-n}$, в которых первые цифры индексов — номера рабочих ЧЭ и этапов, вторые — номера лопаток или лопастей (эпюры 5³, 7-9).

Если измеряются три координаты (схема в Таблице 1, колонка 3, строка 2), то на этапах 1 и 2 (т.е. в течение двух периодов T_p) функционирование ключей K_1, K_2, K_K , аналогично рассмотренному (Таблица 2, колонка 3, строка 2, эпюры 2, 3, 5, 6, 7, 9-11), но на этапе 3 работает пара ЧЭ₃-Р и ЧЭ-К. Синхронные замыкания ключей K_3 и K_K вызывают соответствующие преобразования в ГИЦ, которые завершаются формированием кодов $C_{3-1}, C_{3-2}, \dots, C_{3-n}$, (эпюры 4, 8-11)⁴.

На основе полученных значений кодов, а также градуировочных характеристик (ГХ) составляется система уравнений (для каждой лопатки или лопасти) и вычисляются искомые координаты смещений.

Длительность получения информации о координатах смещений составляет три периода вращения ($3T_p$).

В Таблице 2 (колонка 4, строки 1 и 2) изображены временные диаграммы, поясняющие функционирование схем, представленных в Таблице 1 (колонка 4).

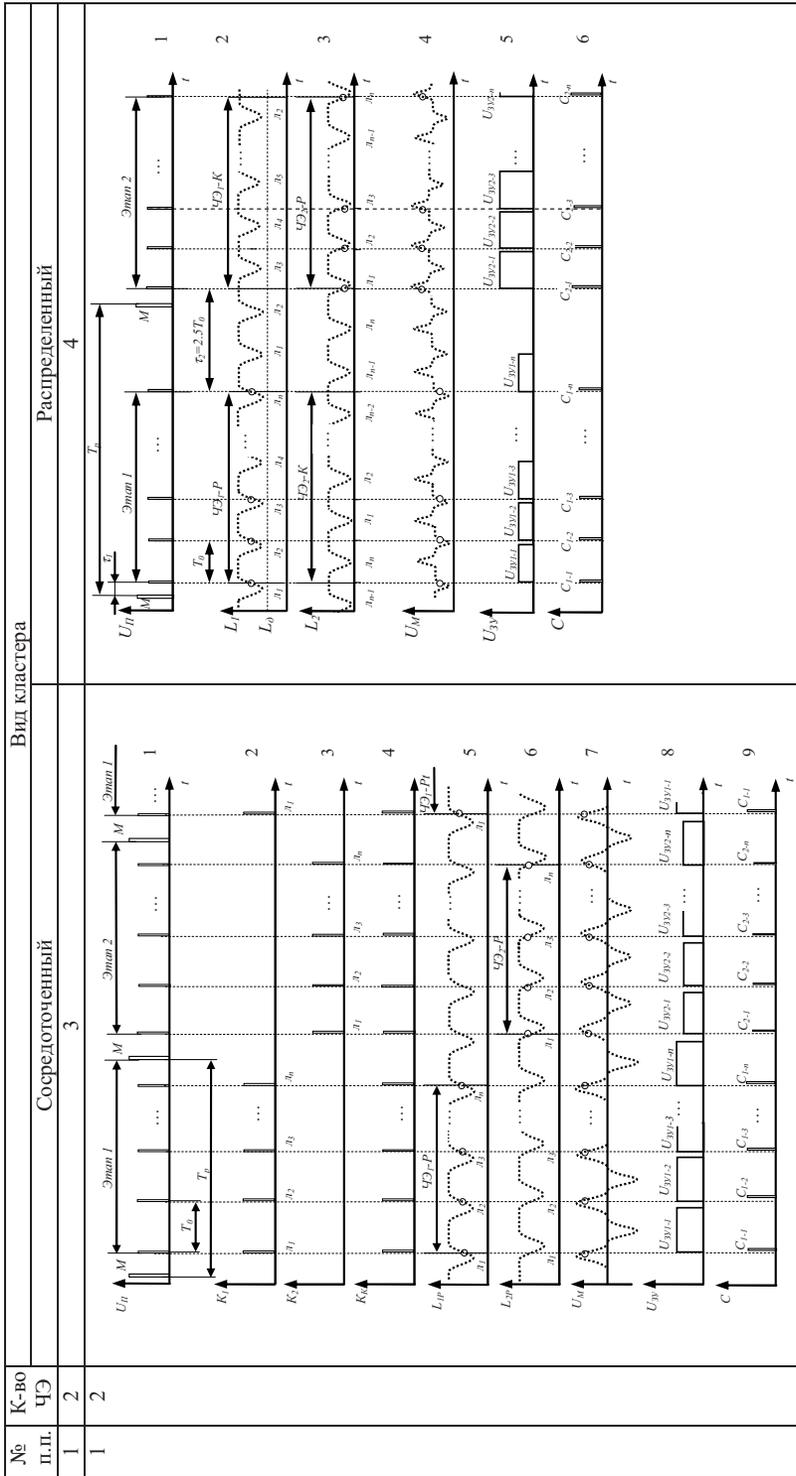
При измерении двух координат (схема в Таблице 1, колонка 4, строка 1) на этапе 1 (первый период T_p) рабочие функции выполняет ЧЭ₁ (ЧЭ₁-Р), а компенсационные — ЧЭ₂ (ЧЭ₂-К). Ключ К формирует импульсы питания U_{Π} (эпюра 1), которым соответствуют напряжения на выходе моста (U_M), ЗУ (U_{3y}) и коды С по всем лопаткам на выходе АЦП (эпюры 2-6).

На этапе 2 (второй период T_p) рабочие функции выполняет ЧЭ₂ (ЧЭ₂-Р), а компенсационные ЧЭ₁ (ЧЭ₁-К). Ключ К формирует импульсы питания с тем же периодом (с задержкой τ_2),

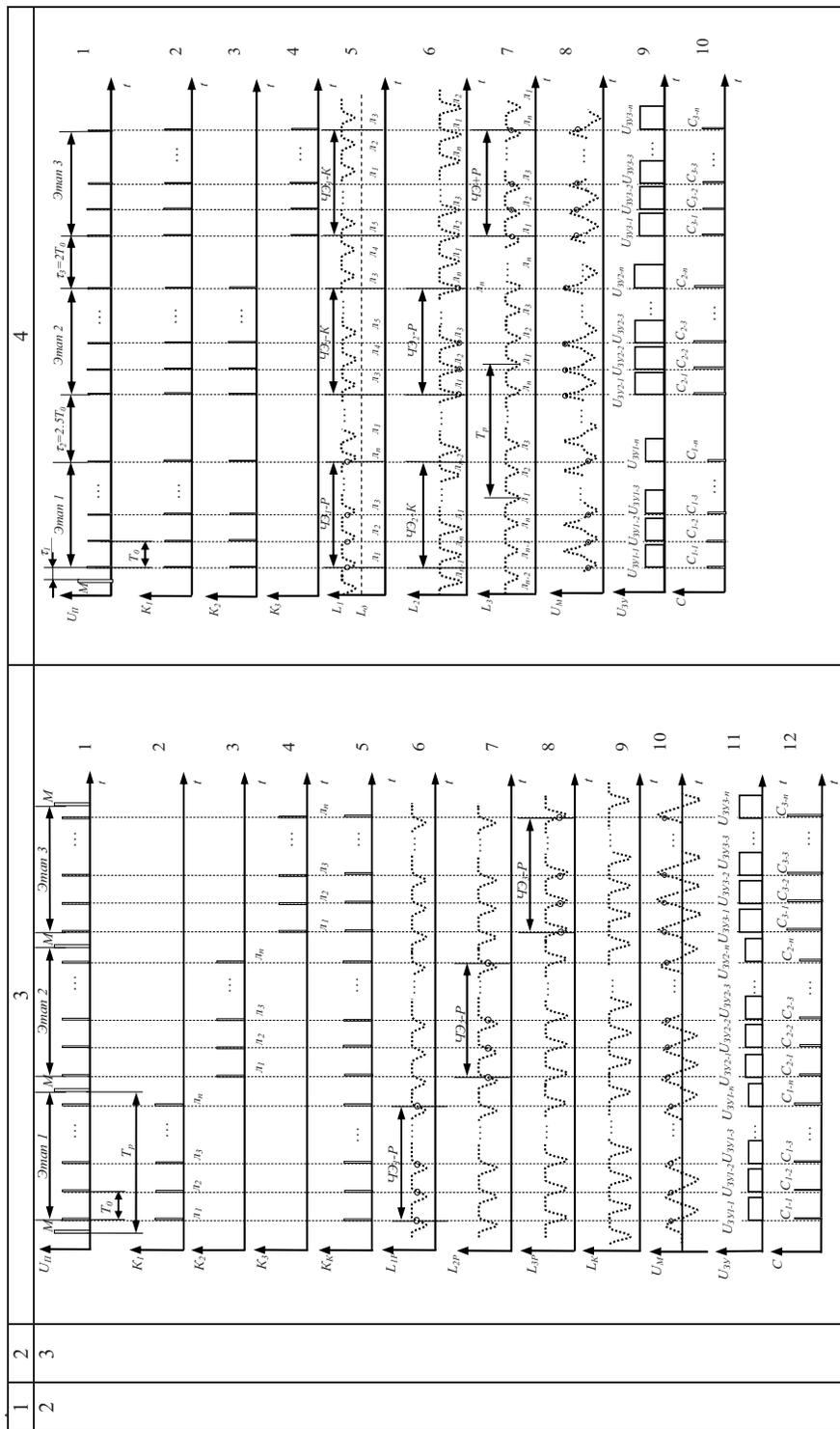
³ Здесь и далее в изображениях эпюр, характеризующих изменения индуктивностей ЧЭ и напряжения на выходе мостовых схем, пунктирными линиями иллюстрируется их поведение при возможном уменьшении периодов импульсов питания. Эпюры напряжения на выходе ДУ не приводятся, так как время получения информативного сигнала составляет около 100 нс, что на несколько порядков меньше периода T_0 [1, 6, 8].

⁴ Эпюры $L_{1P}(t), L_{2P}(t)$ (таблица 2, колонка 3, строка 1) соответствуют параллельному размещению ЧЭ, а эпюры $L_{1P}(t), L_{2P}(t), L_{3P}(t)$ (строка 2) – размещению ЧЭ типа «треугольник» со смещенным центром [6].

Таблица 2 – Временные диаграммы функционирования групповых измерительных цепей



Продолжение таблицы 2



которым соответствуют изменившиеся напряжения U_M , U_{3Y} и коды C по всем лопаткам (те же эпюры 2-6).

При измерении трех координат (схема в Таблице 1, колонка 4, строка 2) на этапе 1 (первый период T_p) функционирует пара ЧЭ₁-Р и ЧЭ₂-К, а синхронное замыкание ключей K_1 и K_2 обеспечивает импульсное питание U_{II} мостовой схемы (Таблица 2, колонка 4, строка 2, эпюры 1-3). В результате последовательных преобразований в ГИЦ формируются цифровые коды $C_{1-1}, C_{1-2}, \dots, C_{1-n}$ (эпюры 5, 6, 8-10).

На этапе 2 функционирует та же пара ЧЭ, но их роли меняются на противоположные (ЧЭ₂-Р, ЧЭ₁-К). Благодаря задержке τ_3 импульсы управления появляются в момент прохождения лопатки 1 ЧЭ₂ (в тот же момент ЧЭ₁ удален от лопаток и функционирует как компенсационный). Синхронные замыкания ключей K_1, K_2 вызывают аналогичные последующие преобразования в ГИЦ, которые завершаются формированием кодов $C_{2-1}, C_{2-2}, \dots, C_{2-n}$ (эпюры 1-3, 5, 6, 8-10). При этом из-за асимметрии плеч моста значения напряжений U_M и U_{3Y} на этапе 2 отличаются от аналогичных значений на этапе 1.

На этапе 3 работает пара ЧЭ₁ и ЧЭ₃, причем компенсационные функции по-прежнему выполняет ЧЭ₁ (ЧЭ₁-К), а рабочим становится ЧЭ₃ (ЧЭ₃-Р). С задержкой τ_3 появляются импульсы управления в момент прохождения лопаткой 1 ЧЭ₃, а синхронные замыкания ключей K_1, K_3 вызывают соответствующие преобразования в ГИЦ, включая формирования кодов $C_{3-1}, C_{3-2}, \dots, C_{3-n}$ (эпюры 1, 2, 4, 5, 7-10)⁵.

На основе полученных кодов, а также ГХ составляется система уравнений (для каждой лопатки) и вычисляются искомые координаты смещений. Длительность получения информации составляет около трех периодов ($\approx 3T_p$).

На рисунке 1 представлены временные диаграммы работы тех же схем, но обеспечивающих уменьшение длительности получения информации о двух и трех координатах [9].

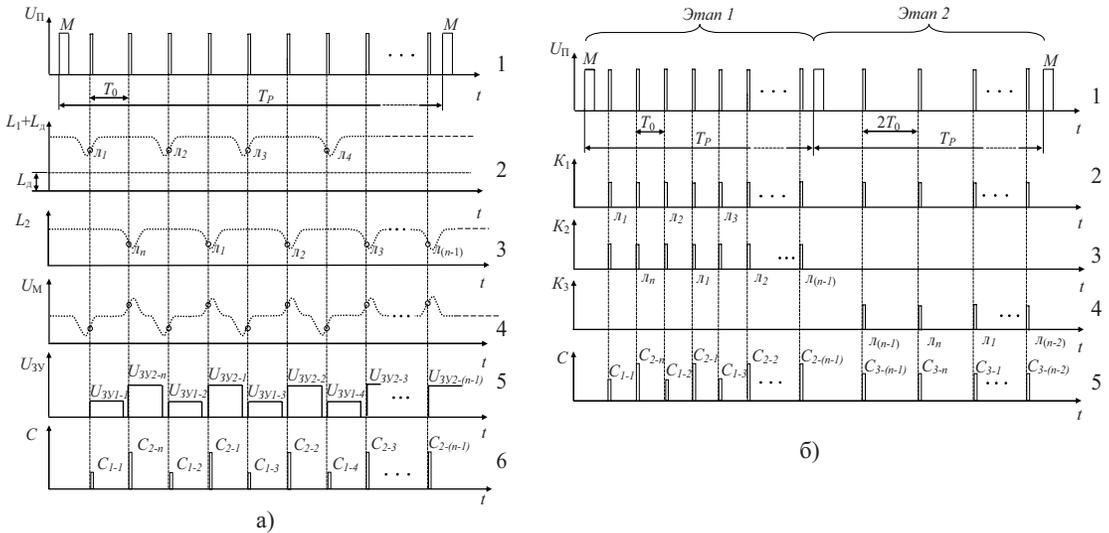


Рисунок 1 – Временные диаграммы схем, обеспечивающих уменьшение длительности получения информации при измерении двух (а) и трех (б) координат

Если измеряются две координаты (Таблица 1, колонка 4, строка 1), то по результатам измерения периода T_p вычисление периода T_0 производится по формуле $T_0 = T_p/2n$, причем период T_0 строго соответствует повороту лопаточного колеса на угол $0.5\Delta\Psi_{л}$ ($\Delta\Psi_{л}$ – угловой шаг

⁵ Эпюры $L_1(t), L_2(t), L_3(t)$ (Таблица 2, колонка 4, строка 2) соответствуют размещению ЧЭ типа «звезда» в СК со смещением центра и углом между ЧЭ₁ и ЧЭ₃, равным 90 град. [8].

установки лопаток) и смене функций каждого ЧЭ с рабочей на компенсационную (Рисунок 1, а, эпюра 1).

Все последующие преобразования сигналов обоих датчиков происходят в течение одного периода T_p (эпюры 2-6), т.е. обеспечивается вдвое меньшая длительность получения информации.

Если измеряются три координаты (схема в Таблице 1, колонка 4, строка 2), то в течение первого периода синхронное замыкание пары ключей K_1, K_2 обеспечивает подачу питания на датчики ОВТД₁, ОВТД₂ (ЧЭ₁, ЧЭ₂), а в течение второго — пары ключей K_1, K_3 и соответственно питание ($U_{п1}$) датчиков ОВТД₁, ОВТД₃ (ЧЭ₁, ЧЭ₃) (рисунок 1, б, эпюры 1-4).

Однако, в течение первого периода вращения лопаточного колеса подача $U_{п1}$, получение напряжений U_m, U_{3y} и соответствующих кодов C (эпюра 5) происходит с интервалом T_0 , а в течение второго периода вращения — с вдвое большим интервалом ($2T_0$).

Заключение

Предложена классификация существующих ГИЦ, ориентированных на ОВТД (КОВТД), которая определяется видом кластера (сосредоточенный и распределенный), а также числом ОВТД (чувствительных элементов КОВТД). Дано систематизированное описание структурных вариантов ГИЦ и их функционирования на уровне временных диаграмм.

Список литературы

- [1] Методы и средства измерения многомерных перемещений элементов конструкций силовых установок / Под ред. Секисова Ю.Н., Скобелева О.П. – Самара: Самарский НЦ РАН, 2001. - 188 с.
- [2] Скобелев О.П. Методы преобразования информации на основе тестовых переходных процессов// Измерение, контроль, автоматизация. 1978. № 4 (16). С. 11 – 17.
- [3] Беленький Л.Б., Райков Б.К., Секисов Ю.Н., Скобелев О.П. Вихретоковые экранные датчики механических параметров для систем автоматизации экспериментальных исследований и испытаний// Автометрия. - 1994. № 5. С. 111 – 116.
- [4] Скобелев О.П. Методы и средства группового преобразования сигналов однородных датчиков в информационно-измерительных системах. Дис. докт. техн. наук. – Куйбышев, 1987. – 565 с.
- [5] Беленький Л.Б., Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Метод получения измерительной информации о многокоординатных смещениях торцов лопастей винтовентилятора с одновременным преобразованием сигналов кластерных датчиков. Часть 1. Обоснование метода и его описание// Вестник СамГТУ, Серия «Технические науки». - 2009. № 1(23). С. 89 – 94.
- [6] Беленький Л.Б., Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Метод получения измерительной информации о многокоординатных смещениях торцов лопастей винтовентилятора с одновременным преобразованием сигналов кластерных датчиков. Часть 2. Реализуемость методов// Вестник СамГТУ, Серия «Технические науки». - 2009. № 2(24). С. 47-54.
- [7] Беленький Л.Б., Боровик С.Ю., Логвинов А.В., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Методы измерения смещений торцов лопаток в компрессорах и турбинах на основе распределенных кластеров датчиков. Часть 1. Обоснование предлагаемых методов и их описание//Мехатроника, автоматизация, управление. - 2009. № 4. С. 16-19.
- [8] Беленький Л.Б., Боровик С.Ю., Логвинов А.В., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Методы измерения смещений торцов лопаток в компрессорах и турбинах на основе распределенных кластеров датчиков. Часть 2. Реализуемость методов// Мехатроника, автоматизация, управление. - 2009. № 5. С. 21-30.
- [9] Беленький Л.Б., Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Уменьшение длительности получения информации о смещениях торцов лопаток //Мехатроника, автоматизация, управление. - 2010. № 5. (В печати).
- [10] Беленький Л.Б., Боровик С.Ю., Райков Б.К., Скобелев О.П. Анализ основных характеристик групповых измерительных цепей одновитковых вихретоковых датчиков// Настоящий сборник.

АНАЛИЗ ОСНОВНЫХ ХАРАКТЕРИСТИК ГРУППОВЫХ ИЗМЕРИТЕЛЬНЫХ ЦЕПЕЙ ОДНОВИТКОВЫХ ВИХРЕТОКОВЫХ ДАТЧИКОВ¹

Л.Б. Беленький, С.Ю. Боровик, О.П. Скобелев

Институт проблем управления сложными системами РАН

443020, Самара, ул. Садовая, 61, Россия

belenki@iccs.ru

тел: +7 (846) 333-26-77, факс: +7 (846) 333-27-70

Ключевые слова: групповая измерительная цепь, эквивалентная схема, одновитковый вихретоковый датчик, линии связи, ключи, дифференцирующее устройство, функции преобразования, влияние параметров

Abstract

The paper describes two equivalent circuits for measuring circuits with single-coil eddy-current sensors. The first equivalent circuit characterizes the most important processes in the group measuring circuits. During analysis take place the assumption about idealness of switches. The second equivalent circuit represents influence of resistance and capacitance of on and off switches. Results of analysis are given.

Введение

Известно, что основной характеристикой любого измерительного преобразователя является функция преобразования (ФП), выражающая связь выходного и входного параметра в статике.

Применительно к групповым измерительным цепям (ГИЦ) с одновитковыми вихретоковыми датчиками (ОВТД) [1] ФП можно представить в виде зависимости выходного напряжения (или цифрового кода) от изменений индуктивности чувствительного элемента (ЧЭ) датчика.

Поскольку ГИЦ построены на основе метода первой производной [2], то ФП в идеале определяется амплитудными значениями напряжений на выходе дифференцирующего устройства (ДУ) в момент подачи импульсного питания, а в реальности – спустя небольшое время, которое зависит от основных и паразитных параметров элементов ГИЦ – ОВТД, линий связи, ключей и операционных усилителей (ОУ) в режиме дифференцирования.

Исследования индивидуальных измерительных цепей с ОВТД, проведенные с учетом LRC-параметров согласующих трансформаторов (СТ) ОВТД, линий связи и RC-параметров ДУ [3] показали, что амплитудного значения выходное напряжение достигает за десятые доли мкс и это время частично определяет длительность преобразования².

В работе [4] исследовались характеристики ГИЦ, в которых в дополнение к перечисленным параметрам элементов цепи учитывалось влияние остаточных параметров бесконтактных ключей. Однако, влияние параметров входного контура цепи, включающего RLC-параметры ЧЭ, токовода датчика и объемного витка не рассматривалось³. Что же касается влияния остаточных параметров ключевых элементов в ГИЦ, то полученные оценки ограничивались цепя-

¹ Работа выполнена при частичной поддержке РФФИ (грант №08-08-00422а).

² Длительность преобразования зависит и от времени рассеяния энергии в измерительной цепи.

³ Исключение составляет работа [5], но проведенные в ней исследования ограничивались мостом Блумлейна, во входном контуре которого автор исследовал только влияние индуктивных параметров, оставив без внимания емкости тоководов и объемного витка СТ.

ми, включающими традиционные индуктивные датчики с магнитопроводами или так называемыми «экранными» (многовитковыми токовихревыми датчиками), индуктивность которых не менее, чем на порядок превышает индуктивность современных конструкций ОВТД и их кластерных разновидностей (КОВТД). При этом не учитывалось влияние межэлектродных емкостей в бесконтактных ключевых элементах, роль которых возрастает с уменьшением индуктивности датчика.

В работах [6, 7] для ГИЦ, ориентированных на ОВТД и КОВТД, исследовано влияние межэлектродных емкостей бесконтактных ключевых элементов, но полученные результаты требуют коррекции в связи с дальнейшим уменьшением индуктивностей СТ в усовершенствованных конструкциях ОВТД, а также остаточных параметров ключей, предлагаемых производителями в последнее время.

И, наконец, в работе [8] исследовалось не только влияние LRC-параметров СТ ОВТД и линий, параметров ОУ в ДУ, но LRC-параметров ЧЭ датчиков и тоководов. Полученные результаты можно использовать и в анализе основных характеристик ГИЦ, полагая, что применяемые ключевые элементы идеальны, но при этом необходимо дополнительно исследовать влияние емкостей линий связи в цепях «неопрашиваемых» датчиков, в которых ключевые элементы разомкнуты.

Настоящая статья, посвященная анализу основных характеристик ГИЦ, ориентированных на ОВТД (КОВТД), призвана устранить перечисленные пробелы. В качестве инструмента для анализа любого из вариантов ГИЦ, представленных в работе [1], предлагаются модели, построенные на основе эквивалентных схем, описание которых разбросаны по отдельным источникам [6-8]. Рассматриваются частично доработанные эквивалентные схемы и соответствующие модели, причем для сравнения и сопоставления результатов моделирования исследования выполняются при одних и тех же исходных данных по элементам ГИЦ, в том числе для последних разработок применяемой электронной элементной базы.

1 Эквивалентные схемы измерительных цепей

Несмотря на разнообразие существующих ГИЦ [1], для анализа основных характеристик достаточно всего двух эквивалентных схем и соответствующих моделей. Первая эквивалентная схема отражает наиболее значимые процессы в ГИЦ при условии, что ключевые элементы идеальны (сопротивления в замкнутом и разомкнутом состоянии равны нулю и бесконечности соответственно, а межэлектродные емкости не учитываются ввиду их малости). Вторая – процессы в ГИЦ, которые связаны с неидеальностью ключевых элементов (ненулевые и конечные сопротивления в замкнутом и разомкнутом состоянии, ненулевые межэлектродные емкости). При этом предполагается, что линии связи отсутствуют, а в отношении параметров датчиков принято упрощающее допущение - учитываются только эквивалентные R , L , C -параметры СТ ОВТД.

Первая эквивалентная схема изображена на рисунке 1,а. Чтобы показать ее отличительные особенности по сравнению со схемой, приведенной в работе [8], на рисунке 1,б в упрощенном виде представлена ГИЦ с датчиками и линиями связи (ЛС), выполненными проводами с заземленными экранами, источником питания (E), а также ключами (K), один из которых в замкнутом, а остальные в разомкнутом состоянии. Здесь изображены только одно плечо неравновесного моста [1] с рабочим («опрашиваемым») ОВТД, включенным последовательно с замкнутым ключом, а также прямая и обратная линия связи (ЛС). При этом ЛС представлены Г-образными схемами замещения, которые характеризуются сосредоточенными параметрами - сопротивлением $R_{л}$, индуктивностью $L_{л}$ и емкостью $C_{л}$. Следует отметить, что элементы ГИЦ, (в том числе K , ДУ и запоминающее устройство [1]), как правило, «встроены» в ЛС и находятся на небольшом расстоянии от датчиков (от 0.5 до 1 м).

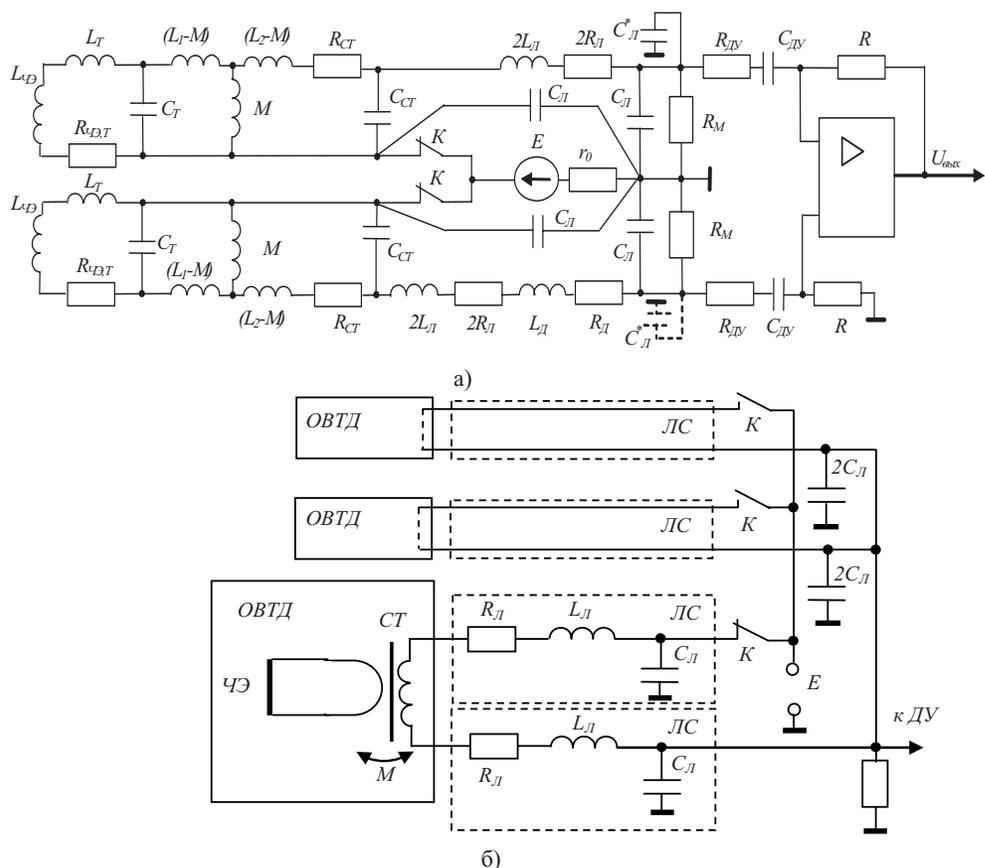


Рисунок 1 – Первая эквивалентная схема (а) и упрощенная схема ГИЦ с одним плечом неравновесного моста, ключами и линиями связи (б)

Поскольку в существующих ГИЦ [1] число ОВТД в кластере (ЧЭ в КОВТД) не превышает трех, то число разомкнутых ключей всего два, а отрезков ЛС, включенных последовательно с разомкнутыми ключами и “неопрашиваемыми” ОВТД (ЧЭ), соответственно четыре. При этом на переходные процессы в мостовой схеме (в плече моста с замкнутым ключом и “опрашиваемым” ОВТД) оказывают влияние только емкости ЛС C_L в цепях разомкнутых ключей и “неопрашиваемых” ОВТД (рисунок 1,б).

Наличие емкости C_L^* на одном из входов ОУ (ДУ) и является основной отличительной особенностью эквивалентной схемы на рисунке 1,а. Ее величина зависит от исследуемого варианта ГИЦ и варьируется в пределах от 0 до $4C_L$ [1].

Кроме того, в рассматриваемой эквивалентной схеме используются следующие обозначения: $L_{ЧЭ}$ – индуктивность ЧЭ (предполагается, что в рабочем ЧЭ индуктивность $L_{ЧЭ} = L - \Delta L$, а в компенсационном $\Delta L = 0$ и его индуктивность постоянна и равна L ; L_T, C_T – индуктивность и емкость тоководов; $R_{ЧЭ,Т}$ – сопротивление входных контуров ОВТД, равное сумме сопротивлений ЧЭ и токоиода; L_1 и L_2 – индуктивности одновитковых первичных и многovitковых вторичных обмоток СТ; M – коэффициент взаимной индукции; L_D, R_D – сопротивление и индуктивность добавочной катушки, обеспечивающей однополярный выходной сигнал на выходе моста; R_M – сопротивления плеч неравновесного моста; R_{CT} – сопротивления вторичных обмоток СТ; $R_{ДВ}, R, C_{ДВ}$ – параметры RC – цепочек, обеспечивающие работу ОУ в режиме ДУ; r_0 – внутреннее сопротивление источника питания.

Вторая эквивалентная схема, предназначенная для исследований влияния остаточных параметров ключей, изображена на рисунке 2. Здесь ОВТД (КОВТД) представлены эквивалентными параметрами вторичной обмотки СТ – индуктивностью $L_{СТ} = (w_2/w_1)^2 \cdot (L_{чЭ} + L_T)$, где w_1 , w_2 – число витков первичной и вторичной обмотки СТ ($w_1=1$), и ее сопротивлением $R_{СТ}$. При этом схема на рисунке 2 отражает максимально возможное число разомкнутых ключей в рабочем плече моста рассматриваемых ГИЦ [1]. Следует также отметить, что важным условием

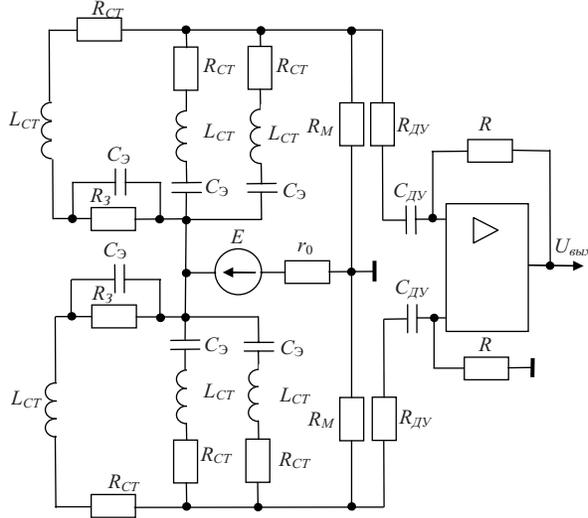


Рисунок 2 – Эквивалентная схема для исследования влияния остаточных параметров

снижения влияния паразитных параметров цепей, работающих в импульсном режиме по методу первой производной, является симметрирование RL -параметров плеч мостовой схемы. Поэтому с целью симметрирующей коррекции параллельно компенсирующему плечу моста через всегда разомкнутые ключи вводятся катушки с параметрами $L_{СТ}$, $R_{СТ}$, соответствующими эквивалентным параметрам датчика, не взаимодействующего с лопаткой (лопастью).

Остаточные параметры ключей в замкнутом состоянии определяются сопротивлением R_3 и эквивалентной емкостью C_3 , в разомкнутом – только емкостью C_3 (предполагается, что сопротивление в разомкнутом состоянии $R_p \rightarrow \infty$ и его влиянием можно пренебречь). Эквивалентная емкость определяется емкостью между стоком и истоком ($C_{СИ}$) и параллельной ей цепочкой из последовательно соединенных емкостей между стоком и затвором ($C_{3С}$) и истоком и затвором ($C_{3И}$) [6].

2 Моделирование и полученные результаты

Математические модели эквивалентных схем, изображенных на рисунках 1,а и 2, представляют собой системы дифференциальных уравнений высокого порядка, для решения которых приходится использовать численные методы. Вместе с тем, если предположить, что в первой эквивалентной схеме линии связи отсутствуют, остальные емкостные параметры эквивалентной схемы пренебрежимо малы, а частота единичного усиления ОУ не ограничена ($f_i \rightarrow \infty$), то можно получить решение в аналитическом виде (в виде формулы). Дальнейшие упрощения (допущение $R_{ДУ} = 0$) позволяет получить искомое решение – функцию $U_{ДУ}(t)$ в наиболее наглядном и компактном виде: график этой функции представлены на рисунке 3⁴ (1), а

⁴ В расчетах $U_{ДУ}(t)$ здесь и далее использовались следующие значения параметров $E = 7$ В, $r_0 = 0.1$ Ом, $L_{чЭ} = L = 5 \cdot 10^{-9}$ Гн, $\Delta L = 1.25 \cdot 10^{-10}$ Гн, $R_{чЭ,Т} = 6.25 \cdot 10^{-3}$ Ом, $L_T = 1.5 \cdot 10^{-9}$ Гн, $w_2 = 40$, $R_{СТ} = 0.2$ Ом, $R_D = 0.1$ Ом, $L_D = 0.5 \cdot 10^{-6}$ Гн, $R = 33$ Ом, $R = 3000$ Ом, $C_{ДУ} = 680 \cdot 10^{-12}$ Ф, $R_{ДУ} = 68$ Ом.

ее максимальное значение (ему соответствует момент времени $t = 0$) в предположении, что $L_{\text{чЭ}} \gg L_{\text{T}}$ определяется выражением:

$$U_{\text{ДУм}}|_{t \rightarrow 0} = E \cdot \frac{T}{\tau_{\text{ЭКВ}}} \cdot \frac{\Delta L}{L_{\text{чЭ}}} \cdot \frac{1}{1 - (\Delta L / L_{\text{чЭ}})}, \text{ где } T = R \cdot C_{\text{ДУ}}, \tau_{\text{ЭКВ}} = \frac{(w_2 / w_1)^2 \cdot L_{\text{чЭ}}}{R_{\text{M}}}$$

w_1, w_2 — число витков первичной и вторичной обмоток СТ.

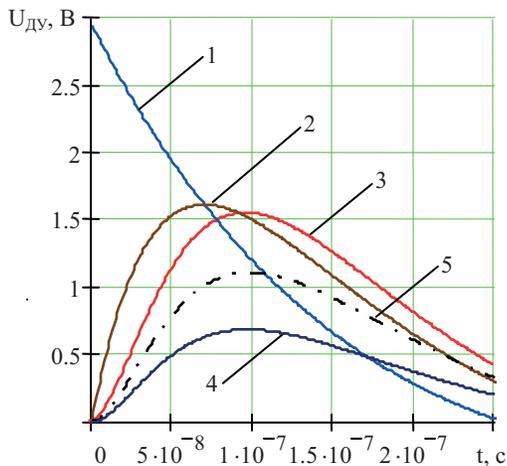


Рисунок 3 – Изменение выходного напряжения ДУ

Как следует из графика 2 ($f_i \rightarrow \infty$), амплитудное значение $U_{\text{ДУм}}$ уменьшается и смещается вправо по оси времени ($t_m \approx 0.07$ мкс). Графики 3, 4, 5 получены при конечных значениях f_i^5 и $R_{\text{ДУ}} = 0$. При этом значения $U_{\text{ДУ}}(t)$ на графиках 1-3 рассчитывались в предположении, что рабочим в мостовой схеме является плечо, в котором отсутствует добавочная катушка индуктивности ($L_{\text{д}}, R_{\text{д}}$) (рисунок 1,а). График 4 показывает изменения $U_{\text{ДУ}}(t)$, когда рабочим является другое плечо (с добавочной катушкой), а график 5 (пунктирная линия) соответствует $\Delta L = 0$.

Отмена ранее принятых допущений в отношении параметров эквивалентной схемы (рисунок 1,а) приводит к необходимости использования численного моделирования, которое осуществлялось в среде Mathcad. В частности, учет емкости токовода⁶ приводит к появлению высокочастотной составляющей в выходном напряжении ДУ, которая растет с ростом f_i (рисунок 4, графики 1 (300 МГц) и 2 (1500 МГц))⁷. Увеличение f_i ожидаемо приводит к более раннему наступлению максимума $U_{\text{ДУ}}$.

На рисунке 4 представлены также результаты моделирования ГИЦ с распределенным кластером из трех ОВТД, в котором используются ЛС (график 3 и 4)⁸. Моделирование велось с учетом симметрирующей емкости $C_{\text{л}}^*$, которая обозначена пунктиром на рисунке 1,а. При этом уменьшение амплитуды $U_{\text{ДУм}}$ не превышает 2.5%, а время наступления максимума увеличивается приблизительно на 5 нс.

Для той же схемы ГИЦ определялись ФП, которые представляют собой зависимость $U_{\text{ДУм}}$ от отношения $\Delta L / L$, где $L = L_{\text{чЭ}} + L_{\text{T}}$ (рисунок 5). График 1 получен в предположении, что роль рабочего выполняет датчик, включенный в плечо, где отсутствует добавочная индуктив-

В существующих датчиках полезные изменения выходного параметра незначительны и обычно составляют единицы процентов. Поэтому максимум выходного напряжения ДУ $U_{\text{ДУм}}|_{t \rightarrow 0} \approx E \cdot \frac{T}{\tau_{\text{ЭКВ}}} \cdot \frac{\Delta L}{L_{\text{чЭ}}}$ и не

зависит от изменений сопротивлений ЧЭ и тоководов датчика ($R_{\text{чЭ,Т}}$), связанных с изменениями температуры окружающей среды, которые составляют сотни процентов.

Решение системы дифференциальных уравнений в аналитическом виде можно получить и при $R_{\text{ДУ}} \neq 0$, сохранив допущение $f_i \rightarrow \infty$, и даже при конечных значениях f_i . Однако, полученные формулы громоздки и поэтому здесь не приводятся. В то же время результаты расчета $U_{\text{ДУ}}(t)$ по этим формулам представлены на рисунке 3.

⁵ Предполагается, что в ДУ используют операционный усилитель с $f_i = 300$ МГц.

⁶ Емкость токовода $C_{\text{T}} = 100 \cdot 10^{-12}$ Ф.

⁷ Предполагается, что в ДУ используется операционный усилитель LMH6552 и AD8056.

⁸ Погонные параметры линии $R_{\text{л}} = 0.1$ Ом/м, $L_{\text{л}} = 0.1 \cdot 10^{-6}$ Гн/м, $C_{\text{л}} = 100 \cdot 10^{-12}$ Ф/м. Длина линии 0.5 м.

ность, а график 2 – когда роль рабочего выполняет датчик, включенный в плечо с добавочной индуктивностью. В пределах $\Delta L/L$ от 0 до 0.025 эта нелинейность не превышает 0.3%.

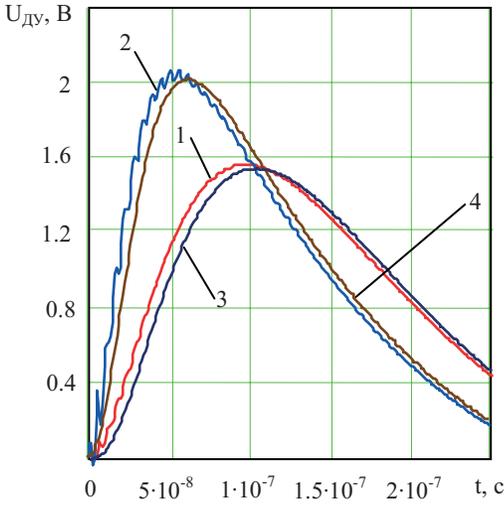


Рисунок 4 – Результаты моделирования с учетом емкостных параметров датчиков и R_3, L_3 и C_3

Модель эквивалентной схемы, изображенной на рисунке 2, также представляет собой систему дифференциальных уравнений, для решения которой используются численные методы и среда Mathcad.

В ходе вычислительных экспериментов изучалось влияние остаточных параметров ключей на форму $U_{дв}(t)$, его амплитудное значение, а также его смещение во времени от момента $t=0$. При этом предполагалось синхронное срабатывание ключей в плечах моста, идеальность ДУ ($R_{дв}=0, f_i=\infty$) и исследовалось влияние конечных значений f_i реально существующих ОУ. Кроме того, оценивалось влияние неустойчивости остаточных параметров ключей из-за изменения температуры и влияние несинхронного срабатывания ключей из-за разброса длительности задержки в выбранных парах.

Как показывает анализ результатов моделирования в варианте идеальных ключей (синхронное срабатывание, $R_3=0, C_3=0$) и идеальности ДУ (рисунок 6) максимум $U_{дв}(t)$ наблюдается при $t=0$ (1). Емкости реальных ключей вносят заметную колебательность (2) при сохранении максимума $U_{дв}$ в момент времени $t=0$.

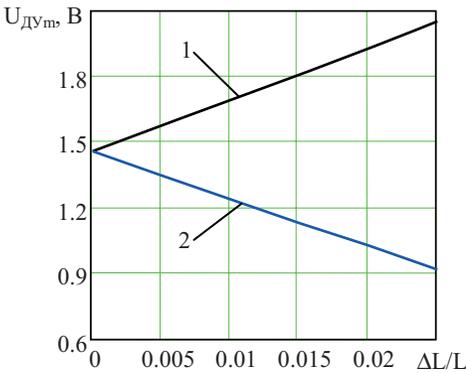


Рисунок 5 – Функции преобразования

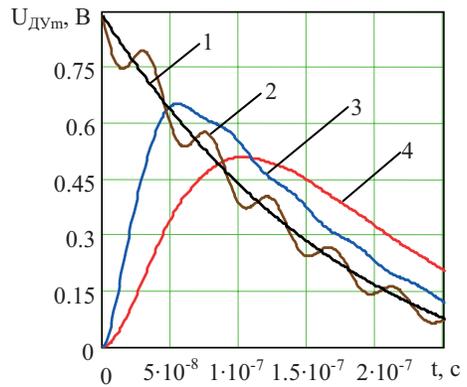


Рисунок 6 – Результаты моделирования эквивалентной схемы, учитывающей остаточные параметры ключей

Неидеальность ДУ и, в частности, уменьшение частоты f_i приводит к снижению $U_{двм}$ и увеличению его смещения во времени: графики $U_{дв}(t)$ получены при $f_i=1500$ МГц (3) и $f_i=300$ МГц (4) и остаточных параметрах реально существующих ключей⁹). Изменения остаточных параметров ключей невелики (например, изменения R_3 составляет около 20% при изменении

⁹ При моделировании для ключа FDC3601N использовались максимальное значение $R_3=0.5$ Ом и $C_3=5 \cdot 10^{-12}$ Ф.

температуры на 25°C) и им соответствуют весьма незначительные изменения $U_{\text{двм}}$ (на 0.1% для ОУ с $f_{\text{г}}=300$ МГц и на 0.06% для ОУ с $f_{\text{г}}=1500$ МГц).

В ходе исследований несинхронного срабатывания ключей из-за разброса в задержках выявлено резкое усиление колебательности $U_{\text{дв}}(t)$ [7], которое, как показали натурные эксперименты, устраняется с помощью RC -цепочек в цепи управления ключами.

Заключение

Представленные эквивалентные схемы и построенные на их основе модели обеспечивают исследования основных характеристик всего существующего разнообразия ГИЦ, ориентированных на ОВТД и КОВТД. В результате вычислительных экспериментов для исходных данных по наиболее современным датчикам и последним разработкам электронной элементной базы получены, в частности, функции преобразования и оценки влияния ряда неинформативных параметров, от которых зависит точность преобразования.

Список литературы

- [1] Беленький Л.Б., Боровик С.Ю., Райков Б.К., Скобелев О.П. Систематизация и классификация групповых измерительных цепей одновитковых вихретоковых датчиков // Настоящий сборник.
- [2] Методы и средства измерения многомерных перемещений элементов конструкций силовых установок / Под ред. Секисова Ю.Н., Скобелева О.П. – Самара: Самарский НЦ РАН, 2001. – 188 с.
- [3] Беленький Л.Б. Системы измерения многомерных перемещений и деформаций элементов конструкций кривошипно-шатунного механизма в поршневых силовых установках. Дисс. канд. техн. наук. – Самара, 2003. – 198 с.
- [4] Скобелев О.П. Методы и средства группового преобразования сигналов однородных датчиков в информационно-измерительных системах. Дисс. докт. техн. наук. – Куйбышев, 1987. – 565 с.
- [5] Хритин А.А. Система измерения радиальных зазоров в турбомашинах. Дисс. канд. техн. наук. – Самара, 1994. – 175 с.
- [6] Беленький Л.Б., Боровик С.Ю., Логвинов А.В., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Методы измерения смещений торцов лопаток в компрессорах и турбинах на основе распределенных кластеров датчиков. Часть 2. Реализуемость методов // Мехатроника, автоматизация, управление. - 2009. № 5. С. 21-30.
- [7] Беленький Л.Б., Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Метод получения измерительной информации о многокоординатных смещениях торцов лопастей винто-вентилятора с одновременным преобразованием сигналов кластерных датчиков. Часть 2. Реализуемость методов // Вестник СамГТУ, Серия «Технические науки». - 2009. № 2(24). С. 47-54.
- [8] Беленький Л.Б., Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Уменьшение длительности получения информации о смещениях торцов лопаток // Мехатроника, автоматизация, управление. - 2010. № 5. (В печати).

ОБ АДЕКВАТНОСТИ АНАЛИТИЧЕСКОЙ МОДЕЛИ ЭЛЕКТРОМАГНИТНОГО ВЗАИМОДЕЙСТВИЯ ЧУВСТВИТЕЛЬНОГО ЭЛЕМЕНТА ОДНОВИТКОВОГО ВИХРЕТОКОВОГО ДАТЧИКА С ТОРЦОМ ЛОПАТКИ И ЕЁ СОВЕРШЕНСТВОВАНИЕ¹

В.В. Тулупова

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
tulupova@iccs.ru
тел: +7 (846) 333-26-77, факс: +7 (846) 333-27-70

Ключевые слова: *одновитковый вихретоковый датчик, функция преобразования, модель электромагнитного взаимодействия, применение модели, адекватность модели*

Abstract

The criteria of conformity for analytical model of electromagnetic interaction of sensing element of the single-coil edge-current sensor with blade tip are given. The method for this model upgrading is considered.

Введение

Известны аналитические модели электромагнитного взаимодействия чувствительных элементов (ЧЭ) одновитковых вихретоковых датчиков (ОВТД) в составе кластера с торцом лопатки (лопасти) газотурбинного двигателя [1]. В процессе построения такой модели ЧЭ и объект (лопатка или лопасть) заменяются электропроводными контурами с допущениями, что параметры электромагнитного поля в пределах каждого контура не зависят от координат, а проводники, образующие контур, имеют ничтожно малый диаметр. По закону Био-Савара определяются величина индукции магнитного поля, которое создает каждая из сторон контуров в окружающем пространстве, и магнитные потоки, связанные с контурами. Далее на основе второго закона Кирхгофа для всех контуров, как электрических цепей с сосредоточенными параметрами, составляется система уравнений, позволяющая (в рамках дополнительных упрощающих допущений) найти семейство функций преобразования (ФП) в виде аналитических формул, связывающих индуктивности ЧЭ с координатами смещений торца.

В настоящее время аналитические модели широко используются на ранних стадиях проектирования датчиков, что иллюстрируется рядом публикаций. В работе [1] рассматривается применение аналитической модели для оценки взаимного влияния в составе сосредоточенного кластера ОВТД (кластерного ОВТД) в процессе одновременного преобразования координат смещений торца в изменения выходных параметров ЧЭ. Примеры использования моделей для проверки и обеспечения реализуемости методов одновременного преобразования координат в сосредоточенном и распределённом кластерах ОВТД за счёт выбора топологии размещения ЧЭ и смещения геометрического центра кластера (г.ц.) даны в работах [2, 3].

Следует заметить, что при практическом использовании аналитических моделей ещё не проводились исследования их адекватности. В настоящей статье формулируются критерии адекватности аналитических моделей для использования их на ранних стадиях проектирования, выполняется проверка ранее разработанных моделей и рассматривается способ совершенствования для лучшего соответствия этим критериям.

¹ Работа выполнена при частичной поддержке РФФИ (грант № 08-08-00422а).

1 Об адекватности аналитических моделей

Представляется очевидным, что адекватность рассматриваемых моделей должна оцениваться путем сравнения расчетных и экспериментальных ФП. Однако, при сравнении возможно использование различных критериев, выбор которых зависит от задач, решаемых с помощью моделирования.

Например, на этапе конструирования ОВТД (КОВТД) в процессе поиска проектного решения важно получить достоверные расчетные данные об эквивалентной индуктивности ОВТД (L_p)², зависящей от множества геометрических и физических параметров элементов конструкции. Её величина должна быть максимально приближена к индуктивности, найденной в результате эксперимента (L_s). В качестве критерия в этом случае уместно использование разности расчетных и экспериментальных значений индуктивностей ($\Delta L=L_p-L_s$), полученных для нулевых и «бесконечных» радиальных зазоров (РЗ).

Вместе с тем, как было показано в [2, 3], наиболее сложные вопросы реализуемости методов преобразования координат в кластерах ОВТД связаны с разрешимостью и сходимостью решений систем уравнений, составленных на основе ФП и выходных параметров датчиков в составе кластера. Решение тех же вопросов в системах измерения, реализующих рассматриваемые методы, сводится к замене ФП на градуировочные характеристики (ГХ), и выходных параметров датчиков на соответствующие значения кодов в измерительных каналах системы. Необходимо подчеркнуть, что реализуемость методов в этом случае в основном определяется видом ГХ, т.е. характером функциональной зависимости кодов от искомых координат в заданном диапазоне, который, в свою очередь, практически полностью определяется относительными изменениями³. Поэтому об адекватности модели можно судить по разности расчетных и экспериментальных значений относительных индуктивностей (\bar{L}_p и \bar{L}_s), вычисляемых по

формуле $\bar{L} = \frac{L - L_{\min}}{L_{\max} - L_{\min}}$, где L_{\max} и L_{\min} – границы диапазона изменения индуктивности.

Поскольку сравниваются монотонные функции, то из множества значений в качестве критерия выбирается максимальное отклонение ($\Delta \bar{L}_{\max} = \bar{L}_p - \bar{L}_s$).

В экспериментальных исследованиях объект измерения представлен реальной лопаткой с длиной торцевой части 30 мм, высотой 50 мм и максимальной толщиной 1.7 мм, а также имитатором лопатки в виде клинообразной пластины максимальной толщиной 3.75 мм (лопатка выполнена из нержавеющей стали, имитатор – из алюминия). Длина ЧЭ ОВТД – 9 мм. Измерения индуктивности датчика проводились с помощью измерителя индуктивности Е7-12. Перемещения объекта задавались на специальной градуировочной установке, обеспечивающей перемещение объекта в направлении оси Y , и контролировались индикатором промышленного изготовления с разрешающей способностью 10 мкм. Г.ц. верхних торцов объекта расположены на оси Y , т.е. проекции г.ц. на плоскость OYZ совпадают с началом системы отсчета.

Экспериментальные ФП ($\bar{L}_s(y)$) определялись в диапазоне изменений y от 0.4 до 3.2 мм (рисунок 1). При этом экспериментальные значения индуктивности L_s в абсолютных единицах на верхней границе диапазона ($y_{\max}=3.2$ мм) одинаковы для лопатки и имитатора ($L_{s\max}=13.010$ мкГн), а на нижней границе ($y_{\min}=0.4$ мм) $L_{s\min}$ отличаются и составляют 12.634 мкГн (лопатка) и 12.587 мкГн (имитатор).

Следует отметить, что в относительных единицах характер изменения ФП, несмотря на различия в форме и материалах обоих объектов, идентичен и, более того, обе функции $\bar{L}_s(y)$ практически совпадают.

² $L_p = w^2 L_{\text{ЧЭ}}$, где w – коэффициент трансформации.

³ С учетом нормирования сигналов в измерительных каналах системы.

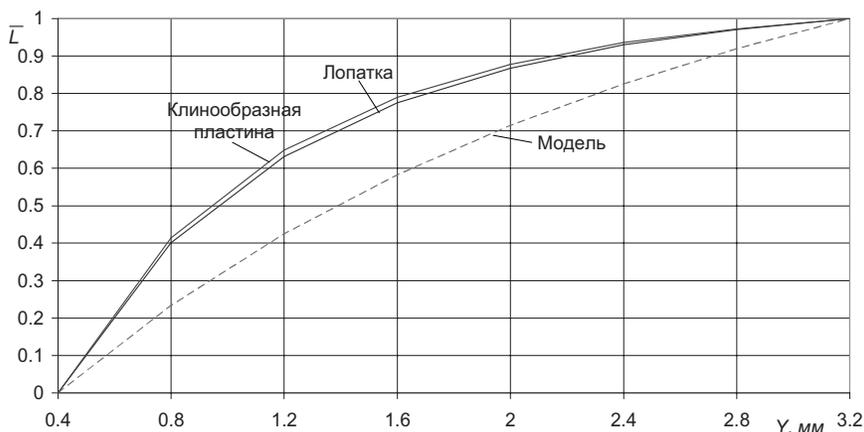


Рисунок 1 – Экспериментальные и расчетные ФП

На том же рисунке представлены ФП, рассчитанные с помощью модели ($\bar{L}_p(y)$).

Сравнение расчетных и экспериментальных ФП показывает в целом идентичный характер изменений: в исследуемом диапазоне y отсутствуют экстремумы, ФП монотонно возрастают, и, наконец, результаты сравнения не превышают 20 % ($\Delta\bar{L}_{\max} \approx 20\%$).

В то же время наблюдается существенная разница расчетных значений индуктивности на границах диапазона изменений y , выраженных в абсолютных единицах: $L_{p\min}=3.832$ мкГн и $L_{p\max}=3.852$ мкГн.

Возвращаясь к сравнению расчетных и экспериментальных ФП по критерию $\Delta\bar{L}_{\max}$ и анализируя вероятные причины их несовпадения, необходимо в первую очередь подчеркнуть существенное упрощение расчета индукции магнитного поля и соответственно магнитного потока контура, вычисляемой в центре площади контура⁴, которое предусмотрено в рассматриваемой модели [1].

2 Совершенствование модели

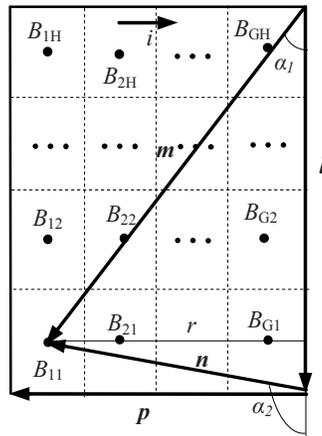
Повысить точность вычисления индукции можно за счет интегрирования по площади контура, в простейшем случае – за счет разбиения площади контура на элементарные площадки и определения индукции в центре каждой из них [1]. Эта операция не меняет существа подхода к моделированию, но немного усложняет вид математического выражения коэффициентов систем дифференциальных уравнений.

Разбиение площади контуров на элементарные площадки иллюстрирует рисунок 2. Коэффициенты систем M_{ik} , $i=1..(N_{ч3}+1)$, $k=1..(N_{ч3}+1)$ соответственно вычисляются по формуле:

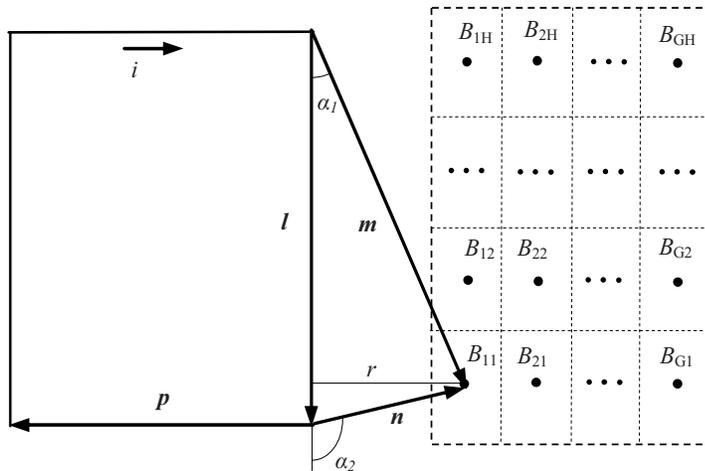
$$(1) \quad M_{ik} = \frac{\mu}{4\pi} \left(S_i \cdot \sum_{g=1}^G \sum_{h=1}^H \sum_{j=1}^4 \frac{l_{kj} \times m_{ghikj}}{|l_{kj} \times m_{ghikj}|} \cdot \frac{l_{kj} \cdot m_{ghikj} - l_{kj} \cdot n_{ghikj}}{l_{kj} \cdot m_{ghikj} \cdot l_{kj} \cdot n_{ghikj}} \right) \cdot m_{ghikj} \left(\frac{m_{ghikj} \cdot p_{kj}}{m_{ghikj} \cdot p_{kj}} \right)$$

⁴ Хотя известно, что зависимость индукции прямолинейного проводника с током, вычисляемой в произвольной точке пространства вблизи проводника, от координат этой точки имеет логарифмический характер.

где S_i - вектор-нормаль элементарной площадке i -го контура, вычисляемый через векторное произведение ($S_i = \frac{l_{i1}}{H} \times \frac{P_{i1}}{G}$), B_{ghik} - индукция в центре этой площадки от токов i_k , B_{ghik}^o - единичный вектор направления индукции от j -го проводника k -го контура к центру выбранной (g, h) площадки i -го контура, вычисляемый через нормализацию векторного произведения вектора направления j -го проводника k -го контура l_{kj} на вектор-радиус от начала j -го проводника k -го контура до центра выбранной площадки i -го контура m_{ghik} ($B_{ghik}^o = \frac{l_{kj} \times m_{ghik}}{|l_{kj} \times m_{ghik}|}$), l_{kj} , m_{ghik} , n_{ghik} , p_{kj} - вспомогательные вектора для вычисления модуля индукции для j -го проводника k -го контура и центра выбранной площадки i -го контура [1].



а)



б)

Рисунок 2 – Определение индукции магнитного поля для прямолинейного проводника с током конечной длины в центре элементарной площадки контура с током (а), в центре элементарной площадки соседнего контура с током (б)

На рисунке 3 представлены экспериментальная ФП (объект – лопатка), а также расчетная ФП, полученная на модели с разбиением контуров на элементарные площадки для вычисления индукции ($G=H=15$).

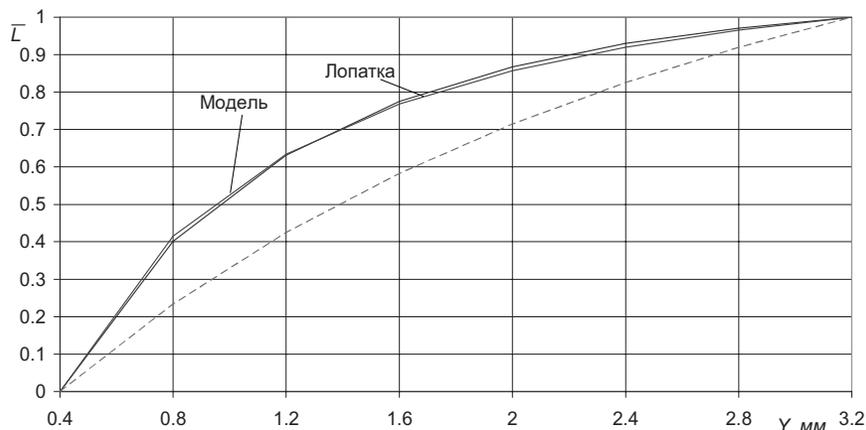


Рисунок 3 – Экспериментальные и расчетные ФП (пунктир – ФП, рассчитанная без разбиения)

Как следует из графиков функции \bar{L}_p и \bar{L}_s практически совпадают и критерий $\overline{\Delta L}_{\max}$ уменьшается почти на порядок ($\overline{\Delta L}_{\max} \approx 2\div3\%$) по сравнению с тем же критерием, полученным на моделях без разбиения. При этом расчетные значения индуктивностей на границах диапазона изменений y , выраженные в абсолютных единицах, составляют $L_{p\min}=10.117$ мкГн и $L_{p\max}=10.270$ мкГн, что «вплотную» приближает их к значениям индуктивностей, найденным экспериментально.

Заклучение

Сформулированы критерии адекватности аналитических моделей электромагнитного взаимодействия чувствительных элементов одновитковых вихретоковых датчиков с торцом лопатки с учетом их практического использования для проверки и обеспечения реализуемости методов преобразования координат в кластерах датчиков. Показан путь совершенствования моделей для обеспечения лучшего соответствия этим критериям.

Список литературы

- [1] Райков Б.К., Секисов Ю.Н., Тулупова В.В. Функции преобразования кластерного одновиткового вихретокового датчика // Автометрия. -2008, №1. – С. 47-58.
- [2] Беленький Л.Б., Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Метод получения измерительной информации о многокоординатных смещениях торцов лопастей винтовентилятора с одновременным преобразованием сигналов кластерных датчиков. Часть 2. Реализуемость методов // Вестник СамГТУ, Серия «Технические науки». - 2009. № 2(24). С. 47-54.
- [3] Беленький Л.Б., Боровик С.Ю., Логвинов А.В., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Методы измерения смещений торцов лопаток в компрессорах и турбинах на основе распределенных кластеров датчиков. Часть 2. Реализуемость методов // Мехатроника, автоматизация, управление. - 2009. № 5. С. 21-30.

РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ КАНАЛОВ АЭРОМЕТРИЧЕСКОЙ СИСТЕМЫ ИЗМЕРЕНИЯ ВОЗДУШНЫХ СИГНАЛОВ ВЕРТОЛЕТА

А.Н. Галяутдинова, Ю.С. Масленникова, А.А. Порунов

Казанский государственный технический университет им. А. Н. Туполева
420111, Казань, ул. К.Маркса, 10, Россия
Adelina603@yandex.ru
тел: +7 (843) 231-03-90

Ключевые слова: винтокрылые летательные аппараты, системы воздушных сигналов, моделирование, измерительные каналы, анализ, синтез, точностные показатели

Abstract

In this article results of simulation modeling channel aerometric air signal system of the helicopter are presented. Most important of them is the estimation of an aerodynamic field indignation and disorder of major converters parameters influence on factors of the basic informative channels quality. Also recommendations and requirements are formulated for constructive and scheme solutions of SVS-V blocks.

Введение

Повышение тактико-технических характеристик современных винтокрылых летательных аппаратов (ВЛА) требует непрерывного совершенствования бортовых средств, позволяющих получать информацию о пилотажных параметрах. Актуальность этой проблемы обусловлена тем, что до настоящего времени большая часть бортовых средств ВЛА заимствована из штатного состава самолетного оборудования, которая не учитывает специфику аэродинамики и режимов полета в ЛА. Поэтому использование этих методов и средств измерения не позволяет измерять воздушную скорость полета вертолета в диапазоне до 70-90 км/ч, а также углы атаки α и скольжения β в диапазоне от $\pm(30-60)$ угол.град. Это ограничивает возможность применения вертолетов на его наиболее важных режимах полета на околонулевых скоростях при изменении угла скольжения в диапазоне ± 180 угол. град [1].

В работе представлены результаты имитационного моделирования перспективного варианта СВС-В с ориентируемым аэрометрическим преобразователем на основе использования компьютерных технологий, которые помогают проводить разработку сложных систем различного назначения в сжатые сроки и при минимальных затратах [2, 3, 4].

Основная часть

Исследование процессов измерительного преобразования, проходящих в каналах СВС-В, является теоретической основой проектирования этих систем с заданными метрологическими характеристиками. Классический метод решения этой задачи основан на решении интегральных уравнений свертки функции. Этому методу присущи существенные недостатки: невозможность аналитического вычисления интеграла для некоторых математических выражений и трудоемкость решения задачи для систем с несколькими устройствами. В операторном методе операция свертки двух функций заменяется произведением их изображений. Основным недостатком этого метода – невозможность аналитического определения полюсов дробной рациональной функции изображения при высоких степенях полиномов функции.

В настоящее время для анализа авиационных приборов и систем широко применяются компьютерные программы [5], реализующие численные методы обратного преобразования

Лапласа. Однако алгоритмы, реализующие эти методы, являются довольно громоздкими. Поэтому представленные в работе результаты исследования статических и динамических характеристик каналов СВС-В (рисунок 1) получены численными методами построения переходных процессов в частности на основе метода Рунге-Кутты.

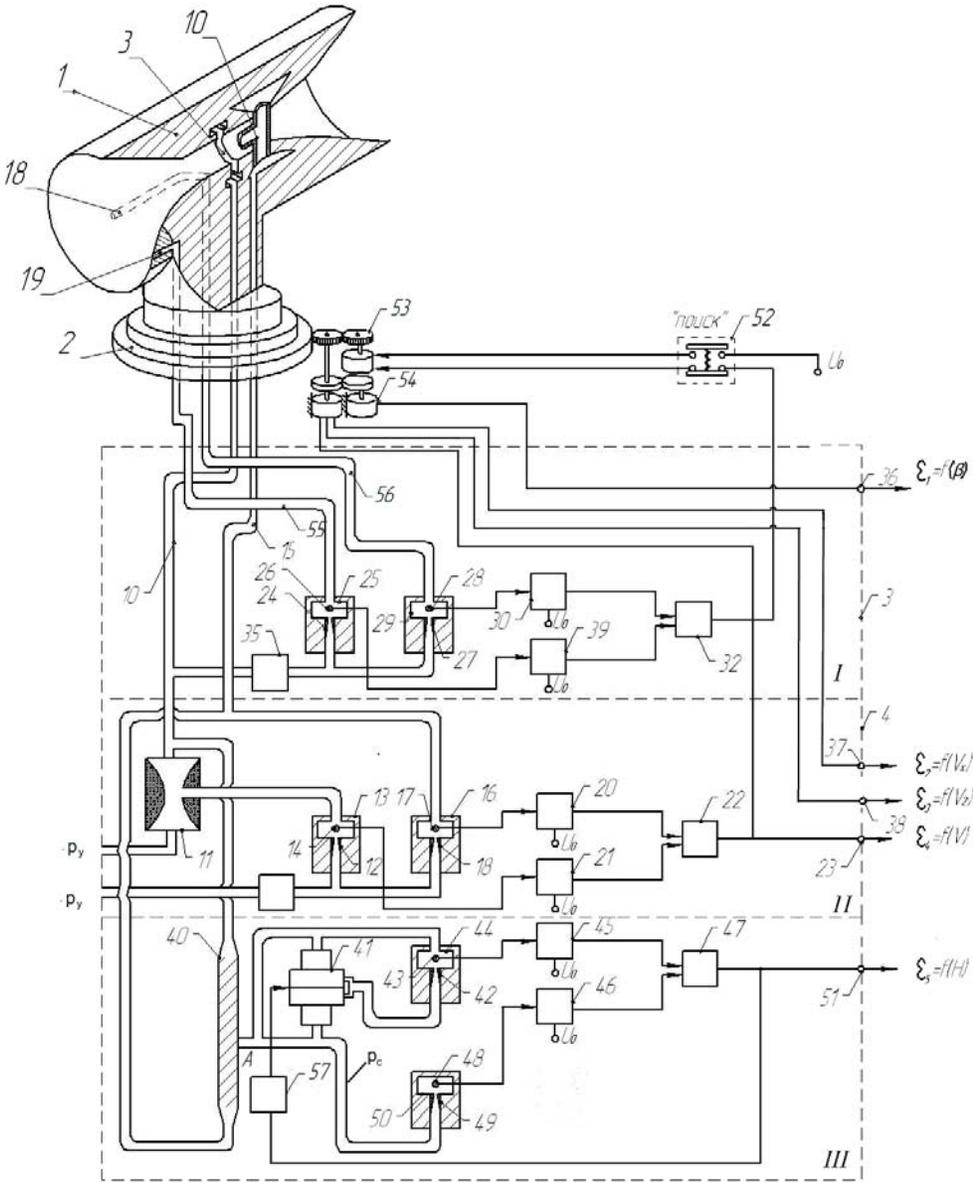


Рисунок 1 – Функциональная схема трехканальной СВС-В с принудительно ориентируемым аэрометрическим преобразователем (обозначено): I – канал угла скольжения; II – канал модуля воздушной скорости V и ее составляющих V_x и V_z ; III – канал барометрической высоты)

Анализ результатов исследования точностных характеристик каналов СВС-В проведем последовательно для каналов угла скольжения, канала воздушной скорости, а затем канала барометрической высоты. Исследование следящей системы канала угла скольжения при мед-

ленно изменяющихся скоростных входных воздействиях и возмущениях типа «порыв ветра» является обязательным этапом при проектировании следящих приборных систем, предназначенных для измерения аэрометрических параметров.

На рисунке 2 представлена структурная схема канала измерения угла скольжения, учитывающая динамические свойства входящих в него преобразователей (табл. П1) Схему можно разделить на три участка. Первый из них включает преобразователи, в которых в качестве информативных сигналов выступают параметры набегающего на АМП-1(рис.1) [6] потока воздуха, затем параметры потока воздуха, движущегося по пневматическому тракту, включающему пневмопроводы 55 и 56 и сообщающему АМП с струйно-конвективными преобразователями 25 и 29. Второй участок включает преобразователь рода энергии информативного сигнала, который выполнен на основе струйно-конвективного преобразователя (СКП), состоящий из анемочувствительных элементов 26 и 28 и схем их включения 30 и 39.

Третий участок включает устройства, осуществляющие повышение мощности информативного сигнала (усилитель напряжения и мощности с коэффициентом передачи $K_{\text{ус}}$) с выхода СКП до уровня, достаточного для его отработки с помощью исполнительного привода (система двигатель-генератор и редуктор), который осуществляет ориентацию АМП по направлению вектора воздушной скорости β в плоскости ее измерения.

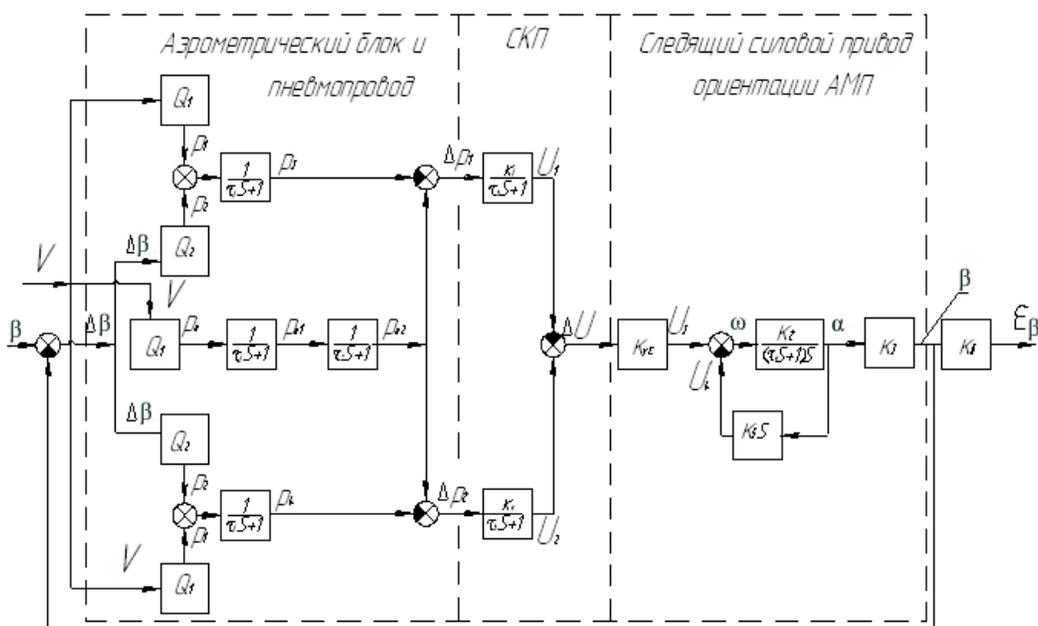


Рисунок 2 – Развернутая структурная схема канала угла скольжения СВС-В

Представленные на структурной схеме преобразователи канала измерения угла скольжения имеют передаточные функции, представленные в табл. 1. Δ

Преобразователи, входящие в пневматические каналы, сообщающие приемные отверстия 18 и 19 АЭП по β с термоанемометрическими модулями, представлены измерительными преобразователями (звеньями) с нелинейной функцией преобразования и коэффициенты передачи которых равны дифференциальной чувствительности Q_1, Q_2 на начальном участке функции преобразования. Усилитель напряжения, пропорционального позиционному сигналу рассогласования $\Delta\beta$, имеет на выходе сигнал U_3 значение которого не может превышать напряжение питания и поэтому усилитель рассогласования является также нелинейным звеном с насыще-

нием. Усилитель мощности имеет аналогичную выходную характеристику. Электродвигатель представлен в виде двух звеньев, первое из которых является апериодическим звеном первого порядка (выход ω – скорость, вращения ротора двигателя), а второй – интегрирующее звено (выход α – угол поворота ротора).

Задача синтеза канала угла скольжения СВС-В является самостоятельной и поэтому выходит за рамки данной работы. Отметим только, что следящая система этого канала нелинейная и поэтому при определении коэффициентов K_3 и K_{VC} используется численный метод, который представляет итеративный процесс и основными критериями при поиске численных значений K_3 и K_{VC} приняты: апериодический характер кривой переходного процесса (исключение колебательности); достижение заданного быстродействия; статическая погрешность канала не должна превышать 1° , а динамическая в режиме отработки скоростного воздействия 20 град/с, не более 5-10 угл.град.

Учет в модели канала угла скольжения большого числа нелинейностей, вообще говоря, не позволяет претендовать на получение каких-либо общих выводов из проведенного исследования. Однако в режимах работы, рассмотренных в процессе имитационного моделирования канала угла скольжения СВС-В, можно определить чувствительность точностных показателей качества этого к вариациям параметров преобразователей, входящих в канал, и качественно оценить эффективность сглаживания скачков входного воздействия в процессе измерительного преобразования информативных сигналов в канал.

Анализ *результатов* моделирования (рисунки 3-5) свидетельствуют о том, что постоянная времени при скачке по углу скольжения, равном 20 угл. град., близка к значению 1,3 с. Переходный процесс следует отнести к апериодическому в независимости от вида возмущений по скорости или углу скольжения.

Одной из важнейших задач имитационного моделирования была оценка влияния разброса чувствительности измерительного канала ΔQ . Анализ переходных процессов (рисунок 3) позволяет сделать вывод, что разброс коэффициента преобразования в пределах 10% не оказывает существенного влияния на форму переходного процесса и постоянную времени, кроме того на основании полученных результатов имитационного моделирования статическая погрешность канала измерения угла скольжения во всех случаях не превышает 0,73 угл.рад.

На рисунке 4 представлены графики переходных процессов канала угла скольжения при воздействии скачка по входному воздействию и возмущению (порыв ветра с 108 км/ч до 18 км/ч и порыв с 18 км/ч до 108 км/ч) для случая как идентичных каналов измерительного преобразования, а также с разбросом $\delta Q = 1\%$.

Влияние порыва ветра на характер переходного процесса состоит, прежде всего, в появлении перерегулирования, которое не превышает 11%. Необходимо отметить, что порыв ветра в сторону уменьшения скорости не оказывает существенного влияния на переходный процесс по углу скольжения при обратном скачке по входному воздействию. Разброс коэффициентов преобразования каналов δQ приводит также к появлению статической ошибки, которая тем меньше, чем больше скорость порыва ветра и для случая на рисунке 3 она не превышает 2,28 угл.град.

Для оценки установившейся погрешности канала угла скольжения в режиме динамического слежения, исследовалась отработка скоростного воздействия с крутизной; 10 и 20 угл. град./с. Эти исследования показали, что значение установившейся погрешности во всем диапазоне скоростей $V = \{18 \dots 108\}$ км/ч не превышает 0,46 угл. град. При разбросе коэффициентов преобразования $\delta Q = 1\%$ выходной сигнал превышает истинное значение на 0,3 угл.град. при $V = 18$ км/ч (5 м/с) и занижен на величину 0,33 угл.град. при $V = 108$ км/ч (30 м/с).

При скоростных воздействиях с крутизной тестового сигнала динамическая погрешность соответственно равна: $\Delta\beta = 0,43$ угл. град. при $\omega_u = 1$ угл.град/с; $\Delta\beta = 3,5$ угл.град. при $\omega_u = 10$ угл. град/с, а для $\omega_u = 20$ угл.град/с установившаяся погрешность, прогрессирующая во времени определяется в соответствии со следующей зависимостью $\Delta\beta = 5,5 + (\omega_u - \omega_m)t$ (где обозначе-

но: (ω_{00} , ω_m – начальная и максимально возможная скорость разворота платформы с АМП, определяемая предельными оборотами выбранного электродвигателя привода).

Для получения более полной оценки динамических характеристик были получены частотные характеристики (рисунок 5) канала угла скольжения при гармонических входных воздействиях, которые позволили определить, что частоты среза канала равна $\sim 0,6$ Гц, запас устойчивости по модулю - 21 дБ, запас устойчивости по фазе – 9 4 град.

В заключение отметим: характер переходных процессов в канале угла скольжения зависит от изменения координат в контуре сервопривода, что заставляет с особым вниманием подходить к проектированию этого узла канала угла скольжения.

Полученные значения угловой скорости электродвигателя сервопривода позволяют сделать обоснованный вывод о пригодности выбранного двигателя к работе в исследуемой следящей системе канала угла скольжения.

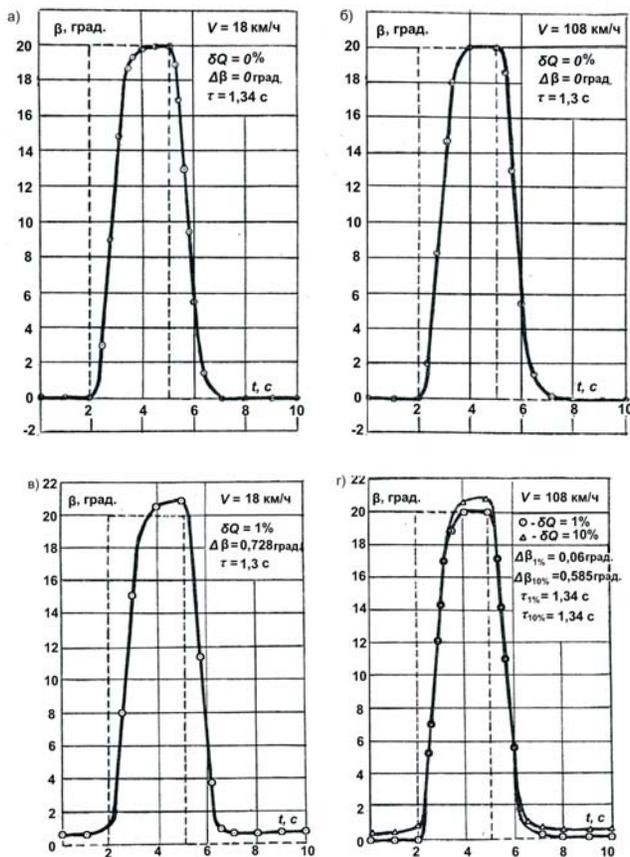


Рисунок 3 – Графики переходных процессов по каналу угла скольжения СВС-В при скорости полета соответственно $V=18\text{км/ч}$ (5м/с) и $V=108\text{км/ч}$ (30м/с) и действии ветровых возмущений в виде скачка угла ветра на 20 угл.град. каналы измерительного преобразования идентичны (а-б); не идентичны (в-г), разброс коэффициентов преобразования соответственно 1% и 10%

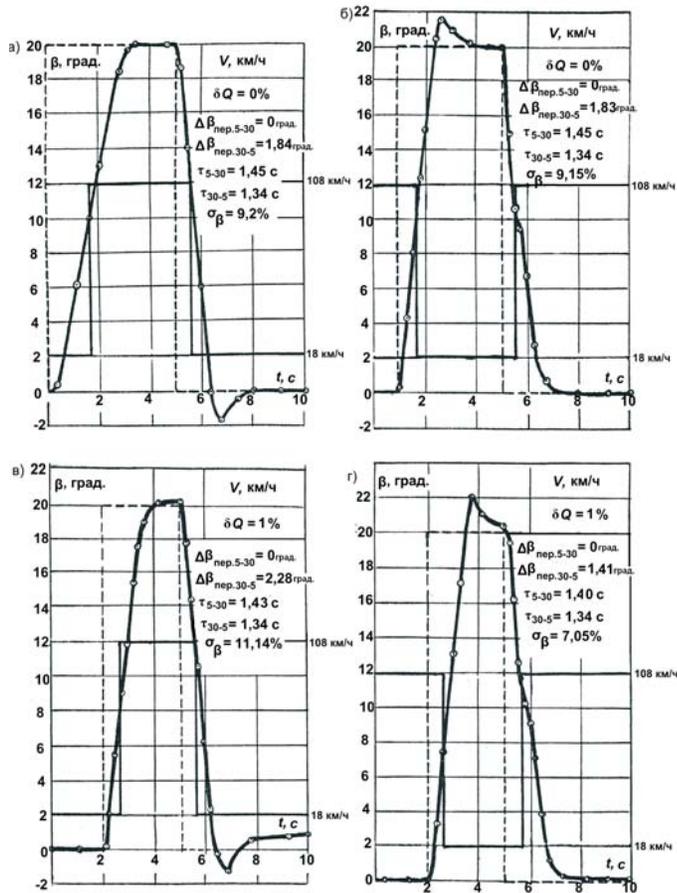


Рисунок 4 – Графики переходных процессов по каналу угла скольжения СВС-В при совместном действии ветровых возмущений в виде скачка угла ветра на 20 угл.град. и сдвинутого по времени прямого и обратного скачков по скорости ветра с 18км/ч до 108 км/ч и с 108 км/ч до 18 км/ч при идентичности параметров измерительных каналов (а -б) и разбросе коэффициентов преобразования каналов, соответственно 1% и 10% (в-г)

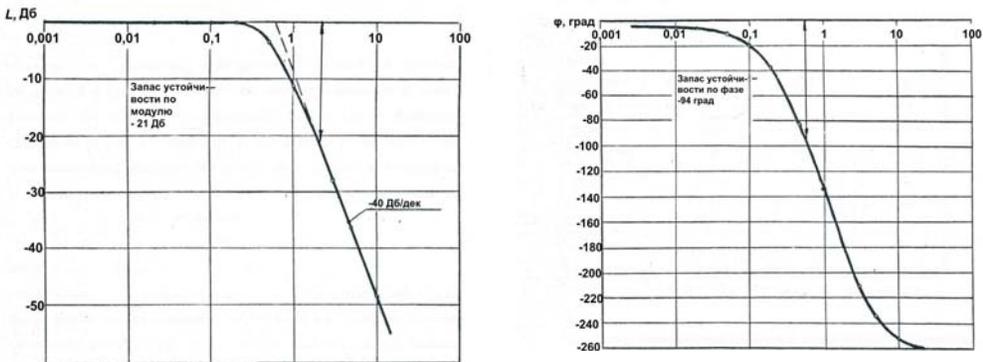


Рисунок 5 – Логарифмические частотные характеристики канала угла скольжения

На рисунке 6 представлена структурная схема канала измерения скорости, учитывающая особенности динамических свойств основных преобразователей (табл. П2). Схему можно разделить на три участка. Первый из них включает преобразователи, в которых в качестве информативных сигналов выступают параметры набегающего на АМП-1(рис.1) потока воздуха, затем параметры потока воздуха, движущегося по пневматическому тракту, включающему пневмопроводы 10 и 15 и сообщающему АМП с анемометрическими модулями 13 и 16. Второй участок включает преобразователь рода энергии информативного сигнала, который выполнен на основе струйно-конвективного преобразователя (СКП), состоящий из анемочувствительных элементов 14 и 17и схем их включения 20 и 21.

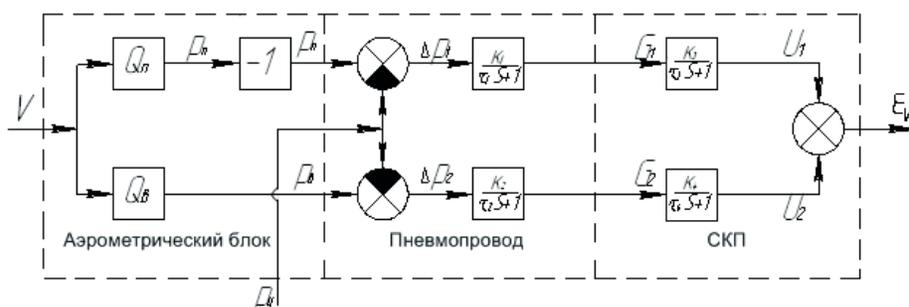


Рисунок 6 – Структурная схема канала воздушной скорости

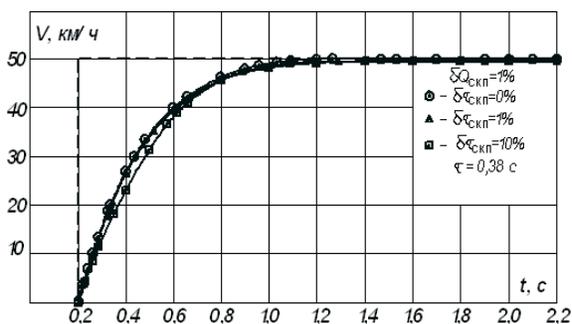
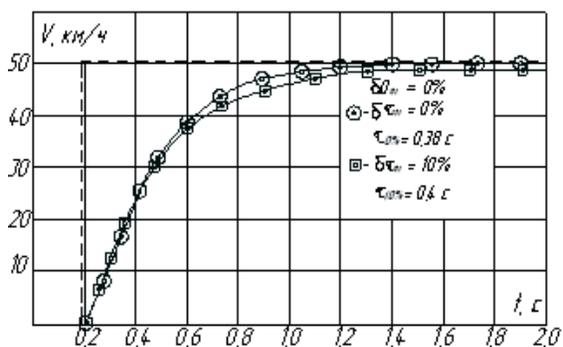


Рисунок 7 – Графики переходных процессов канала воздушной скорости СВС-В: при полной идентичности каналов измерительного преобразования и при разбросе постоянной времени $\delta\tau$ струйно-конвективных преобразователей каналов скорости

Третий участок включает устройства, осуществляющие повышение мощности информативного сигнала (усилитель напряжения и мощности к коэффициентом передачи K_{yc}) с выхода СКП до уровня, достаточного для его обработки с помощью исполнительного привода (система двигатель-генератор и редуктор), который осуществляет ориентацию АМП по направлению вектора воздушной скорости β в плоскости ее измерения.

Анализ результатов моделирования (рисунки 7, 8) свидетельствуют о том, что постоянная времени при полной идентичности каналов измерительного преобразования, равна 0,38 с.

Одной из важнейших задач имитационного моделирования канала скорости была оценка влияния разброса постоянной времени $\delta\tau$ струйно-конвективных преобразователей. Анализ переходных процессов (рисунок 7) позволяет сделать вывод, что разброс постоянной времени в пределах 10% не оказывает существенного влияния на форму переходного процесса.

Информативный канал барометрической высоты

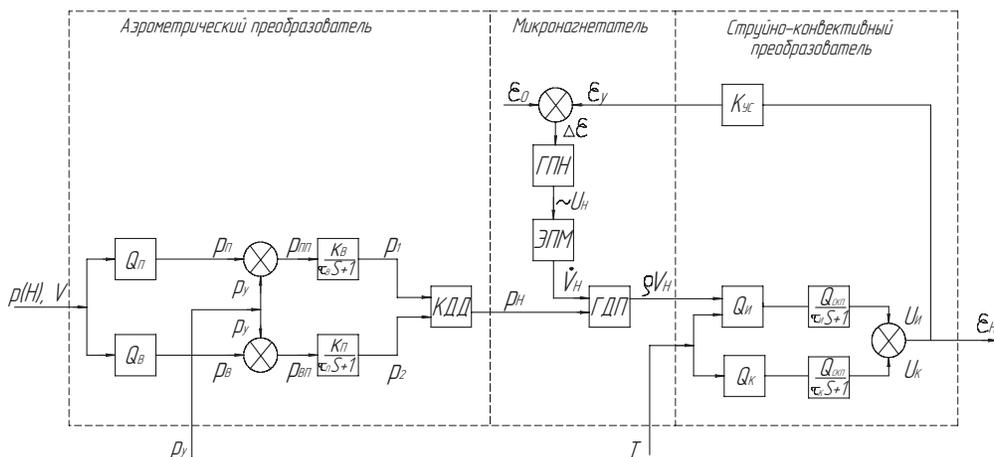


Рисунок 8 – Структурная схема информативного канала барометрической высоты
 КДД – капиллярный делитель давления, ГДП – газодинамический преобразователь,
 ЭПМ – электромагнитный привод мембраны, ГПН – генератор привода нагнетателя

На рисунке 8 представлена структурная схема канала барометрической высоты, отражающая особенности его построения и функционирования, в частности формирования сигнала статического давления, которое осуществляется с помощью капиллярного делителя давлений [А.с. №695288 Преобразователь статического давления на ЛА /Авт.: Порунов, В.А.Ференец.1979.], 40 (здесь и далее рис.1), входы которого сообщены посредством пневмопроводов 10 и 15 с приемными отверстиями АМП 3 и 10 соответственно полного p_n и дросселированного p_b давлений. Этот процесс формирования сигнала статического давления p_n структурно представлен (рис.5) в виде преобразователей имеют нелинейную функцию преобразования $\Delta p = f(V)$ и $\Delta p_{II} = f(V)$ и описываются дифференциальными чувствительностями Q_{II} , Q_B , которые получены по результатам экспериментальных исследований макета АЭП. Для обеспечения возможности моделирования АМП в канале барометрической высоты, экспериментально полученные зависимости аппроксимировались по методу наименьших квадратов полиномом третьего порядка. В результате в характерной точке, расположенной в пределах длины капиллярного делителя 40 формируется сигнал статического давления, которое подается в газодинамический преобразователь микронагнетателя 41, управляемого генератором привода нагнетателя 57, на вход которого подается напряжение рассогласования.

Принцип действия канала барометрической высоты основан на измерении посредством ТАПР массового расхода потока воздуха по пневматическим каналам, вызываемого работой

микронагнетателя, и плотность которого определяется величиной статического давления/(барометрической высотой) поступающего от приемника давления, при условии стабилизации температуры в пневмоканалах.

На рисунке 8 представлена структурная схема канала измерения барометрической высоты, учитывающая динамические свойства входящих в него преобразователей (табл. ПЗ). Схему можно разделить на три участка.

Первый участок из них включает преобразователи, в которых в качестве информативных сигналов выступают параметры набегающего потока воздуха на АМП-1 (рис.8), движущегося по пневматическому тракту, включающему пневмопроводы 55 и 56.

Во второй участок входит микронагнетатель. Микронагнетатель 41 вызывает перетекание воздуха по пневматическому каналу, который формируется соплом в струю, омывающую измерительный пневмочувствительный элемент. Производительность микронагнетателя 41 и расход воздуха в канале а, следовательно, и степень охлаждения измерительного анемочувствительного элемента (АЧЭ) измеряется до тех пор, пока напряжение на выходе электроизмерительной схемы (ЭИС) не уравнивается на входе усилителя напряжению от ЭИС. Таким образом, напряжение на выходе усилителя, регулируемое с помощью потенциометра задатчика начальной высоты определяет величину опорного массового расхода через пневматический канал и тем самым задается начальная высота отсчета по каналу измерения барометрической высоты

Третий участок включает преобразователь рода энергии информативного сигнала, который выполнен на основе струйно-конвективного преобразователя (СКП), состоящий из анемочувствительных элементов 44 и 48 и схем их включения 45 и 46.

При исследовании динамических характеристик канала барометрической высоты необходимо было определить параметры цепи обратного преобразования, а именно: коэффициент усиления усилителя $K_{УС}$ и значение опорного напряжения питания ε_0 определяющее расчетный режим работы микронагнетателя. С этой целью первый машинный эксперимент проводится с моделью канала, включая получение зависимости выходного сигнала ε канала от $K_{УС}$ и ε_0 . Графики этих зависимостей получены на основе параметров преобразователей канала барометрической высоты сведенных в табл. ПЗ и представлены на рисунке 9 и из них можно сделать вывод, что в качестве наиболее рациональных значений $K_{УС}$ и ε_0 могут быть приняты значения $K_{УС} = 50$ и $\varepsilon_0 = 2,8$ В. Для окончательной оценки выбранных значений $K_{УС}$ и ε_0 был проведен прогон модели канала барометрической высоты в процессе которого определялась его статическая характеристика, при вариациях $\varepsilon_0 = 1...3$ В. Этот эксперимент позволил также оценить влияние ε_0 не только на дифференциальную чувствительность и порог реагирования, но и на степень нелинейности статической характеристики канала. Результаты такого прогона модели канала барометрической высоты представлены на рисунке 10 из которого следует, что принятые значения $K_{УС}$ и ε_0 близки к наилучшим, хотя не удалось достигнуть полностью компромисса между дифференциальной чувствительностью и степенью нелинейности статической характеристики канала. Все последующие машинные эксперименты с моделью канала проведены при выбранных значениях $K_{УС}$ и ε_0 .

Динамические свойства канала барометрической высоты оценивались на основе кривых переходных процессов, полученных в результате машинных экспериментов, проведенных при крайних значениях базовой высоты $H_6 = 10$ м. и $H_6 = 10$ км, при этом амплитуда скачка в первом случае не превышала ± 10 м, а во втором ± 100 м. Результаты этих машинных экспериментов приведены на рисунках 11 и 12, из которых следует, что кривые переходных процессов имеют характер, присущий аperiodическим преобразователям. Постоянная времени канала с высотой возрастает с 0,022 до 0,045 (без учета зависимости постоянных времени пневмотрассы от высоты). Статическая погрешность канала пренебрежимо мала и не превышает соответственно для базовых высот:

$$H_6 = 10 \text{ м}; \Delta H \leq 0,05...0,1 \text{ м} \text{ и } H_6 = 10 \text{ км} \quad \Delta H \leq 1...2 \text{ м.}$$

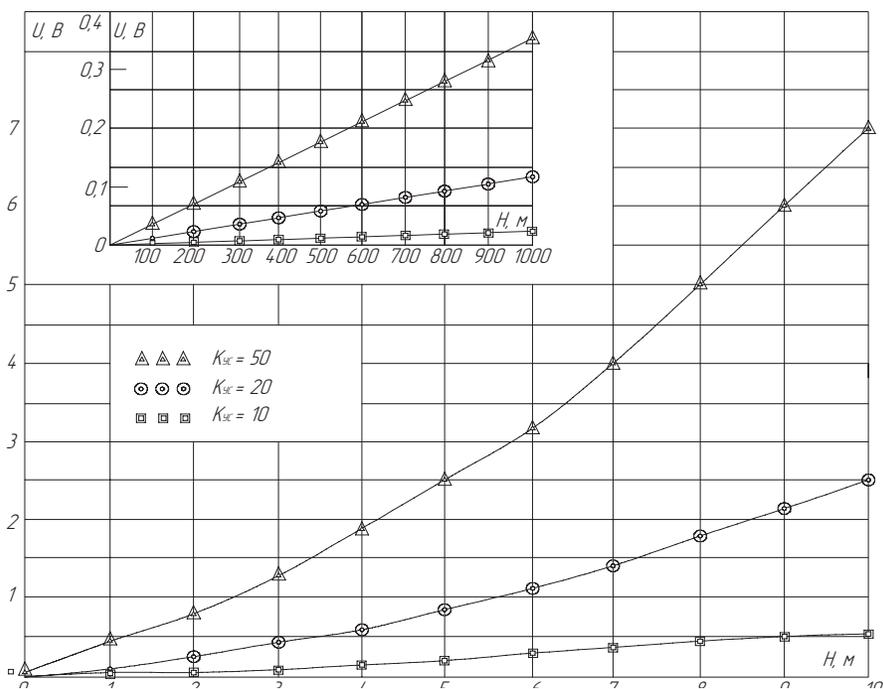


Рисунок 9 – Графики зависимости выходного сигнала канала $U=f(H)$ барометрической высоты от коэффициента усиления $K_{ус}$ усилителя и от величины опорного напряжения питания \mathcal{E}_0 микронагнетателя

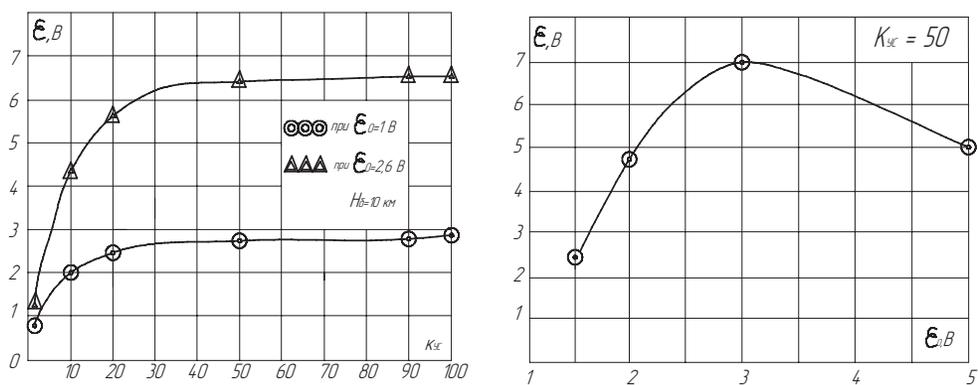


Рисунок 10 – Графики статических характеристик канала барометрической высоты при вариациях величины опорного напряжения питания \mathcal{E}_0 микронагнетателя

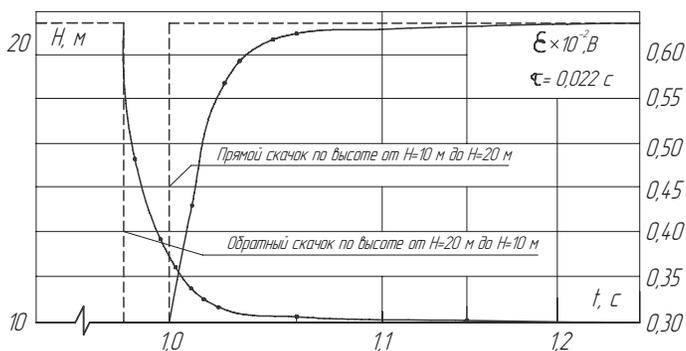


Рисунок 11 – Графики переходных процессов системы СВС по каналу барометрической высоты при полете на базовой высоте $H_0=10$ м и амплитудах скачков ± 10 м

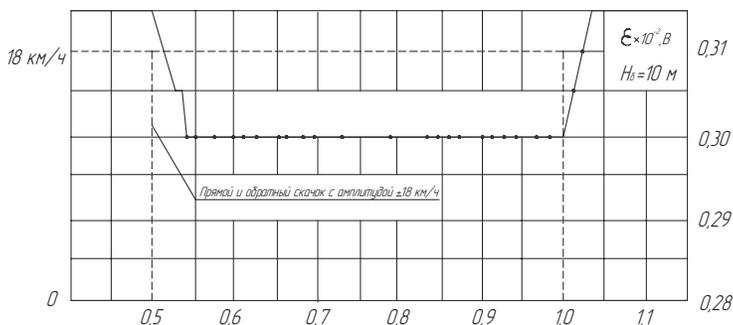


Рисунок 12 – Графики переходных процессов канала барометрической высоты при полете со скоростью 18 км/ч на базовой высоте 10 м и воздействии возмущений - порыв ветра с амплитудой ± 18 км/ч

Сильное влияние на работу канала стабилизации высоты летательных аппаратов, осуществляемую в настоящее время по сигналам штатных измерителей типа КВ-11 и КЗВ, оказывают аэродинамические погрешности аэрметрических преобразователей, обусловленные неустойчивостью приземного слоя атмосферы. Основным возмущающим фактором являются порывы ветра, имеющие переменную амплитуду, частоту и направление. С целью оценки влияния этого возмущающего фактора на канал барометрической высоты был проведен машинный эксперимент, в котором на канал было оказано воздействие порывов ветра как ступенчато возрастающих, так и знакопеременных. Из результатов машинных экспериментов представлены на рисунках 11, 12, из которых следует, что статическая погрешность канала прогрессирует пропорционально возмущающему фактору (амплитуда порыва ветра, см. рисунок 12). Однако для порывов ветра с изменяющимся на $\pm 180^\circ$ направлением (рисунок 12) после прекращения действия возмущения статическая погрешность стремится к нулю.

Большое влияние на метрологические характеристики канала барометрической высоты имеют характеристики системы термостатирования пневмоканала, содержащего анемочувствительные элементы. С целью определения требований к точности стабилизации температуры пневмоканала с анемочувствительными элементами был проведен машинный эксперимент по исследованию зависимости выходного сигнала канала ϵ от температуры окружающей среды при отключенной системе термостатирования и вариациях процента разброса крутизны Q_{HT} зависимостей коэффициента рассеяния от температуры анемочувствительных элементов. Ре-

зультаты машинных экспериментов приведены на рисунке 13, из которого следует, что для обеспечения значения методической температурной погрешности в пределах 1% требуется стабилизировать температуру в пневмоканалах содержащих анемочувствительные элементы с точностью 1...2 °С, что является приемлемым для современных систем термостатирования. Из рисунка 13 также следует вывод, что с увеличением разброса крутизны Q_{HT} зависимости коэффициента рассеяния анемочувствительного элемента в положительную сторону, температурная погрешность уменьшается, а, следовательно, требования к системе термостатирования облегчаются.

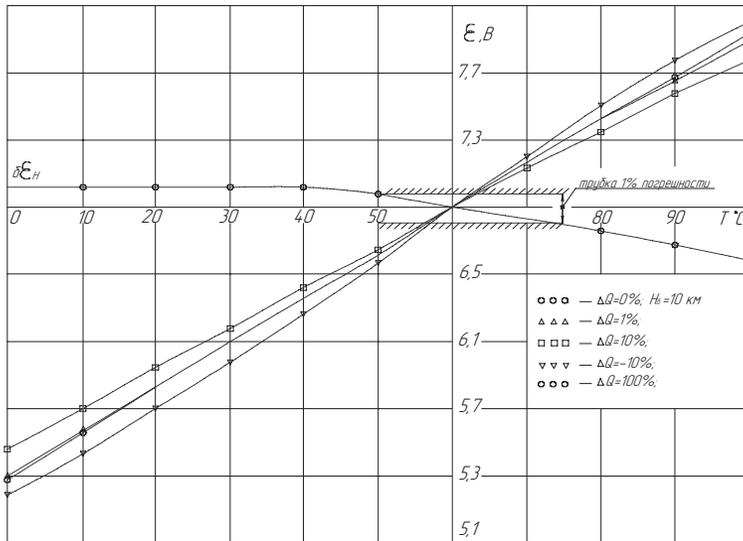


Рисунок 13 – Графики зависимостей выходного сигнала ϵ_H канала барометрической высоты от температуры в измерительной камере для случая отказа системы термостатирования

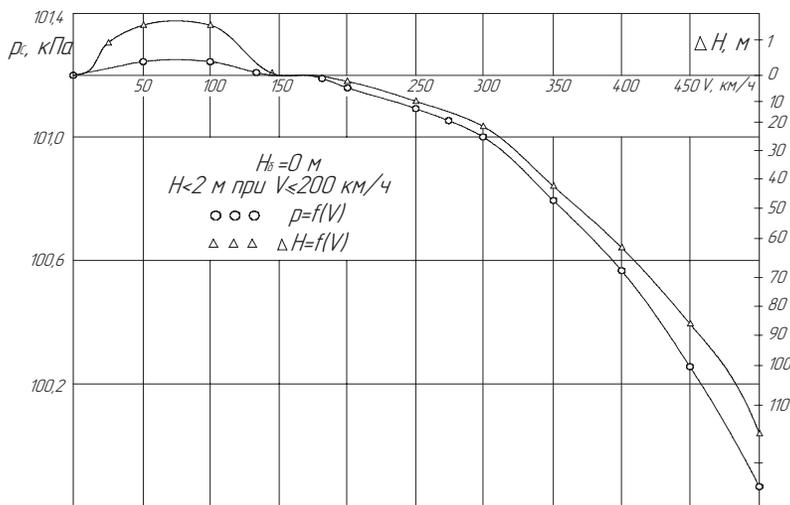


Рисунок 14 – Графики зависимостей выходного сигнала по статическому давлению p и абсолютной погрешности измерения барометрической высоты от вариаций скорости полета

Заключительным машинным экспериментом по исследованию канала барометрической высоты является оценка скоростной погрешности капиллярного делителя давления 40 (рисунок 1) в рабочем диапазоне скоростей. Результаты эксперимента представлены на рисунке 14, из которого следует, что погрешность формирования сигнала статического давления посредством пневмоделителя в диапазоне скоростей 0...500 км/ч изменяется соответственно в пределах 0...1300 Па, а приведенная ко входу скоростная погрешность канала в диапазоне скоростей полета 0...200 км/ч не превышает величины $\approx 2\text{м}$. Основной вклад в скоростную погрешность капиллярного делителя давлений 40 вносит неточность определения точки отбора давления на капилляре делителя давлений 40.

Таким образом, представленный в работе анализ результатов имитационного моделирования каналов СВС-В построенный на основе принудительно ориентированного АМП позволило решить следующие задачи: определить наиболее важные для каждого из каналов СВС-В измерительные преобразователи, определяющие показатели качества, а также сформулировать рекомендации по численным значениям параметров этих преобразователей (коэффициент усиления $K_{yc}=5300$ и коэффициент передачи $K_s=0,01$ гибкой обратной связи) для канала угла скольжения; по каналу барометрической высоты - коэффициент усиления $K_{yc}=50$, опорное напряжение $\epsilon_0=2,8\text{ В}$; оценить динамические свойства каждого из каналов при тестовых воздействиях типа скачок, скоростное и гармоническое воздействие с параметрами, соответствующими рабочим условиям эксплуатации СВС-В, показано, что переходные процессы по каждому из каналов имеют аperiodический характер; оценить влияние разброса параметров СКП (чувствительности и постоянные времени) на показатели качества каналов, постоянную времени и статическую погрешность; определить требования к системе термостатирования канала барометрической высоты, требуемая точность стабилизации температуры измерительной камеры $1-2^\circ$; определить характеристику скоростной погрешности канала барометрической высоты (сигнал p_c) от степени нелинейности характеристик АМП по каналам полного и дросселированного давления, а также точности определения координаты отбора давления p_c , формируемого капиллярным делителем давления.

Список литературы

- [1] Козицын В.К., Макаров Н.Н., Порунов А.А., Солдаткин В.М. Анализ принципов построения СВС вертолета//Авиакосмическое приборостроение, №3, 2003; 12.Порунов А.А., Солдаткин В.В., Солдаткин В.М. Измерительно-вычислительная система определения малых воздушных скоростей вертолета// Авиакосмическое приборостроение, №11, 2004 год
- [2] Порунов А.А.Цифровое моделирование информативных каналов измерителя аэрометрических параметров // Международная научно-техническая конференция "Актуальные проблемы математического моделирования и автоматизированного проектирования. Тезисы доклада (Модель - проект- 95) - Казань:1995, с.71-73.
- [3] Порунов А.А., Солдаткин В.В. Моделирование комплексной системы измерения малых воздушных скоростей вертолета. Современные технологии-ключевое звено в возрождении отечественного авиастроения: Материалы Международной научно-практической конференции.Т.2. Казань: Изд-во Казан. гос. техн. ун-та, 2008 с.76.
- [4] Соболев В.И., Ледяев В.В. Математические аспекты теории аэрометрии высотно-скоростных параметров//Приборы и системы. Управление, контроль, Диагностика. №8, 2000г., с. 50-55.
- [5] Arena – система имитационного моделирования.-2001. (www.interface.ru/sysmod/arena.htm)
- [6] Соболев В.И., Ледяев В.В. Математическое моделирование приемников воздушных давлений дозвуковых скоростей//Приборы и системы. Управление, контроль, Диагностика. №8, 2000г., с. 55-56.

Приложение

Таблица П1

№ п/п	Измерительный преобразователь	Передаточная функция или функция преобразования	Значения параметров
1.	Аэрометрический преобразователь (АМП)	$W_{АЭП}(S) = Q_{АЭП}$, где $Q_{АЭП}$ – чувствительность АМП, $Q_{АЭП} = Q_1 + Q_2$	$Q_1 = 2$ Па/(км/ч); $Q_2 = 2$ Па/град. (при $V=36$ км/ч, $\beta=30$ град)
2.	Пневматический тракт	$W_{ПТ}(S) = \frac{1}{\tau_{ПТ} s + 1} e^{\tau_{a}s}$, $\tau_{ПТ}$ – постоянная времени, определяемая соотношением $\tau_{ПТ} = r C_n$ (r_n – пневматическое сопротивление и емкость пневмоканала);	$r_n = \frac{128 \nu l}{\pi d^4}$; при $V = 3,77 \cdot 10^{-3}$ м ³ ; $l = 0,3$ м; $d = 4 \cdot 10^{-3}$ м, $\tau_{ПТ} = 0,021$ с; $\tau_a = l/a$ (a – местная скорость звука) $\tau_a = 0.001$ с.
3.	Проточная пневматическая камера	$W_{ПК} = \frac{Q}{\tau_{ПК} s + 1}$, где $\tau_{ПК}$ – постоянная времени пневмокамеры, $\tau_{ПК} = \frac{V_{ПК} p_k}{K_g R T (G_1 + G_2)}$ (здесь p_k , V_k – давление и объем пневмокамеры; G_1 , G_2 – расходы воздуха по каналам, содержащим термоанеморезистивные капсулы).	G_1, G_2 соответствуют $V = 18$ км/ч и углу $\beta \leq 5^\circ$, $\tau_{ПК} = 0,0001$ с.
4.	Преобразователь перепада давления в расход	$W_{ПР}(S) = Q_{ДР}$, где $Q_{ДР}$ – дифференциальная чувствительность преобразователя	$Q_{ДР} = 1,43 \cdot 10^{-7} \frac{\text{кг/с}}{\text{Па}}$
5.	Усилитель рассогласования	$W_{УС}(S) = K_{УС}$, где $K_{УС}$ – коэффициент усиления.	$K_{УС} = 50$
6.	Струйно-конвективный преобразователь (СКП)	$W_{СКП}(S) = \frac{Q}{\tau_{СКП} s + 1}$, где $Q_{СКП}$ – дифференциальная чувствительность СКП, В/(кг/с), $\tau_{СКП}$ – постоянная времени термоанеморезистора,	$\tau_{СКП} = 0,10$.
7.	Тахогенератор	$W_{ТГ}(S) = K_S S$, где K_S – коэффициент пропорциональности.	$K_S = 0,1$
8.	Электродвигатель	$W_{ДВ}(S) = \frac{K_{ДВ}}{s(\tau_D s + 1)}$, где $K_{ДВ}$ – коэффициент передачи электродвигателя. Для ДГ-2ТА τ_D – электрохимическая постоянная двигателя, $\tau_D = I_{np} \omega_H / M_H - M_H$ (здесь I_{np} – приведенный момент инерции), ω_H – номинальная скорость вращения	$K_{ДВ} = 27,8$ рад·К/В. $I = I_p + \frac{I_{np}}{ip^2} = 5,38 \cdot 10^{-3}$ $\omega_H = 837,8$ рад/с; $\tau_D = 0,135$.

Окончание таблицы П1

9.	Редуктор	$W_p(S) = K_p$, где K_p – коэффициент передачи редуктора; $K_p = 1/i_p$ (i_p – передаточное отношение редуктора)	$i_p = 2400$
----	----------	--	--------------

Таблица П2

№ п/п	Измерительный преобразователь	Передаточная функция или функция преобразования	Значения параметров
1.	Аэрометрический преобразователь (АМП)	$W_{АЭП}(S) = Q_{АЭП}$, где $Q_{АЭП}$ – чувствительность АМП, $Q_{АЭП} = Q_{II} + Q_B$ $Q_B : \Delta p_B = a_B - b_B \beta - c_B \beta^2$ $Q_{II} : \Delta p_{II} = a_{II} + b_{II} \beta + c_{II} \beta^2$	$a_n = 8,51$; $b_n = 0,99$; $c_n = 0,011$ (при 36 км/ч)
2.	Пневматический тракт	$W_{ПТ}(S) = \frac{1}{\tau_{ПТ} s + 1} e^{\tau_{вс}}$, $\tau_{ПТ}$ – постоянная времени, определяемая соотношением $\tau_{ПТ} = r C_n$ (r_n – пневматическое сопротивление и емкость пневмоканала);	$r_n = \frac{128 \nu l}{\pi d^4}$; при $V = 3,77 \cdot 10^{-3} \text{ м}^3$; $l = 0,3 \text{ м}$; $d = 4 \cdot 10^{-3} \text{ м}$; $\tau_{ПТ} = 0,021 \text{ с}$; $\tau_a = l/a$ (a – местная скорость звука) $\tau_a = 0,001 \text{ с}$.
3.	Проточная пневматическая камера	$W_{ПК} = \frac{Q}{\tau_{ПК} s + 1}$, где $\tau_{ПК}$ – постоянная времени пневмокамеры, $\tau_{ПК} = \frac{V_{ПК} p_k}{K_g R T (G_1 + G_2)}$ (здесь p_k , V_k – давление и объем пневмокамеры; G_1 , G_2 – расходы воздуха по каналам, содержащим струйно-конвективные преобразователи).	G_1, G_2 соответствуют $V = 18 \text{ км/ч}$ и углу $\beta \leq 5^\circ$, $\tau_{ПК} = 0,0001 \text{ с}$.
4.	Преобразователь перепада давления в расход	$W_{ПР}(S) = Q_{\Delta p}$, где $Q_{\Delta p}$ – дифференциальная чувствительность преобразователя	$Q_{\Delta p} = 1,43 \cdot 10^{-7} \frac{\text{кг/с}}{\text{Па}}$
5.	Струйно-конвективный преобразователь (СКП)	$W_{СКП}(S) = \frac{Q}{\tau_{СКП} s + 1}$, где $Q_{СКП}$ – дифференциальная чувствительность СКП, В/(кг/с), $\tau_{СКП}$ – постоянная времени термоанемомонориста,	$\tau_{СКП} = 0,10 \text{ с}$

Таблица П3

№ п/п	Название	Функция преобразования	Числовое значение
1.	Аэрометрический преобразователь	$\Delta p_{II} = (c_2 V + c_1 V^2 + c_0 V^3) \bar{\rho} = Q_{II} \cdot \bar{\rho}$ $\Delta p_B = (b_2 V + b_1 V^2 + b_0 V^3) \bar{\rho} = Q_B \cdot \bar{\rho}$	$Q_{II} \approx 1$ $Q_B \approx 1,5$

Окончание таблицы ПЗ

		$\bar{\rho} = \frac{\rho(H)}{\rho(H_0)}$ <p>где $c_0, c_1, c_2, b_0, b_1, b_2$, – коэффициенты полинома; V – скорость полета ЛА, км/ч; $\bar{\rho}$ – относительная плотность воздуха на высоте H; Q_B, Q_{Π} – дифференциальная чувствительность по каналу Вентури и Пито</p>	
2.	Пневматический тракт	$W_{\Pi T}(S) = \frac{K_B}{\tau_B S + 1},$ $W_{\Pi T}(S) = \frac{K_{\Pi}}{\tau_{\Pi} S + 1},$ <p>где τ – температурный градиент; K_B – коэффициент передачи канала Вентури; K_{Π} – коэффициент передачи канала Пито;</p>	$K_{\Pi} = K_B = 1$ $\tau_B = \tau_{\Pi} = 0,021 \text{ с}$
3.	Капиллярный делитель давления (КДД)	$K_{\partial} = \frac{Q_B}{Q_{\Pi}}$ <p>где Q_B, Q_{Π} – дифференциальная чувствительность по каналу Вентури и Пито; K_{∂} – капиллярный делитель давления</p>	$Q_{\Pi} \approx 1$ $Q_B \approx 1,5$ $K_{\partial} = 1,5$
4.	Микронагнетатель включает: ГДП – газодинамический преобразователь; ГПН – генератор привода нагнетателя ЭПМ – электромагнитный привод мембраны	$\Delta q = Q_p \Delta p + Q_v \Delta \dot{V}$ $Q_{ГДП} = \kappa_H \dot{V}_0;$ $Q_{ГПН} = \kappa; Q_U = \kappa_s v \dot{V}$ <p>κ_n – коэффициент пропорциональности, $\kappa_n = 1/\kappa_g R T_0 f_{\partial p}$. V_0 – начальная скорость изменения объема полости нагнетательного элемента.</p>	$g = 9,8 \text{ м/с}^2;$ $R = 29,27 \text{ м/град};$ $T_0 = 323^\circ \text{К};$ $\kappa_n = 21,4 \text{ с}^2/\text{м}^4;$ $Q_{ГПН} = 10^2$ $Q_U = 2,53 \cdot 10^{-6} \text{ м}^3/\text{с/В}$ $Q_{ГДП} = 162,6 \cdot 10^{-2} \text{ кг/с/м}$
5.	Струйно-конвективный преобразователь (СКП)	$W_{СКП}(S) = \frac{Q_{СКП}}{\tau_H S + 1},$ $W_{СКП}(S) = \frac{Q_{СКП}}{\tau_K S + 1}$ <p>где $Q_{СКП}$ – дифференциальная чувствительность СКП, $Q_{СКП} = \frac{B}{\kappa_2 / \text{с}}$, τ_H – постоянная времени измерительная, τ_K – постоянная времени компенсационная.</p>	$\tau_{СКП} = 0,10 \text{ с}$

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССА ОПРЕДЕЛЕНИЯ НАГРУЗКИ ОТ КОЛЕСНЫХ ПАР ЛОКОМОТИВА

П.Ю. Вильвер, А.В. Протасов

Институт динамики систем и теории управления СО РАН (ИДСТУ СО РАН)
664033, Иркутск, ул. Лермонтова, 134, Россия
wilwer@mail.ru, arprot@icc.ru
тел: (3952) 45-31-57 факс: (3952) 45-31-57

Ключевые слова: имитационная модель, сеть Петри, OPC – сервер, развеска локомотивов

Abstract

In given article the example of creation of the operating program of a measuring complex on the basis of network model working in real mode of time is described.

Введение

Техническое состояние локомотивов является одной из основ надёжности и безопасности движения подвижного состава. Не последнюю роль при этом играет взаимозависимость между состоянием пути и характером его использования. Для системы колесо - рельс характерны высокие и неравномерные динамические нагрузки, накопительный характер нагрузки, являющейся причиной износа и возникновения неисправностей. В связи с этим распределение статических нагрузок по элементам колесных пар на локомотивах требует специальной регулировки, что способствует уменьшению степени износа колесных пар, и как следствие, сокращение эксплуатационных расходов железнодорожного транспорта.

В процессе испытания локомотива одним из основных пунктов является статическая развеска, позволяющая осуществить такое взаимное расположение его узлов и агрегатов, при котором было бы реализовано равномерное распределение нагрузок от колёсных пар на рельсы.

В связи с вышесказанным была сформулирована цель - разработка программно – измерительного комплекса развески локомотива с использованием имитационной модели.

За основу математической модели системы взят сетевой формализм (СФ), берущий свое начало от сетей Петри [1,2].

СФ имеет ряд полезных при моделировании свойств числовых, самомодифицирующихся, временных и приоритетных сетей Петри, и способен представлять локальное управление и параллелизм, многоуровневую иерархию, недетерминизм, синхронизацию, течение времени и потоки данных. Формальное определение СФ приведено в [3], который был расширен введением иерархических сетей, функцией динамического обмена данными модели и первичных датчиков, прогностических функций и механизмом динамического изменения прогностических функций в зависимости от значений параметров меток (данных от датчиков) [4].

Описание СФ

$$NF = \langle P, T, M, CM, LEV, OPC, I, O, R_i, M_0 \rangle,$$

где $P = \{p_1^{prt_1}, p_2^{prt_2}, \dots, p_{N_p}^{prt_{N_p}}\}$ – множество позиций. Позициями сети моделируются компоненты ТС. на котором определена система вложенных сетей NF
 $prt_i \in N^+ \leq N_i, prt_i \neq prt_j, i, j = \overline{1, N_i}$.

$T = \{t_1^{prt_1}, t_2^{prt_2}, \dots, t_{N_t}^{prt_{N_t}}\}$ – множество переходов, на котором определена система приорите-

тетов $prt_i \in N^+ \leq N_t, prt_i \neq prt_j, i, j = \overline{1, N_t}$.

$M: P \rightarrow MET \cup \{M(p) | p \in P\}$ – маркировка сети, где множество меток MET определяется как совокупность векторов с указанием приоритета метки prm_i , типа (цвета) метки tm_i и множества параметров $(at_{i1}, at_{i2}, \dots, at_{iN_a})$, описывающих этот тип, с историей значений:

$$MET = \left\{ \begin{array}{l} (m_1^{prm_1}, m_2^{prm_2}, \dots, m_{N_m}^{prm_{N_m}}) | m_i^{prm_i} = \{tm_k, (at_{k1}, at_{k2}, \dots, at_{kN_{ka}})^{his}\} \\ prm_i \in N^+ \leq N_m, prm_{i1} \neq prm_{i2}, i1, i2, i = \overline{1, N_m}, \\ tm_k \in TM, at_{kj} \in AT_j, k = \overline{1, N_k}, j = \overline{1, N_{ka}}, his = \overline{1, N_{his}} \end{array} \right\},$$

где TM – множество типов меток, AT_j – область допустимых значений параметра at_{ij} , N_k – число типов меток, N_{ka} – число параметров, описывающих тип метки, N_{his} – число предыдущих значений параметров меток. История значений меток N_{his} является основой для прогнозирования значений параметров модели.

$I = (I_m, I_f, I_l, I_{lev})$ – входная функция,

где $I_m: P \times T \rightarrow MET \cup \{M(p) | p \in P\}$ – функция, определяющая возможность срабатывания перехода и изменение параметров меток по каждому из цветов множества TM ;

$$I_f: P \times T \rightarrow \left\{ (f_{k1}, f_{k2}, \dots, f_{N_k N_{ka}})^k | f_{kj} = f(at_{kj}), k = \overline{1, N_k}, j = \overline{1, N_{ka}} \right\},$$

$$I_l: P \times T \rightarrow \left\{ (L_{k1}, L_{k2}, \dots, L_{N_k N_{ka}})^k | at_{kj} \in L_{kj}, k = \overline{1, N_k}, j = \overline{1, N_{ka}} \right\} –$$

функции разметки входных дуг, где f_i – алгебраическая функция, определяющая метод вычисления значения параметра метки в позиции p ; L_i – ограничение на значение параметра метки в позиции p . Использование ограничений позволяет разделить множество состояний функционирования (МСФ) на подмножества с номинальными, допустимыми и недопустимыми значениями параметров и таким образом, становится возможным моделировать отказы и инциденты, возникающие при функционировании ТС.

$I_{lev}: P^{lev_i} \times T^{lev_i} \rightarrow P^{lev_{i+1}} \times T^{lev_{i+1}}$ – функция, описывающая взаимодействие уровней сети на входных дугах.

Аналогично определяется выходная функция:

$$O = (O_m, O_f, O_l, O_{pro}, O_{lev}),$$

где $O_m: T \times P \rightarrow MET \cup \{M(p) | p \in P\}$,

$$O_f: T \times P \rightarrow \left\{ (f_{k1}, f_{k2}, \dots, f_{N_k N_{ka}})^k | f_{kj} = f(at_{kj}), k = \overline{1, N_k}, j = \overline{1, N_{ka}} \right\}$$

$$O_l: T \times P \rightarrow \left\{ (L_{k1}, L_{k2}, \dots, L_{N_k N_{ka}})^k | at_{kj} \in L_{kj}, k = \overline{1, N_k}, j = \overline{1, N_{ka}} \right\}.$$

Функции раскраски I_m и O_m определяют тип и параметры потоков, изменяющихся в данном аппарате. Функции I_f и O_f соответствуют операциям, производимыми над потоками.

$$O_{pro}: T \times P \rightarrow \left\{ (f_{k1}^{pro}, f_{k2}^{pro}, \dots, f_{N_k N_{ka}}^{pro})^k | f_{kj}^{pro} = f^{pro} \left(\{at_{kj}\}^{his} \right), k = \overline{1, N_k}, j = \overline{1, N_{ka}}, his = \overline{1, N_{his}} \right\} –$$

функция прогнозирования значений параметров метки на основе предыдущих (исторических) значений. Необходимо отметить, так как история значений меток динамически изменяется, то обеспечивается возможность динамического прогнозирования.

$O_{lev} : T^{lev_i} \times P^{lev_i} \rightarrow T^{lev_{i+1}} \times P^{lev_{i+1}}$ – функция, описывающая взаимодействие уровней сети на выходных дугах.

$LEV : P \times T \rightarrow \left\{ (lev_1, lev_2, \dots, lev_{N_{lev}}) \mid \sum_{lev_i} NF^{lev_i} = NF, j = \overline{1, N_{lev}} \right\}$ – функция, описывающая иерар-

хические уровни сети. Условия срабатывания переходов между уровнями сети отражают функции I_{lev} , определяющие какие метки будут перемещаться на другой уровень. Правила срабатывания для таких переходов те же, что и для обыкновенных сетей Петри, при этом после срабатывания перехода вложенная сеть (с новой разметкой) остается в прежней позиции.

M_0 – начальная разметка сети,

R_i – структура данных для определения риска (в данном случае не используется).

Изменения формализма позволяют моделировать функционирование и не только диагностировать отказы элементов ДК, но и прогнозировать их появление в реальном режиме времени.

Использование СФ в качестве средства создания модели позволяет при соблюдении определенных ограничений на структуры сети, построить граф называемый *деревом достижимых маркировок* (ДДМ), который полностью отражает картину функционирования модели (сети). Т.е. каждой вершине ДДМ ставится в соответствие достижимая маркировка $M \in M_0$, дуга между двумя вершинами M и M' существует, если M' может быть получена из M в результате срабатывания перехода T . Такая дуга помечается символом (идентификатором) перехода t .

Используя известные алгоритмы анализа ДДМ, модифицируя их для данного сетевого формализма, для модели ТС можно решать следующие задачи:

- 1) достижима ли какая-нибудь заданная маркировка;
- 2) допустима ли какая-либо заданная последовательность срабатываний;
- 3) какова должна быть начальная маркировка, чтобы заданная маркировка была достижима.

Декомпозиция ДК

Отражение специфики поведения отдельных, существенных компонентов модели производится декомпозицией их на более подробном уровне. Для этого технологические системы ДК разделены на 5 подсистем: 1) управление подсистемами; 2) датчики давления; 3) кнопочные пульты; 4) панели светодиодные (LED); 5) гидравлическая система.

Назначение систем:

Управление подсистемами – в соответствии с регламентом проведения измерений производится включение и отключение маслостанции, открытие и закрытие запорной арматуры (гидравлическая система), контроль работоспособности датчиков давления и измерение давления (датчики давления), выбор контролируемой оси и считывание данных (кнопочные пульты), контроль работоспособности и отображение результатов измерения и расчета (LED).

В моделях используются сигналы следующего оборудования: датчиков давления, датчиков уровня, концевых выключателей запорной арматуры, служебных сигналов системы управления светодиодными панелями, служебных сигналов системы управления маслонасосами.

Архитектура

Под архитектурой программной системы понимается в совокупности ее структура, используемые алгоритмы и структуры данных.

Структура программного комплекса представлена на рис. 1. Элементы структуры в дальнейшем называются модулями.

Программное обеспечение включает следующие модули:

- 1) модуль интерфейса пользователя;
- 2) создания, редактирования, модели ТС;
- 3) управляющий и измерительный модуль обмена данных с датчиками и OPC-сервером, другими устройствами не поддерживающими стандарт OPC;
- 4) построения ДО, ДС, ДДМ при возникновении отказа;
- 5) прогностический модуль;
- 6) модуль функционирования модели ТС.

Алгоритм работы ДК представляется следующим образом: набор аналоговых и дискретных сигналов управления от оборудования (датчиков давления, концевых выключателей, запорной арматуры, LED) поступает на вход контроллеров или напрямую на вход компьютера по протоколу RS232. Набор полученных сигналов в цифровой форме обрабатывается OPC-сервером и разделяется на группы кнопочного поста управления панелями LED и управления маслостанцией. Далее, сигналы управления подаются на входы подпрограмм модулей ДК. Синхронизация состояния модели осуществляется только в соответствии с параметрами датчиков и контроллеров, время в модели не используется. Также контролируется состояние запорной арматуры.

Структура программной системы программно - измерительного комплекса отображена на рисунке 1.

Модули выполняют следующие функции:



- 1) Управляющая программа выполняет следующие функции: ввод задания на моделирование, управление интерфейсом пользователя;
- 2) создание и редактирование позиций, переходов, меток, дуг, заполнение деклараций, анализ соответствия ТС и построенной модели;
- 3) связывание объектов (позиций и меток) модели с контроллерами и датчиками, OPC- сервером;
- 4) моделирование функционирования технических систем (ТС);
- 5) вывод прогностических уравнений по состоянию модели;
- 6) построение деревьев событий и деревьев отказов.

Рисунок 1 – Структура программной системы

Алгоритмы функционирования модели используют основное уравнение функционирования СП, динамический обмен данными между и ТС и моделью, ограничения на значения параметров меток. Структуры данных вследствие достаточно большого их описания не приводятся.

В соответствии с декомпозицией ТС созданы модели систем и реализован алгоритм их функционирования (рисунок 2, 3).

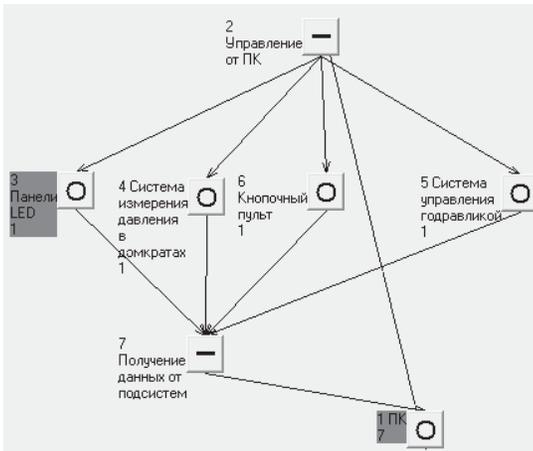


Рисунок 2 – Уровень 1. Взаимодействие подсистем

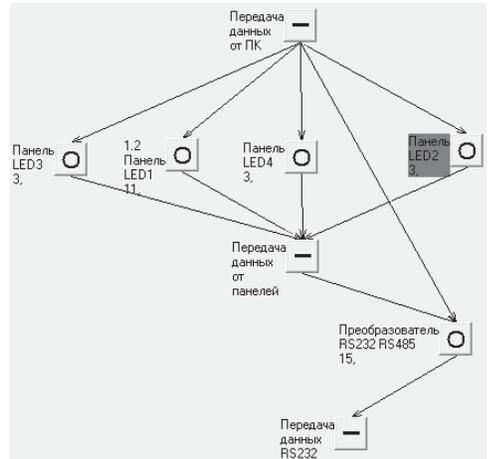


Рисунок 3 – Уровень 2. Панели LED

Каждый элемент (позиция) модели имеет входные и/или выходные сигналы с сервера ОРС, которые являются значениями параметров меток. Сопоставляя значения параметров и ограничения на них, определяется работоспособность системы. Например, при выполнении «перехода» – «Запрос на получение данных» метка поступает сразу на все датчики давления (ДД) и при выполнении «перехода» – «Получение данных с датчиков» каждая из меток, поступивших от ДД1-12 несет значение, которое и сравнивается с заложенными в модель ограничениями на значения.

Скриншот графического интерфейса пользователя подсистемы моделирования приведен на рисунке 4.

При данном расширении СФ переход сработает, если во входных позициях имеются метки такого цвета, который указан в функциях дуг и значение параметров этих меток принадлежат множеству AT_j – (выполнены условия I_m, I_f, I_l), а в выходных позициях разрешено получение меток цвета, который определяется функциями дуг и значение параметров этих меток принадлежат множеству AT_j – (выполнены условия O_m, O_f, O_l).

Моделирование события происходит следующим образом:

- на каждом шаге выполнения, метка $m_i^{prmi} = \{tm_k, (at_{k1}, at_{k2}, \dots, at_{N_k N_{ka}})^{his}\}$, цвета tm_k и находящаяся во входной позиции P_g по отношению к выполняемому переходу T_l получает новые значения атрибутов в результате опроса датчиков, и полученные значения присваиваются параметрам меток $(at_{k1}, at_{k2}, \dots, a_{N_k N_{ka}})$;
- массив значений атрибутов метки цвета tm_k изменяется:
 $(at_{k1}, at_{k2}, \dots, a_{N_k N_{ka}}) := (at_{k1}, at_{k2}, \dots, a_{N_k N_{ka}})$,
 $(at_{k1}, at_{k2}, \dots, a_{N_k N_{ka}})^{his+1} := (at_{k1}, at_{k2}, \dots, a_{N_k N_{ka}})^{his}$, $his = \overline{N_{his} - 1}$;
- при выполнении предусловия перехода T_l связанного с входной позицией P_g метка изымается и происходит коррекция коэффициентов уравнения $f_{kj}^{pro} = f^{pro}(\{at_{kj}\}^{his})$;

- при выполнении постусловий O_l тип и атрибуты метки изменяются в соответствии с функциями выходных дуг O_f ;
- в выходной позиции P_{g+1} , для выполняющегося перехода T_l , создается метка $m_{i+1}^{pr_{i+1}}$ с заданным цветом и атрибутами.

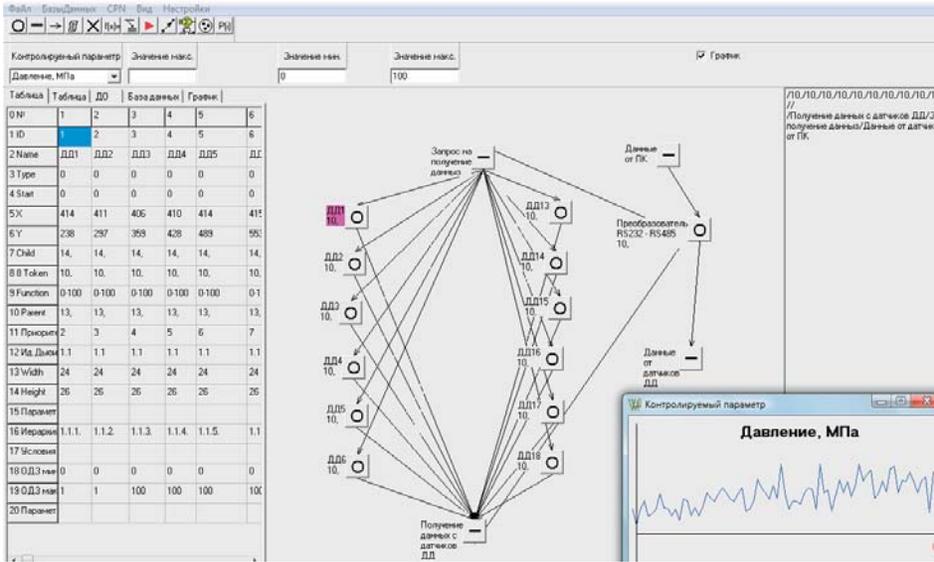


Рисунок 4 – Интерфейс пользователя подсистемы моделирования

Заключение

Результатами выполнения работы являются:

- создание имитационной модели с использованием расширенного формализма сетей Петри;
- разработан программно – измерительный комплекс для процесса определения нагрузки от колесных пар локомотив с системой самодиагностики отказов.

Данный программно-измерительный комплекс внедрен в ОАО «Нижеудинское ремонтное локомотивное депо» - филиала ОАО РЖД .

Список литературы

- [1] Питерсон Дж. Теория сетей Петри и моделирования систем: Пер. с англ. -М.: Мир, 1984. -264 с., ил.
- [2] Котов. Сети Петри. - М.: Наука,1984, 160 с.
- [3] В. Н. Дубинин, С. А. Зинкин. Проектирование вычислительных систем и сетей на основе сетевых формализмов. Книга 1: Сетевые технологии проектирования и реализации распределенных вычислительных систем на программно-аппаратных платформах локальных и глобальных сетей Ethernet/Internet. - Пенза: Изд-во Пенз. гос. ун-та, 1998. - 322 с.: 160 ил., 19 табл., библиогр. 63 назв.
- [4] Берман А.Ф., Николайчук О.А., Вильвер П.Ю. Моделирование функционирования сложных технологических комплексов на основе модифицированной сети Петри // Современные технологии. Системный анализ. Моделирование. – 2009. – №4. - С.23-29.

СИНТЕЗ ОПТИМАЛЬНОГО ПО БЫСТРОДЕЙСТВИЮ СЛЕДЯЩЕГО ЭЛЕКТРОПРИВОДА ПОСТОЯННОГО ТОКА

Н.В. Фалдин, А.Ю. Липатов

Тульский государственный университет
300600, Тула, пр. Ленина, 95, Россия
info@sau.tsu.tula.ru
тел: +7 (4872) 35-38-35

Ключевые слова: электропривод, оптимальное управление, быстродействие, следящий привод

Abstract

The question of the hydraulic actuator control's optimization by the time response criteria is considered. The resulting control function enables the work of the actuator in tracking mode.

Введение

Электроприводы постоянного тока находят широкое применение в технике. Существует большое количество работ, посвященных оптимизации таких приводов по быстродействию. Однако в указанных работах рассматривается только режим позиционирования, т.е. предполагается, что привод должен обрабатывать ступенчатые входные сигналы.

Закон управления, синтезированный на ступенчатые входные сигналы, не может рассматриваться в качестве оптимального для следящих систем. Далее, в электроприводах постоянного тока для защиты обмотки якоря двигателя, как правило, используются ограничители тока. Ограничители оказывают большое влияние на длительность переходного процесса и их необходимо учитывать при синтезе оптимального по быстродействию управления. В известных из литературы работах, посвященных оптимизации электроприводов по быстродействию, ограничители тока не рассматриваются.

В настоящей работе излагается процедура синтеза оптимального по быстродействию следящего электропривода постоянного тока по его нелинейной динамической модели, т.е. с учетом ограничителя тока. Именно в такой постановке задачи синтеза удастся получить оптимальный закон управления, который представляет наибольший интерес для практики.

1 Синтез базового закона управления

На рисунке 1 представлена структурная схема электропривода как объекта управления. Привод работает на инерционную нагрузку.

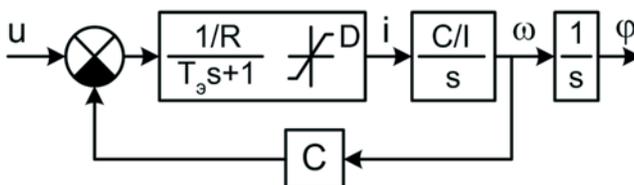


Рисунок 1 – Структурная схема базовой модели электропривода

Здесь: $R = 0.15$, $T_{\Sigma} = 0.0015$ с, $C = 0.052$, $I = 0.00926$, $D = 120$ А. На напряжение u наложено ограничение $|u| \leq A$, где $A = 27$.

Постоянная времени T_3 является малой величиной. Это позволило использовать базовый метод синтеза [1]. Базовой называется система, которая получается из исходной, если в ней приравнять нулю малые постоянные времени. Таким образом, в изображенной на рисунке 1 структурной схеме следует положить $T_3 = 0$. Ограничитель тока при этом окажется безынерционным.

Синтез оптимальной системы выполняется по следующей схеме. Сначала для базовой системы синтезируется ее оптимальный по быстродействию закон управления, а затем в полученном законе приближенно учитывается динамика звена с малой постоянной времени. Таким образом удастся получить закон управления, который оказывается близким к строго оптимальному.

Чтобы упростить выделение оптимального управления и оптимальных траекторий, ограничение тока в базовой системе будем обеспечивать соответствующим выбором управления. Это приведет к переменной области управления, вид которой представлен на рисунке 2.

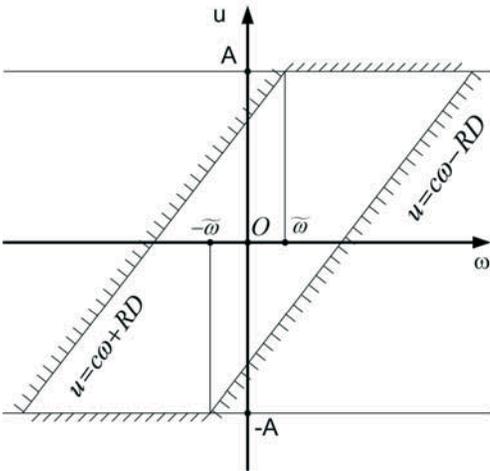


Рисунок 2 – Переменная область управления

Выпишем неравенства, задающие переменную область управления:

$$(1) \begin{cases} -u - A \leq 0, & -u + C\omega - RD \leq 0; \\ u - A \leq 0, & u - C\omega - RD \leq 0. \end{cases}$$

Рассмотрим сначала автономный электропривод, ориентированный на отработку начальных рассогласований.

Для нахождения оптимального управления и оптимальных траекторий воспользуемся необходимыми условиями оптимальности в форме принципа максимума, как они формулируются для систем с переменной областью управления [2,3]. Поскольку граница области управления является кусочно-гладкой, то в точках излома границы необходимо также учитывать условие скачка. На рисунке 3 изображен вид оптимальных по быстродействию траекторий для автономной базовой системы. Здесь $u_1(t) = C\omega + RD$, $u_2(t) = C\omega - RD$. Траектория MON является линией переключения.

Покажем, что траектория $LSEO$ удовлетворяет всем необходимым условиям оптимальности. Пусть точке S соответствует момент времени t_1 , точке E – момент t_2 , а точке O – момент t_3 .

Как следует из рисунка 1, движение базовой системы задается уравнениями

$$(2) \frac{d\omega}{dt} = k(u - C\omega), \quad \frac{d\phi}{dt} = \omega,$$

где $k = C/RI$. Запишем функцию Гамильтона $H(\bar{\psi}, \omega, \phi, u) = \psi_1 k(u - C\omega) + \psi_2 \omega$.

В точке S (момент t_1), которая является угловой точкой границы области управления, должно выполняться условие скачка [3]:

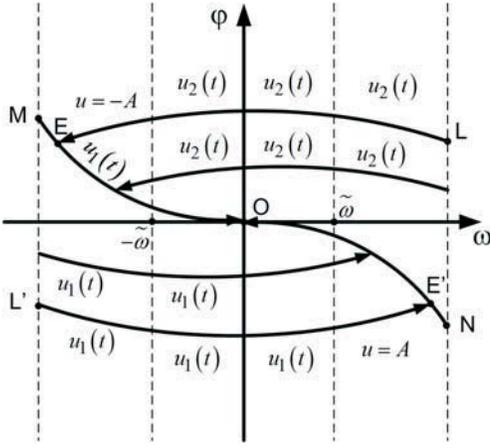


Рисунок 3 – Оптимальные по быстродействию траектории

$$\psi_1(t_1 + 0) = \psi_1(t_1 - 0) + \mu,$$

$$\psi_2(t_1 + 0) = \psi_2(t_1 - 0),$$

где μ – некоторое число. Из условия непрерывности функции Гамильтона найдем, что $\mu = 0$. Оптимальное управление $u(t)$ в каждый момент времени t доставляет функции Гамильтона максимум. В соответствии с рисунком 3 в интервале $0 \leq t < t_2$ должно иметь место неравенство $\psi_1(t) < 0$, а в интервале $t_2 < t \leq t_3$ – неравенство $\psi_1(t) > 0$. Так как $\psi_1(t)$ – непрерывная функция, то

$$(3) \quad \psi_1(t_2) = 0$$

В соответствии с принципом максимума, как нетрудно установить, в интервалах $0 \leq t < t_1$ и $t_2 < t \leq t_3$ вспомогательные переменные $\psi_1(t)$ и $\psi_2(t)$ задаются уравнениями

$$(4) \quad \frac{d\psi_1}{dt} = -\psi_2, \quad \frac{d\psi_2}{dt} = 0,$$

а в интервале $t_1 < t < t_2$ – уравнениями

$$(5) \quad \frac{d\psi_1}{dt} = kC\psi_1 - \psi_2, \quad \frac{d\psi_2}{dt} = 0.$$

Если задано начальное условие $\vec{\psi}(0)$, то уравнения (4) и (5) полностью определяют вспомогательный вектор $\vec{\psi}(t)$.

Принципом максимума вспомогательный вектор $\vec{\psi}(t)$ определяется с точностью до постоянного положительного множителя. Положим $\psi_1(0) = -1$. Начальное значение $\psi_2(0)$ легко находится с помощью уравнений (4) и (5) и равенства (3):

$$(6) \quad \psi_2(0) = \frac{BkC}{1 - kB - kCB}, \quad B = \frac{e^{kCt_2}}{e^{2kCt_1}}.$$

Рассчитав соответствующий полученным начальным условиям $\psi_1(0)$ и $\psi_2(0)$ вспомогательный вектор $\vec{\psi}(t)$, легко установить, что траектория *LSEO* удовлетворяет всем необходимым условиям оптимальности. Аналогичным образом обосновывается оптимальность остальных траекторий.

Рассмотрим оптимизацию следящего электропривода. В следящих системах входные сигналы заранее неизвестны. Будем в качестве входных сигналов рассматривать произвольные линейно нарастающие (убывающие) функции. Это, с одной стороны, позволяет существенно упростить синтез оптимальной системы, а, с другой стороны, как показывает выполненное исследование, полученный закон управления обеспечивает хорошее быстродействие и при других возможных входных сигналах.

Пусть входной сигнал $y = g_1 t + g_0$, где g_1 и g_0 – любые числа. Введем ошибки $\Delta\varphi = g_1 t + g_0 - \varphi$, $\Delta\omega = g_1 - \omega$. В пространстве ошибок переменная область управления U , как следует из (1), задается неравенствами

$$(7) \quad -u - A \leq 0, \quad -u - C\Delta\omega + Cg_1 - RD \leq 0,$$

$$(8) \quad u - A \leq 0, \quad u + C\Delta\omega - Cg_1 - RD \leq 0.$$

Уравнения движения (2) принимают вид

$$(9) \quad \frac{d\Delta\omega}{dt} = -ku + kCg_1 - kC\Delta\omega, \quad \frac{d\Delta\varphi}{dt} = \Delta\omega.$$

Как и в рассмотренном выше автономном случае, оптимальное управление принимает граничные значения области (7)-(8). Входящий в неравенства (7)-(8) параметр g_1 , изменяет

расположение угловых точек на границе области U . Будем расположить угловые точки задавать значением ошибки $\Delta\omega$. Расположение угловых точек находится из неравенств (7)–(8).

Если система вышла на режим слежения, то $\Delta\varphi(t) = 0$, $\Delta\omega(t) = 0$. Таким образом, необходимо найти управление, которое за минимально возможное время переводит фазовую точку системы (9) в начало координат.

Если в процессе движения управление изменяет свой знак, то соответствующий момент времени будем называть точкой переключения. Аналогично тому, как это было доказано для автономной системы (рисунок 3), оптимальное управление для системы (9) может иметь на траектории только одну точку переключения, т.е. при любом фиксированном g_1 оптимальный закон управления задается линией переключения. Нетрудно установить, что при $\Delta\omega < 0$ линия переключения представляет собой идущую в начало координат траекторию системы (9), на которой $u(t)$ задается равенством (7), а при $\Delta\omega > 0$ – линию, на которой $u(t)$ определяется равенством (8).

При расчете линий переключения целесообразно воспользоваться принципом попятного движения Фельдбаума [4]. На уравнения, задающие линии переключения, существенное влияние оказывают угловые точки границы области U . Далее, необходимо также учитывать, что максимальная скорость привода $\omega_{max} = A/C = 519.2$.

Ниже приводятся уравнения, задающие линию переключения:

1) При $|g_1| < \bar{g}_1$:

$$(10) \quad \Delta\varphi = -\frac{(\Delta\omega)^2}{2bD} \operatorname{sgn}(\Delta\omega).$$

2) При $|g_1| > \bar{g}_1$:

$$(11) \quad \Delta\varphi = \begin{cases} \frac{1}{kC^2}(A - C|g_1|) \ln\left(\frac{C|\Delta\omega| + A - C|g_1|}{A - C|g_1|}\right) \operatorname{sgn}(\Delta\omega) - \frac{1}{kC} \Delta\omega, & \text{если } |\Delta\omega| < \Delta\bar{\omega} \text{ и } \Delta\omega g_1 > 0; \\ -\left[\Delta\bar{\omega} + \frac{1}{2bD} \left[(\Delta\omega)^2 - (\Delta\omega)^2 \right] \right] \operatorname{sgn}(\Delta\omega), & \text{если } |\Delta\omega| > \Delta\bar{\omega} \text{ и } \Delta\omega g_1 > 0; \\ -\frac{(\Delta\omega)^2}{2bD} \operatorname{sgn}(\Delta\omega), & \text{если } \Delta\omega g_1 < 0. \end{cases}$$

Здесь $\bar{g}_1 = \frac{A - RD}{C} = 173$, $b = C/I = 5.616$, $\Delta\bar{\omega} = \frac{-A + C|g_1| + RD}{C}$,

$$\Delta\bar{\omega} = \frac{1}{kC^2} \left[(A - C|g_1|) \ln\left(\frac{C\Delta\bar{\omega} + A - C|g_1|}{A - C|g_1|}\right) - C\Delta\bar{\omega} \right].$$

Обозначим

$$(12) \quad \Delta\varphi = f(\Delta\omega, g_1)$$

уравнение переключения. Выше, чтобы упростить обоснование процедуры синтеза оптимального по быстрдействию управления, ограничение тока обеспечивалось соответствующим выбором управления. Принимая во внимание, что в базовой системе присутствует физический ограничитель тока, то оптимальный закон управления можно записать в виде

$$(13) \quad u = A \operatorname{sgn}(\Delta\varphi - f(\Delta\omega, g_1)).$$

На рисунках 4-5 изображены полученные с помощью компьютерного моделирования осциллограммы отработки оптимальной базовой системой входных сигналов $y(t) = 150 \cdot 1(t)$ и $y(t) = -200t + 200$.

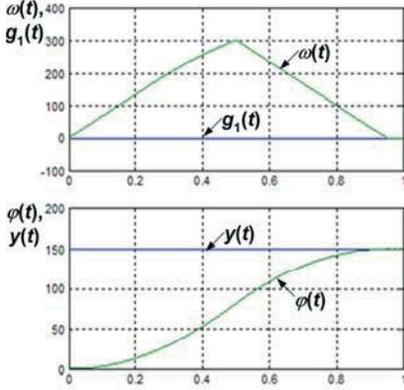


Рисунок 4 – Осциллограммы отработки базовой системой начального рассогласования

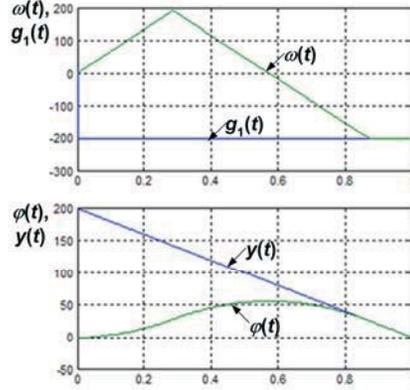


Рисунок 5 – Осциллограммы отработки базовой системой линейно убывающего сигнала

2 Оптимальный закон управления электроприводом

Учтем в базовом законе управления (13) динамику звена с малой постоянной времени T_3 . Для этого воспользуемся способом, предложенным в [1]. Указанный способ основан на предположении о том, что звено с малой постоянной времени вносит в систему некоторое запаздывание. Для компенсации запаздывания заменим в законе управления (13) фазовые переменные на упрежденные на величину "эквивалентного" запаздывания τ . Закон управления примет вид

$$(14) \quad u = A \operatorname{sgn} [\Delta \varphi + \Delta \omega \tau - f(\Delta \omega - b i \tau)].$$

Величина запаздывания τ подбиралась путем компьютерного моделирования электропривода с законом управления (14) и оказалась равной $0.8T_3$.

На рисунках 6-8 представлены осциллограммы отработки синтезированным приводом входных сигналов $y(t) = 150 \cdot 1(t)$, $y(t) = -200t + 200$ и $y(t) = 120 \sin(1.26t)$.

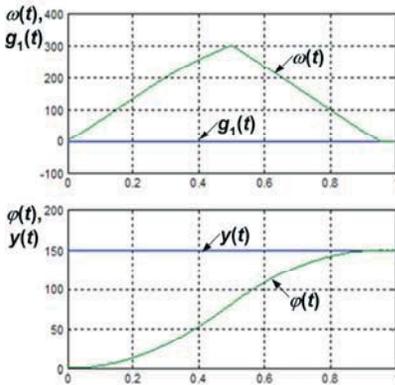


Рисунок 6 – Осциллограммы отработки электроприводом начального рассогласования

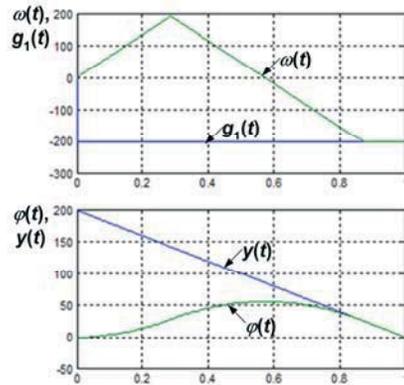


Рисунок 7 – Осциллограммы отработки электроприводом линейно убывающего сигнала

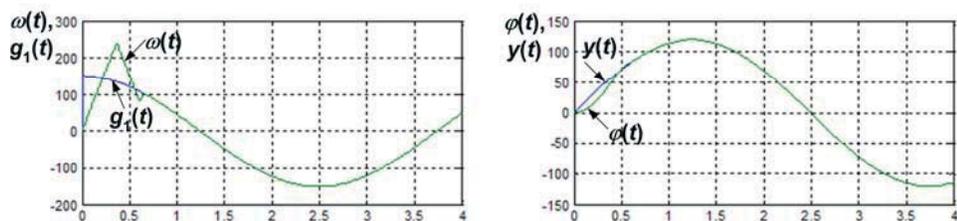


Рисунок 8 – Осциллограммы обработки электроприводом гармонического входного сигнала

Из сравнения осциллограмм, представленных на рисунках 6-7 с осциллограммами на рисунках 4-5 следует, что закон управления (14) можно вполне рассматривать как оптимальный по быстродействию. Осциллограммы, приведенные на рисунке 8, показывают, что закон управления (14) обеспечивает хорошее качество переходных процессов и высокую точность режима слежения (ошибка не превышает 0.04 мрад) и при гармонических входных сигналах.

Заключение

В настоящей работе фактически излагается методика синтеза оптимальных по быстродействию электроприводов постоянного тока. Очень важно, что оптимальный закон управления в аналитическом виде. Это делает его пригодным для целого класса приводов.

Благодарности

Работа выполнена при финансовой поддержке по гранту РФФИ (грант № 09-08-00332).

Список литературы

- [1] Пупков К.А., Фалдин Н.В., Егупов Н.Д. Методы синтеза оптимальных систем автоматического управления. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2000. – 510 с.
- [2] Болтянский В.Г. Математические методы оптимального управления. – М.: Наука, 1969. – 407 с.
- [3] Иванов В.А., Фалдин Н.В. Теория оптимальных систем автоматического управления. – М.: Наука, 1981. – 331 с.
- [4] Фельдбаум А.А. Основы теории оптимальных автоматических систем. – М.: Физматгиз, 1963. – 552 с.

СТРУЙНЫЙ ИЗМЕРИТЕЛЬ КИНЕМАТИЧЕСКИХ ПАРАМЕТРОВ СИСТЕМЫ УПРАВЛЕНИЯ ПОДВИЖНЫМ ОБЪЕКТОМ

В.А. Яшков

Казанский государственный технический университет им. А. Н. Туполева
420111, Казань, ул. К.Маркса, 10, Россия
yavitalic@yandex.ru
тел: +7 (843) 231-03-90

***Ключевые слова:** струйный измеритель, система управления, угловая скорость, ускорение, подвижный объект, кинематика движения*

Abstract

Development and current state of methods and means of measurement of kinematic parameters of mobile objects of various classes is investigated. Designs of jet measuring instruments of the kinematic parameters constructed on a combination of jet, thermal and acoustic effects are described.

Введение

Развитие современной авиационной и ракетной техники характеризуется значительным ростом скорости, высоты полета и маневренности летательного аппарата (ЛА). Как показывает опыт эксплуатации отечественной и зарубежной авиационной техники, эффективность использования информации о параметрах подвижных объектов (ускорение, угловая скорость) в ряде случаев ограничивается применением гироскопических датчиков угловой скорости (ДУС) и акселерометров, которые обладают недостаточной вибропрочностью, виброустойчивостью, значительным и непостоянным по величине моментом трения (для устройств с механическим подвесом) в подшипниках рамок подвеса, сложностью осуществления и поддержания постоянным демпфирования. Гироскопическим ДУС свойственны существенные погрешности аддитивного характера, среди основных источников которых известны: кориолисово ускорение, дебаланс массы, вариации температуры, магнитные поля, влияния гибких проводников (видов), обратная реакция катушек чувствительных элементов, моментных датчиков и др. Для гироскопических акселерометров наиболее существенны также погрешности от заданной динамической реакции в виде резонансных явлений, а также значительные температурные погрешности. Наличие достоверной информации по линейному и угловому ускорению позволяет наиболее полно использовать лётно - технические возможности ЛА, обеспечить безопасность и эффективность полетов [1].

В связи с этим в последнее время стали развиваться струйные измерители параметров подвижных объектов (ПО) и системы, построенные на их основе, где в качестве инерционной массы используется струя жидкости или газа. Это позволило повысить надёжность и долговечность, уменьшить стоимость, вес и габариты из-за отсутствия подвижных механических частей. Такие измерители не чувствительны к радиационным воздействиям, электрическим и магнитным полям, что приводит к упрощению конструкции. Характеристики струйных измерителей не зависят от условий внешней среды.

Основная часть

Одними из наиболее информативно ценными для управления и контроля ПО и, в том числе транспортных средств (ТС), вне зависимости от области их применения и решаемых

задач, являются измерители угловой скорости и ускорения. Современные ТС в зависимости от характера взаимодействия с окружающей средой, относительно которой проводится их перемещение, а также с учетом характера подстилающей поверхности, принято делить на космические, воздушные, водные (надводные и подводные) и наземные. Виды ТС и спектр их возможного применения показан на рисунке 1. В настоящее время известно более 100 различных физических принципов, которые можно положить в основу создания систем измерения и автономных измерителей кинематических параметров (ИКП) ТС относительно инерциального пространства. Реализация многих из них связана со схемной, конструкторской, технологической и эксплуатационной сложностью. Однако, по мере появления новых материалов, элементов совершенствования технологии возникает реальная возможность создания ИКП и вариантов датчиков первичных сигналов КП на новых физических принципах с последующей доводкой их до уровня серийности. Эта необходимость связана с требованиями к уменьшению габаритов, повышению точности, снижению себестоимости и увеличению надежности. Одними из наиболее информативно ценных для управления и контроля ПО и, в том числе, ТС, вне зависимости от области их применения и решаемых задач являются, измерители угловой скорости и ускорения [2].

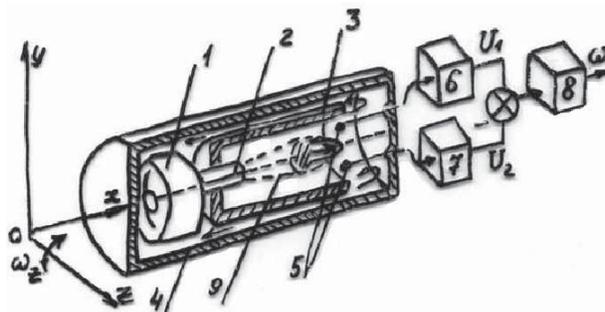


Рисунок 1 – Спектр подвижных объектов (транспортных средств) и задачи информационного обеспечения, управления и контроля

Важнейшим параметром, определяющим динамические характеристики ПО, является его угловая скорость. Датчики угловой скорости используются в качестве элементов систем автоматического управления и стабилизации ПО. В системах стабилизации длительного действия используют гироскопические ДУС, которые не всегда удовлетворяют возрастающим требованиям по надежности (вероятность безотказной работы за 12 часов более 0,999), сроку службы (10 000...30 000 час), себестоимости, технологичности и экономичности. Этим обусловлена высокая актуальность создания датчиков угловой скорости, основанных на новых физических принципах преобразования сигнала, применение которых позволяет получить качественно новые характеристики системы автоматического управления (малый порог реагирования, высокие динамические характеристики, большой ресурс работы и т.п.).

В качестве измерителя струйной системы управления параметрами подвижных объектов рассматривается струйный ДУС, основной целью которого является повышение точности из-

мерения (рисунок 2). При отсутствии угловой скорости струя газа расположена симметрично относительно оси прибора и анемочувствительные элементы 5 обдуваются участками струи с одинаковой скоростью. В этом случае сигналы U_1 и U_2 на выходах электрических цепей 6 и 7 одинаковы и разностный электрический сигнал, пропорциональный угловой скорости, равен нулю. При вращении датчика с угловой скоростью ω происходит отклонение струи газа относительно оси симметрии прибора. При этом скорости обдува анемочувствительных элементов становятся различными, что приводит к изменению сигналов U_1 и U_2 , разность которых на выходе дифференциального усилителя 8 образует сигнал, пропорциональный угловой скорости. Источник пневмопитания содержит пьезокерамический преобразователь 2, резонансную акустическую камеру 3, образованную пьезопреобразователем и стенкой 4 с центральным отверстием – соплом, и приёмный канал 5. Одним из перспективных путей уменьшения дрейфа пневматического нуля, характерного для большинства аналогов, является использование принципов струйной техники и широтно – импульсной модуляции (ШИМ) пневматического сигнала[3].



1 – источник пневмопитания, 2 – формирующее сопло, 3 – рабочая камера,
4 – герметичный корпус, 5 – анемочувствительные элементы,
6,7 – дифференциальные электрические цепи, 8 – усилитель, 9 – струя

Рисунок 2 – Структурно-функциональная схема струйного ДУС

Решение задач инженерного проектирования и расчета струйного ДУС в значительной степени затрудняется в связи с отсутствием теоретических исследований по методам их структурного и параметрического синтеза. Это обусловлено тем, что процессы измерительного преобразования в струйных ДУС имеют сложное математическое описание, исключающее прямое применение аналитических методов анализа и расчета.

Наибольшее влияние на суммарную погрешность в данной схеме будут оказывать нагнетатель (влияет на стабильность работы) и термоанемочувствительные элементы (вводят функциональную погрешность). Мультипликативная погрешность датчика будет определяться выражением[4]:

$$(1) \quad \eta_i = \left| 1 - \left(\frac{v_{\bar{n}}}{v_0} \right)^{0,4} \frac{R_{cro} [\exp(Kx/d_c Re_c^{0,9}) - 1] (T - T_c) \lambda_c Pr_c^{0,33}}{R_{cr} [\exp(Kx/d_c Re_c^{0,9}) - 1] (T - T_0) \lambda_0 Pr_0^{0,33}} \right|,$$

где ν – коэффициент кинематической вязкости; R_{cr} – радиус струи газа; d_c – диаметр сопла; λ – длина волны; Pr – критерий Прандтля; параметры с индексом «с» соответствуют температуре T_c – температура окружающей среды, а с индексом «0» - температуре T_0 ; T – температура тела термоанеморезистора. Дополнительная погрешность η_ϕ фазового сдвига, обусловленная изменением температуры среды, определяется соотношением[4]:

$$(2) \quad \eta_\phi = \left| 1 - \left(\frac{v_c}{v_0} \right)^{0,9} \frac{1 - \exp(-Kx/d_c Re_c^{0,9})}{1 - \exp(-Kx/d_c Re_c^{0,9})} \right|.$$

Мультипликативная погрешность δ_{ω_y} датчика угловой скорости при действии на него перекрёстной угловой скорости ω_y определяется выражением[4]:

$$(3) \quad \delta_{\omega_y} = 1 - \frac{y_z}{y} \sqrt{\frac{1 - (l - y)^2 / R_{cr}^2}{1 - (l - y_z)^2 / (R_{cr}^2 - z^2)}},$$

где y_z – отклонение струи в рабочей камере при действии ω_y .

В случае воздействия на преобразователь ω_x и ω_y получили выражение, определяющее перекрёстную погрешность[4]:

$$(4) \quad \Delta y_{\omega_x} = - \frac{2x^3 K \omega_z \sin(\omega_x t + \varphi_H)}{v_0 d_c \text{Re}^{0,9} [1 - \exp(-Kx / d_c \text{Re}^{0,9})]},$$

где φ_H – начальная фаза угловой скорости ω_x , t – время действия угловой скорости.

Оценка влияния перекрёстных угловых скоростей проведена при задании по измерительной оси прибора и по перпендикулярной ей оси угловых скоростей равных диапазону измерения (0,1 рад/с). В результате исследования установлено, что величина изменения выходного сигнала составляет $\pm 5\%$ от диапазона измерения.

Таким образом, перекрёстная угловая скорость в плоскости измерения влияет незначительно. Вопрос устойчивости работы нагнетателя при воздействии вибрации решается путем увеличения частоты его возбуждения.

На основе экспериментальных исследований оценено влияние на метрологические характеристики конструктивных параметров основных элементов струйного ДУС и параметров режимов его работы (рисунок 3).

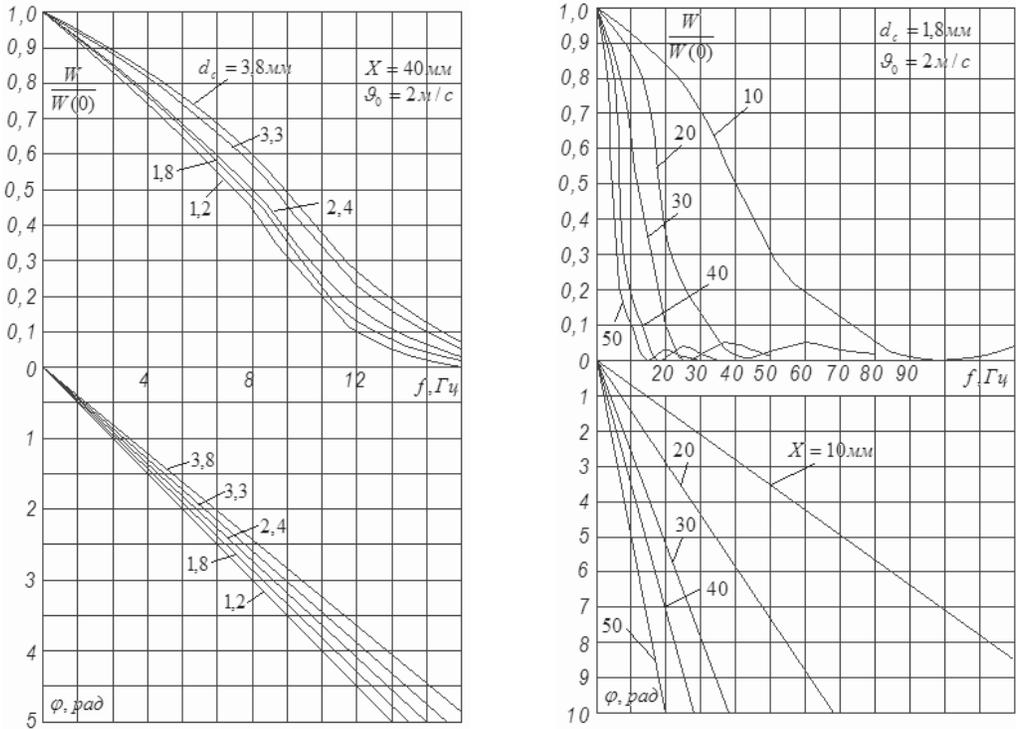


Рисунок 3 – Частотные характеристики струйного измерителя для различных диаметров сопла и расстояний от среза сопла до термоанеморезисторов

Из проведённого выше анализа динамических характеристик струи воздуха известно, что при воздействии угловой скорости на струю газа, истекающую из формирующего сопла, центр струи отклоняется на расстояние y относительно оси симметрии измерителя, которое определяется выражением [5]:

$$(5) \quad y = -\frac{2x^3 k \omega_0}{V_0 d_c \text{Re}^{0,9} [1 - \exp(-kx / d_c \text{Re}^{0,9})]}.$$

Передаточная функция струйного канала представляется в виде $W_{CK}(p) = \sqrt{0,98\pi^2 f^2 - 0,16}$.

Оценка динамических характеристик струйного датчика угловой скорости производилась по амплитудно-частотным и фазо-частотным характеристикам приборов с диапазоном измерения 0,1 рад/с и 1,6 рад/с. Амплитудно-частотные и фазо-частотные характеристики струйного датчика угловой скорости показаны на рисунке 4, где $A = U_{\text{ВЫХ}} / U_0$, при этом $U_{\text{ВЫХ}}$ – выходное напряжение прибора при $f \neq 0$, U_0 – при частоте $f = 0$. Полоса пропускания приборов без амплитудного искажения выходного сигнала для диапазона 0,1 рад/с равна 6 Гц, для 1,6 составляет 15 Гц, а по уровню 0,7 соответственно 11 Гц и 35 Гц. Сдвиг по фазе на частоте 5 Гц равен 0,52 рад и 0,31 рад, соответственно, для диапазонов измерения 0,1 рад/с и 1,6 рад/с.

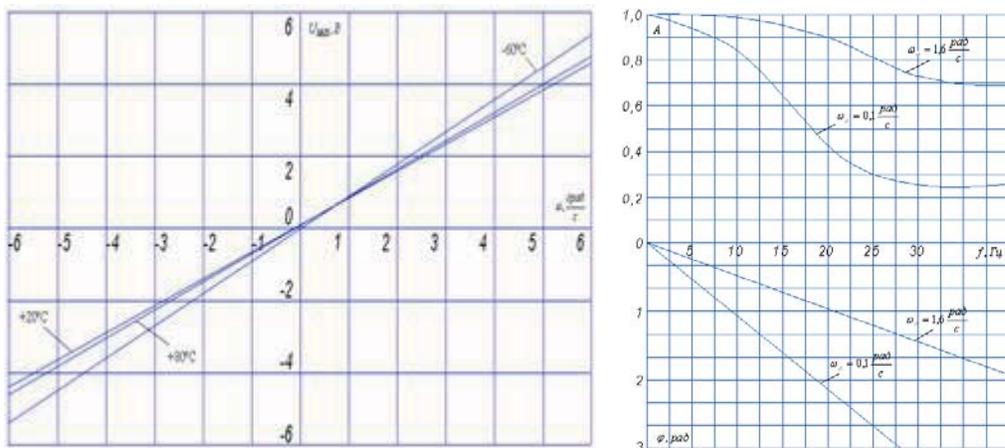


Рисунок 4 – Статические (с компенсацией нулевого сигнала) и динамические характеристики струйного измерителя

На рисунке 5 показано изменение нулевого сигнала прибора без компенсатора (кривая 1) и с компенсатором (кривая 2).

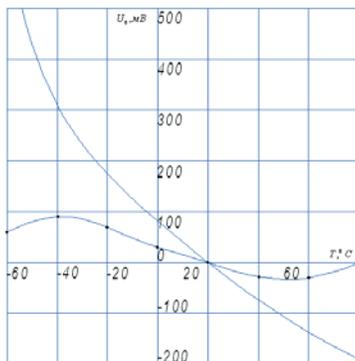


Рисунок 5 – Изменение нулевого сигнала

Основными влияющими факторами, с которыми приходится иметь дело при проектировании, испытаниях и эксплуатации струйного измерителя КП ПО, являются температура, вибрации, ускорения, уровень акустического давления, влажность, давление, колебания питающих напряжений, электромагнитные поля, агрессивная среда, радиация, воздействие грибок, тропиков и т.д. Многие факторы уменьшают своё воздействие, если предусмотреть при производстве прибора ряд конструктивных и технологических мероприятий. Кроме того, исследуемый струйный измеритель позволяет исключить влияние значительной части факторов, т.к. имеет герметичный объём, позволяющий заполнять прибор качественным газом (без пыли, влаги, определённого состава и т.д.).

В результате проведенных экспериментов была получена передаточная функция струйного измерителя кинематических параметров[5]:

$$(6) \quad W_{СК}(p) = \frac{2V_0(x)V_{cp}}{R_{сг}} \left[\frac{1 - \exp(p\tau_3)}{p} \right]^2,$$

где V_0 – максимальная скорость истечения струи газа из формирующего сопла; l – длина рабочей камеры; V_{cp} – средняя скорость потока, омывающего термоанемомрезисторы; $R_{сг}$ – радиус струи газа; τ_3 – время, определяющее инерционные свойства струи.

Заключение

Дальнейшее совершенствование струйного измерителя кинематических параметров системы управления подвижных объектах связано с оптимизацией его конструктивных параметров, миниатюризацией, интеграцией и повышением серийноспособности при одновременном снижении стоимости и использовании структурных методов повышения точности.

Список литературы

- [1] Яшков В.А. Инерциальные системы на основе струйных акселерометров // Туpoleвские чтения: Труды XVI международной молодежной научной конференции (28-29 мая 2008 г, Казань, Россия). - Казань: КГТУ, 2008. - С. 387-389.
- [2] Яшков В.А., Корнилов Л.А., Галяутдинов Б.Н. Принципы и схемы построения струйных измерительных устройств для определения параметров подвижных объектов // Решетневские чтения: Труды XII международной научной конференции (10-12 ноября 2008 г, Красноярск, Россия). – Красноярск: ДВГТУ, 2008. - С.196-197.
- [3] Яшков В.А. Разработка и исследование базовых элементов струйно-конвективной системы измерения параметров движения летательного аппарата // Приоритетные направления современной российской науки глазами молодых ученых: Труды всероссийской научно-практической конференции (4-6 ноября 2009 г, Рязань, Россия). – Рязань: РГУ, 2009. - С.117-121
- [4] Яшков В.А. Эволюция струйных измерителей кинематических параметров подвижных объектов и этапы коммерциализации // Наука. Технологии. Инновации-2009: Труды всероссийской научно-технической конференции (4-5 декабря 2009 г, Новосибирск, Россия). – Новосибирск: НГТУ, 2009. - С. 38-40.
- [5] Яшков В.А. Принципы построения струйного датчика кинематических параметров для систем управления летательным аппаратом // Молодежь, техника, космос: Труды общероссийской молодежной научно-технической конференции (17-19 марта 2010 г, Санкт-Петербург, Россия). - Санкт-Петербург: БГТУ “Военмех”, 2010. – С. 121-122.

КАЛИБРОВКА АКУСТИЧЕСКИХ АНТЕННЫХ РЕШЕТОК

Г.К. Чигарьков, А.В. Кочемасов

Государственное образовательное учреждение высшего профессионального образования
«Московский государственный институт радиотехники, электроники и автоматики
(технический университет)»
119454 Москва, просп. Вернадского д. 78, Россия
kotchemassov@mirea.ru
тел: +7 (495) 434-9685

Ключевые слова: антенная решетка, фазовая неидентичность, лучевой фазовый набег, фазовая коррекция, калибровка

Abstract

In this paper we propose a method for calibration of multi-element acoustic antenna arrays, based on an assessment of the coordinates of the emitter test signal with Calibration of the array. The effectiveness of the method of confirming the simulations.

Введение

Акустические антенные решетки являются основным средством излучения и приема акустических сигналов в системах различного назначения. С их помощью выполняется пространственная обработка сигналов. При изготовлении антенных решеток возникает технологический разброс амплитудно-фазовых параметров её элементов. Неидентичность элементов решетки существенно снижает угловую разрешающую способность. Если амплитудную неидентичность устранить сравнительно легко с помощью элементарных измерений уровня сигналов на решетке, то с фазовой неидентичностью возникают определенные проблемы. Это связано с тем, что при экспериментах измерению доступен суммарный фазовый сдвиг: фазовый сдвиг, обусловленный прохождением сигналов от излучающего устройства до каждого элемента антенной решетки (лучевой фазовый набег), и фазовый сдвиг в тракте приема (аппаратный фазовый сдвиг). Аппаратный фазовый сдвиг включает фазовую неидентичность. Таким образом, для измерения и компенсации фазовой неидентичности необходимо из суммарного фазового сдвига исключить лучевые фазовые набег. Расчет лучевых фазовых набегов можно выполнить при точном позиционировании антенной решетки и излучающего устройства. Однако часто необходимые для этого средства измерения отсутствуют. В данном докладе предлагается для позиционирования источника тестового сигнала относительно антенной решетки использовать саму решетку.

1 Математическая модель

При проведении эксперимента на определенном расстоянии от антенной решетки примерно на нормали к ней располагается излучающее устройство, которое излучает тональный импульс. Принимаемые антенной решеткой сигналы регистрируются и затем подвергаются обработке. При первичной обработке оценивается комплексная амплитуда сигнала в стробе, соответствующем приходу импульса. В результате формируется следующий массив сигналов:

$$(1) \quad u_k = A_k \cdot e^{-j\varphi_k}, \quad k = 1, 2, \dots, K,$$

где A_k – амплитуда импульса на k -м элементе решетки, φ_k – фазовый сдвиг сигнала k -го элемента решетки, измеренный относительно сигнала произвольного опорного сигнала.

Коррекцию амплитудной неидентичности можно выполнить сразу путем нормировки уровня всех принятых сигналов, например, по среднему значению амплитуды на решетке. Для этого необходимо в каждом канале установить масштабирующие коэффициенты, устраняющие амплитудный разброс

$$(2) \quad |\mu_k| = \frac{1}{A_k K} \sum_{n=1}^K A_n.$$

При необходимости можно также учесть и различное ослабление сигналов при распространении в среде до элементов решетки. С фазовой неидентичностью дело обстоит сложнее, так как необходимо разделять измеряемый суммарный фазовый сдвиг сигналов на лучевые фазовые набег и аппаратные фазовые сдвиги, для чего требуется позиционирование источника тестового сигнала. Таким образом, задача калибровки антенной решетки включает в себя задачу оценки координат источника тестового сигнала. Рассмотрим возможности оценки координат излучателя по фазовому распределению на апертуре плоской антенной решетки.

Пусть имеется плоская антенная решетка и излучающее устройство (рисунок 1).

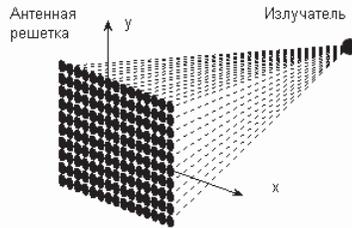


Рисунок 1 – Постановка эксперимента по фазовой коррекции антенной решетки

Декартовы координаты излучающего устройства связаны с его полярными координатами следующими соотношениями:

$$(3) \quad \begin{cases} x = r \cos \beta \sin \alpha, \\ y = r \sin \beta, \\ z = r \cos \beta \cos \alpha. \end{cases}$$

Координаты элементов решетки равны:

$$(4) \quad \begin{cases} x_{nk} = (k-1)d, & k = 1, 2, \dots, N_x, \\ y_{nk} = (n-1)d, & n = 1, 2, \dots, N_y, \\ z_{nk} = 0, \end{cases}$$

где k – номер столбца, n – номер строки антенной решетки. Лучевые фазовые набег при этом равны:

$$(5) \quad \varphi_{nk}^0 = 2\pi \rho_{nk}(r, \alpha, \beta) / \lambda, \quad k=1, 2, \dots, N_x, n=1, 2, \dots, N_y,$$

где

$$(6) \quad \rho_{nk}(r, \alpha, \beta) = \sqrt{(x - x_{nk})^2 + (y - y_{nk})^2 + z_{nk}^2},$$

длина луча для элемента решетки, расположенного в k -ом столбце и в n -ой строке. Расчетные сигналы имеют следующий вид:

$$(7) \quad v_{nk}^0 = A_0 e^{-j\varphi_{nk}^0}, \quad k = 1, 2, \dots, N_x, n = 1, 2, \dots, N_y.$$

Реальные сигналы будут отличаться от расчетных вследствие неидентичности и иметь определенные амплитуды и фазы:

$$(8) \quad v_{nk} = A_{nk} e^{-j\varphi_{nk}}, k = 1, 2, \dots, N_x, n = 1, 2, \dots, N_y$$

Фазу сигнала целесообразно измерять относительно гипотетического опорного сигнала с осредненной фазой:

$$(9) \quad \begin{cases} \Delta\varphi_{nk}^0 = \varphi_{nk}^0 - \frac{1}{N_x N_y} \sum_{\kappa=1}^{N_x} \sum_{\nu=1}^{N_y} \varphi_{\kappa\nu}^0, \\ \Delta\varphi_{nk} = \varphi_{nk} - \frac{1}{N_x N_y} \sum_{\kappa=1}^{N_x} \sum_{\nu=1}^{N_y} \varphi_{\kappa\nu}, \end{cases} k = 1, 2, \dots, N_x, n = 1, 2, \dots, N_y$$

Приравнивая расчетные и реальные дифференциальные фазы сигналов антенной решетки, получим систему уравнений относительно координат источника тестового сигнала:

$$(10) \quad \Delta\varphi_{nk} = \frac{2\pi}{\lambda} \rho_{nk}(r, \alpha, \beta) - \frac{1}{N_x N_y} \frac{2\pi}{\lambda} \sum_{\kappa=1}^{N_x} \sum_{\nu=1}^{N_y} \rho_{\kappa\nu}(r, \alpha, \beta), k = 1, 2, \dots, N_x, n = 1, 2, \dots, N_y,$$

включающую в качестве параметров измеренные фазовые сдвиги $\Delta\varphi_{nk}$.

Учитывая случайный характер разброса параметров при большом числе элементов антенной решетки, можно с определенной точностью вычислить координаты излучателя на основе метода минимума среднеквадратичной погрешности. В этом случае минимизируется среднеквадратичная погрешность (разность левой и правой частей системы уравнений (10)). Система (10) является нелинейной. Поэтому для нахождения приближенного решения целесообразно использовать итерационную процедуру [1]:

$$(11) \quad \begin{cases} r(v+1) = r(v) - \gamma \frac{\partial}{\partial r} J(r, \alpha, \beta), \\ \alpha(v+1) = \alpha(v) - \chi \frac{\partial}{\partial \alpha} J(r, \alpha, \beta), \\ \beta(v+1) = \beta(v) - \xi \frac{\partial}{\partial \beta} J(r, \alpha, \beta), \end{cases}$$

$$(12) \quad J(r, \alpha, \beta) = \sum_{n=1}^{N_y} \sum_{k=1}^{N_x} \left(\frac{2\pi}{\lambda} \rho_{nk}(r, \alpha, \beta) - \frac{1}{N_x N_y} \frac{2\pi}{\lambda} \sum_{\kappa=1}^{N_x} \sum_{\nu=1}^{N_y} \rho_{\kappa\nu}(r, \alpha, \beta) - \Delta\varphi_{nk} \right)^2 - \text{среднеквадратичная погрешность.}$$

Явное выражение составляющих градиента функционала погрешности имеет следующий вид:

$$(13) \quad \begin{aligned} & \frac{\partial}{\partial r} J(r, \alpha, \beta) = 2 \sum_{n=1}^{N_y} \sum_{k=1}^{N_x} \left(\frac{2\pi}{\lambda} \rho_{nk}(r, \alpha, \beta) - \frac{1}{N_x N_y} \sum_{\kappa=1}^{N_x} \sum_{\nu=1}^{N_y} \frac{2\pi}{\lambda} \rho_{\kappa\nu}(r, \alpha, \beta) - \Delta\varphi_{nk} \right) \times \\ & \times \frac{2\pi}{\lambda} \left(\frac{(x_{nk} - x) \cdot (-\cos\beta \sin\alpha) + (y_{nk} - y) \cdot (-\sin\beta) + z \cos\beta \cos\alpha}{\rho_{nk}(r, \alpha, \beta)} - \right. \\ & \left. - \frac{1}{N_x N_y} \sum_{\kappa=1}^{N_x} \sum_{\nu=1}^{N_y} \frac{(x_{\kappa\nu} - x) \cdot (-\cos\beta \sin\alpha) + (y_{\kappa\nu} - y) \cdot (-\sin\beta) + z \cos\beta \cos\alpha}{\rho_{\kappa\nu}(r, \alpha, \beta)} \right) \end{aligned}$$

$$(14) \quad \frac{\partial J(r, \alpha, \beta)}{\partial \alpha} = 2 \sum_{n=1}^{N_y} \sum_{k=1}^{N_x} \left(\frac{2\pi}{\lambda} \rho_{nk}(r, \alpha, \beta) - \frac{1}{N_x N_y} \sum_{\kappa=1}^{N_y} \sum_{\nu=1}^{N_x} \frac{2\pi}{\lambda} \rho_{\kappa\nu}(r, \alpha, \beta) - \Delta\varphi_{nk} \right) \times$$

$$\times \frac{2\pi}{\lambda} \left(\frac{(x_{nk} - x) \cdot (-r \cos \beta \cos \alpha) - rz \cos \beta \sin \alpha}{\rho_{nk}(r, \alpha, \beta)} - \frac{1}{N_x N_y} \sum_{\kappa=1}^{N_x} \sum_{\nu=1}^{N_y} \frac{(x_{\kappa\nu} - x) \cdot (-r \cos \beta \cos \alpha) - rz \cos \beta \sin \alpha}{\rho_{\kappa\nu}(r, \alpha, \beta)} \right)$$

$$(15) \quad \frac{\partial J(r, \alpha, \beta)}{\partial \beta} = 2 \sum_{n=1}^{N_y} \sum_{k=1}^{N_x} \left(\frac{2\pi}{\lambda} \rho_{nk}(r, \alpha, \beta) - \frac{1}{N_x N_y} \sum_{\kappa=1}^{N_y} \sum_{\nu=1}^{N_x} \frac{2\pi}{\lambda} \rho_{\kappa\nu}(r, \alpha, \beta) - \Delta\varphi_{nk} \right) \times$$

$$\times \frac{2\pi}{\lambda} \left(\frac{(x_{nk} - x) \cdot r \sin \beta \sin \alpha + (y_{nk} - y) \cdot (-r \cos \beta) - zr \sin \beta \cos \alpha}{\rho_{nk}(r, \alpha, \beta)} - \frac{1}{N_x N_y} \sum_{\kappa=1}^{N_x} \sum_{\nu=1}^{N_y} \frac{(x_{\kappa\nu} - x) \cdot r \sin \beta \sin \alpha + (y_{\kappa\nu} - y) \cdot (-r \cos \beta) - zr \sin \beta \cos \alpha}{\rho_{\kappa\nu}(r, \alpha, \beta)} \right)$$

Итерационная процедура позволяет найти оценку координат излучающего устройства, обеспечивающих минимизацию среднеквадратичной фазовой погрешности. При известных координатах источника тестового сигнала можно оценить лучевые фазовые набег $\Delta\varphi_{nk}^0(r, \alpha, \beta)$ и фазовую неидентичность

$$(16) \quad \Delta\varphi_{nk}^{\text{НИ}} = \Delta\varphi_{nk} - \Delta\varphi_{nk}^0(r, \alpha, \beta),$$

и, таким образом, построить массив корректирующих амплитудно-фазовых множителей:

$$(17) \quad \mu_{nk} = \frac{\sum_{\kappa=1}^{N_x} \sum_{\nu=1}^{N_y} A_{\kappa\nu}}{N_x N_y A_{nk}} e^{-j\Delta\varphi_{nk}^{\text{НИ}}}.$$

2 Результаты имитационного моделирования

С целью оценки возможностей данного алгоритма было выполнено имитационное моделирование. При моделировании использовалась гипотетическая антенная решетка размером 10×15 элементов с шагом $d = \lambda/2$. Излучающее устройство располагалось на расстоянии 1 м от решетки с угловым положением, близким к нормали решетки. Фазовая неидентичность имитировалась как случайный массив с нормальным центрированным распределением и стандартным отклонением 20° .

Результаты оценки координат излучающего устройства приведены в таблице 1.

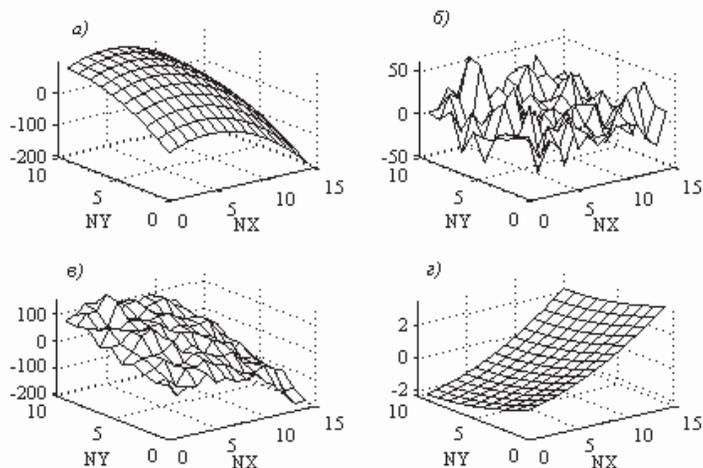
Таблица 1 – Параметры итерационного процесса

	г (см)	α (град)	β (град)
Нулевое приближение	1.000	0°	0°
Конечный результат (500 итераций)	0.991	-3.104°	5.051°
Точное значение	1.000	-3.000°	5.000°

По расчетным координатам можно в соответствии с предлагаемой методикой определить лучевое фазовое распределение, выделить фазовую неидентичность и выполнить фазовую

коррекцию. Погрешность оценки координат излучающего устройства приводит к не полной компенсации неидентичности. Остаточная фазовая неидентичность представляет собой случайно сложившуюся регулярную составляющую исходной неидентичности, которая и приводит к погрешности оценки координат излучающего устройства.

На рисунке 2 показаны полученные при моделировании фазовые распределения на решетке: идеальный волновой фронт, фазовая неидентичность, искаженный неидентичностью фронт, остаточная неидентичность после коррекции.



а) – фазовое распределение на решетке, град., б) – фазовая неидентичность решетки, град., в) – фазовое распределение на решетке с неидентичностью решетки, град., г) – остаточная фазовая неидентичность, град.

Рисунок 2 – Фазовые характеристики антенной решетки

В результате калибровки в данном случае среднее квадратичное значение исходной фазовой неидентичности снижается от $\overline{\Delta\varphi_0} = 20^\circ$ до $\overline{\Delta\varphi} = 1.5^\circ$. При другой реализации неидентичности на решетке остаточная неидентичность может принимать другие значения. Чтобы оценить осредненную остаточную неидентичность по множеству реализаций, проведено статистическое моделирование по 100 реализациям. Среднее значение остаточной неидентичности составляет 2.56° , при этом остаточная неидентичность изменяется от 0.40° до 5.45° . Таким образом, неидентичность снижается в среднем почти на порядок.

Заключение

В данной работе предложен метод калибровки многоэлементных антенных решеток, позволяющий существенно снизить фазовую неидентичность элементов и соответственно повысить качество пространственной селекции сигналов. Основное достоинство данного метода заключается в том, что при постановке экспериментов не требуется дополнительных средств измерения. Их роль выполняет калибруемая решетка.

Список литературы

- [1] Ортега Дж., Рейнболд В. Итерационные методы решения нелинейных систем уравнений с многими неизвестными. М.: Мир, 1975.

МЕТОД ОПРЕДЕЛЕНИЯ ЮСТИРОВОЧНЫХ УГЛОВ ПРИ ОТСУТСТВИИ ТОЧНЫХ ДАННЫХ О ПАРАМЕТРАХ ОРИЕНТАЦИИ ОБЪЕКТА

А.П. Колеватов, А.В. Куликов

Пермский Государственный Технический Университет
614990, г. Пермь, ул. Комсомольский пр., 29, Россия
kavik2@yandex.ru
тел: +7 (912) 887-44-64

Ключевые слова: юстировка, инерциальная система ориентации, погрешности, точность

Abstract

The method adjustment inertial system of the orientation, applied in the conditions of the limited possibilities of use of the measuring equipment is developed. The method is based on use of control positions and allows to calculate analytically values adjustment corners.

Введение

Для бесплатформенной инерциальной системы ориентации (далее по тексту прибор или ИСО), построенной на динамически-настраиваемых гироскопах, возникала необходимость определения углов рассогласования между собственной системой координат и системой координат объекта, на который она устанавливается. Поскольку прибор размещается на установочной площадке объекта неточно, оси ориентации не совпадают. Привести их в соответствие можно путём введения в вычислительные алгоритмы ИСО корректировочных параметров, определяющих данное рассогласование, описываемое тремя углами, которые будем называть юстировочными. Методика их определения называется юстировкой. В общем случае, под ней понимается совокупность операции по приведению средств измерения в состояние обеспечивающие их правильное функционирование. Найденные юстировочные углы должны обеспечить точность определения угловых координат объекта при помощи инерциальной системы ориентации с ошибкой составляющей не более $0,06^\circ$. Решение данной задачи сильно затрудняется ограниченностью необходимого пространства на объекте для размещения дополнительной контрольно-измерительной аппаратуры фиксирующей положение прибора на объекте. По этой причине, юстировка не может быть сделана по традиционной методике, применяемой для авиационных и морских судов.

В настоящей статье предлагаются метод юстировки основанный на анализе данных от инерциальной системы ориентации и данных об угловом положении объекта. Сравнение измерений углов наклона объекта, сделанные ИСО, и полученные от внешних измерителей, предлагается выполнять в заданных контрольных положениях.

Существенно осложняют процесс юстировки следующие факторы:

- 1) При больших углах тангажа отсутствуют данные об угле крена объекта;
- 2) Истинный курс, формируемый инерциальной системы ориентации, при начальной выставке определяется с недопустимо большой погрешностью (ошибка до $0,4^\circ$);
- 3) Наличие инструментальных погрешностей у инерциальных датчиков прибора: смещение нуля у акселерометров (случайная составляющая смещения до $10^{-3}g$), дрейф датчиков угловых скоростей (случайная составляющая дрейфа $0,5^\circ/ч$);
- 4) Контрольные положения объекта формируются со случайной ошибкой до $0,06^\circ$.

Предлагаемый в данной статье подход, в основе которого лежит математическая модель взаиморасположения осей объекта и прибора, а так же принцип использования контрольных положений, призван определить юстировочные углы с требуемой точностью.

Математическая модель

Для расчётов используются следующие системы координат (СК):

- Нормальная СК (горизонтально – географическая) обозначается: X_B, Y_B, Z_B , где X_B – задаёт направление на север; Y_B – направление местной вертикали; Z_B – направление на восток.
- Система координат связанная с прибором: x_p, y_p, z_p . Направление её осей совпадает с направлением осей инерциальных датчиков: ось x_p – продольная ось прибора, y_p – вертикальная ось, перпендикулярная основанию прибора; z_p – поперечная ось прибора.
- Система координат связанная с объектом: x_t, y_t, z_t . Направление осей следующее: ось x_t – продольная ось объекта, y_t – вертикальная ось, перпендикулярная установочной площадке объекта; z_t – поперечная ось.

Ориентация объекта и прибора в пространстве задаётся тремя углами ориентации: углом тангажа (ϑ) – угол между продольной осью объекта и горизонтальной плоскостью; углом крена (γ) – угол наклона продольной плоскости объекта к вертикали, углом курса (ψ) – угол в горизонтальной плоскости, заключенный между северным направлением меридиана, принятого за начало отсчета и проекцией продольной оси самолета на горизонтальную плоскость. Последовательность поворотов $\psi \rightarrow \vartheta \rightarrow \gamma$. За положительное направление угла курса выбран поворот в плоскости горизонта по часовой стрелке для углов тангажа и крена данное направление соответствуют поворотам против часовой стрелки вокруг поперечной и продольной осей прибора соответственно.

Для получения формульных зависимостей СК прибора с объектом объединяются и обозначаются как связанная система координат: x_c, y_c, z_c . Преобразование вектора, координаты которого заданы в нормальной СК, в координаты вектора, заданные в связанной СК, задаётся матрицей направляющих косинусов – В

$$(1) \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = B \cdot \begin{bmatrix} X_B \\ Y_B \\ Z_B \end{bmatrix};$$

Данная матрица получается путём последовательного перемножения матриц поворота для каждого из углов ориентации в следующей последовательности:

$$(2) B = B_\gamma \cdot B_\vartheta \cdot B_\psi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & \sin \gamma \\ 0 & -\sin \gamma & \cos \gamma \end{bmatrix} \cdot \begin{bmatrix} \cos \vartheta & \sin \vartheta & 0 \\ -\sin \vartheta & \cos \vartheta & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos \psi & 0 & \sin \psi \\ 0 & 1 & 0 \\ -\sin \psi & 0 & \cos \psi \end{bmatrix} =$$

$$\begin{bmatrix} \cos \vartheta \cos \psi & \sin \vartheta & \cos \vartheta \sin \psi \\ -\sin \vartheta \cos \gamma \cos \psi - \sin \gamma \sin \psi & \cos \vartheta \cos \gamma & -\sin \vartheta \cos \gamma \sin \psi + \sin \gamma \cos \psi \\ \sin \gamma \sin \vartheta \cos \psi - \cos \gamma \sin \psi & -\cos \vartheta \sin \gamma & \sin \vartheta \sin \gamma \sin \psi + \cos \gamma \cos \psi \end{bmatrix}$$

При преобразовании векторов одной из СК, входящих в группу связанных, (это СК объекта и СК прибора) в вектор нормальной, в обозначении углов добавляются индексы, определяющие принадлежность к нужной СК. Для СК объекта – $\psi_t, \vartheta_t, \gamma_t$; для СК прибора – $\psi_p, \vartheta_p, \gamma_p$, матрицы преобразования B_t, B_p соответственно.

Прибор устанавливается на установочной площадке объекта неточно и это означает, что положения СК прибора и СК объекта не совпадают. Их рассогласование будет описываться тремя углами Эйлера, которые и называются юстировочными: $\alpha_x, \alpha_y, \alpha_z$ – нижний индекс угла есть обозначение оси, вокруг которой выполняется поворот. За положительный угол α_y , выберем угол поворота в плоскости основания прибора по часовой стрелке (аналогично положи-

тельному направлению угла курса). Положительные углы α_x, α_z соответствуют повороту против часовой стрелки (аналогично углам крена и тангажа соответственно). Последовательность поворотов $\alpha_y \rightarrow \alpha_z \rightarrow \alpha_x$. Оценивание этих углов будет выполняться в результате решения задачи юстировки. Преобразование координат вектора, заданного в СК прибора в координаты вектора СК объекта задаётся матрицей A , которая называется юстировочной.

$$(3) \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = A \cdot \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix}$$

Находится A путём последовательного перемножения матриц поворота на соответствующие юстировочные углы:

$$(4) A = A_{\alpha_x} \cdot A_{\alpha_z} \cdot A_{\alpha_y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha_x & \sin \alpha_x \\ 0 & -\sin \alpha_x & \cos \alpha_x \end{bmatrix} \cdot \begin{bmatrix} \cos \alpha_z & \sin \alpha_z & 0 \\ -\sin \alpha_z & \cos \alpha_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos \alpha_y & 0 & \sin \alpha_y \\ 0 & 1 & 0 \\ -\sin \alpha_y & 0 & \cos \alpha_y \end{bmatrix} =$$

$$\begin{bmatrix} \cos \alpha_z \cos \alpha_y & \sin \alpha_z & \cos \alpha_z \sin \alpha_y \\ -\sin \alpha_z \cos \alpha_x \cos \alpha_y - \sin \alpha_x \sin \alpha_y & \cos \alpha_z \cos \alpha_x & -\sin \alpha_z \cos \alpha_x \sin \alpha_y + \sin \alpha_x \cos \alpha_y \\ \sin \alpha_x \sin \alpha_z \cos \alpha_y - \cos \alpha_x \sin \alpha_y & -\cos \alpha_z \sin \alpha_x & \sin \alpha_z \sin \alpha_x \sin \alpha_y + \cos \alpha_x \cos \alpha_y \end{bmatrix}$$

Преобразование нормальной СК в СК объекта и прибора отражается следующими формульными зависимостями (5) и (6):

$$(5) \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = B_t \cdot \begin{bmatrix} X_B \\ Y_B \\ Z_B \end{bmatrix},$$

$$(6) \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} = B_p \cdot \begin{bmatrix} X_B \\ Y_B \\ Z_B \end{bmatrix},$$

где, элементы матриц B_t и B_p получены из элементов матрицы B введением в обозначение углов индекса t и p для объекта и прибора соответственно.

Поскольку вектора СК прибора преобразуются в координаты вектора СК объекта с помощью юстировочной матрицы A , то справедливо следующее преобразование:

$$(7) \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = A \cdot \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} \Rightarrow \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = A \cdot B_p \cdot \begin{bmatrix} X_B \\ Y_B \\ Z_B \end{bmatrix}$$

Тождественность преобразований (5) и (7) позволяет записать матричное уравнение в виде:

$$(8) B_t = A \cdot B_p \Rightarrow A = B_t \cdot B_p^T$$

Тождественно из (8):

$$(9) B_p = A^T \cdot B_t$$

Таким образом, из (8) может быть найдена юстировочная матрица, а вместе с ней и юстировочные углы.

Метод аналитического определения юстировочных углов

В том случае, если бы все угловые координаты объекта и прибора, соответственно элементы матрицы направляющих косинусов B_l и B_p , были достоверно известны, то аналитическое нахождение значений юстировочных углов α_x , α_y , α_z при помощи выражения (8) не представляло бы труда. При решении задачи юстировки на практике остаются неизвестными достоверные значения углов крена объекта и курса прибора.

Для однозначного определения юстировочных углов и других неизвестных параметров ориентации, необходимо решить систему уравнений (9). В результате перемножения транспонированной юстировочной матрицы и матрицы направляющих косинусов связанной с объектом получается система, состоящая из девяти уравнений, где элементы в правой части – измерения сделанные прибором. Ниже приведено одно из таких уравнений:

$$(10) \begin{aligned} & \cos \alpha_x \cdot \cos \alpha_y \cdot \sin \alpha_z \cdot \cos \gamma_l \cdot \cos \psi_l \cdot \sin \vartheta_l + \cos \alpha_x \cdot \cos \alpha_y \cdot \sin \alpha_z \cdot \sin \gamma_l \cdot \sin \psi_l + \\ & \cos \alpha_x \cdot \sin \alpha_y \cdot \cos \gamma_l \cdot \sin \psi_l - \sin \alpha_y \cdot \sin \gamma_l \cdot \cos \psi_l \cdot \sin \vartheta_l \cdot \cos \alpha_x + \\ & \sin \alpha_x \cdot \sin \alpha_y \cdot \cos \gamma_l \cdot \cos \psi_l \cdot \sin \vartheta_l + \sin \alpha_x \cdot \sin \alpha_y \cdot \sin \gamma_l \cdot \sin \psi_l - \\ & \sin \alpha_x \cdot \cos \alpha_y \cdot \sin \alpha_z \cdot \cos \gamma_l \cdot \sin \psi_l + \sin \alpha_x \cdot \cos \alpha_y \cdot \sin \alpha_z \cdot \sin \gamma_l \cdot \cos \psi_l \cdot \sin \vartheta_l + \\ & \cos \alpha_y \cdot \cos \alpha_z \cdot \cos \psi_l \cdot \cos \vartheta_l = \cos \vartheta_p \cdot \cos \psi_p \end{aligned}$$

Как можно видеть на примере данного уравнения, каждое из них содержат большое количество параметров. Неизвестными среди них являются следующие углы: крена объекта - γ_l , курса прибора - ψ_p , юстировочные по каналу курса - α_y , крена - α_x , и тангажа - α_z . Решить систему уравнений с таким количеством неизвестных параметров, которые к тому же взаимно перемножаются – не представляется возможным.

Несколько упрощает решение задачи возможность определить углы тангажа и крена соответственно α_x и α_z , когда объект находится в плоскости горизонта. В этом контрольном положении значения указанных углов объекта устанавливаются равными нулю при помощи, например, оптического квадранта, обеспечивающего высокую точность измерений. Таким образом, α_z имеет численное значение равное углу тангажа измеренному прибором в горизонтальном положении взятое с противоположным знаком. Аналогично определяется юстировочный угол по каналу крена - α_x . Определить α_y в горизонте не представляется возможным по причине низкой точности определения курса инерциальной системой ориентации. Теме не менее, истинный курс объекта может быть определён с требуемой точностью по ориентирам на местности.

При известных юстировочных углах α_x и α_z решение (9) несколько упрощается. Не смотря на то, что оставшиеся неизвестные встречаются в уравнениях как с **sin** так и с **cos**, в результате линеаризации получается система, состоящая из девяти уравнений при десяти неизвестных. Решить такую систему для одного контрольного положения объекта в плоскости горизонта нельзя, но сделать это становится возможным, собрав данные об углах ориентации во втором контрольном положении, которое формируется путём изменения угла тангажа до максимально возможного значения близкого к 90° . В этом случае, для решения системы необходимо, чтобы неизвестный угол курса прибора ψ_p в обоих контрольных положениях был неизменен.

Таким образом, имея данные об углах ориентации объекта и прибора в двух контрольных положениях, становится возможным получить систему, состоящую из десяти уравнений при таком же количестве неизвестных. Её решение позволяет оценить неизвестные параметры ориентации объекта с установленным на нём прибором и определить все искомые юстировочные углы. Данная методика юстировки была математически промоделирована в среде Mathcad при задании двух тестовых контрольных положений. Полученные результаты подтвердили корректность математической модели и правильность составления системы уравнений. Однако в

результате проведения анализа чувствительности модели к ошибкам в задании углов ориентации были сделаны выводы, что модель весьма чувствительна к ошибкам.

Заключение

Построена математическая модель взаиморасположения осей ориентации объекта с установленным на нём прибором. На основании её разработана методика аналитического определения юстировочных углов, которая в связи с высокой чувствительностью к ошибкам не может определять параметры ориентации объекта с максимально возможной точностью, на которую потенциально способна инерциальная система ориентации.

Список литературы

- [1] Бранец В.Н., Шмыглёвский И.П. Применение кватэрнионов в задачах ориентации твёрдого тела. М.: Наука, 1973 – 320с.
- [2] Ишлинский А.Ю. Ориентация, гироскопы и инерциальная навигация. М.: Наука, 1976 - 672с.
- [3] Ривкин С.С., Берман З.М., Окон И.М. Определение параметров ориентации объекта бесплатформенной инерциальной системой. СПб:ГНЦ РФ – ЦНИИ «Электроприбор», 1996 – 226с.
- [4] Salychev O.S. Inertial Systems in Navigation and Geophysics, Moscow: Bauman MSTU Press, 1998 – 216с.

Biography

Andrey P. Kolevatov graduated from the Perm Polytechnical Institute in 1979. He got a Ph. D. in CSRI “Electropribor”, Leningrad, in 1986. He is the assistant professor of the Perm State Technical University, Chair of Measurement and Computer Systems for Aircrafts. He is engaged in development of algorithms and software for different navigation systems.

Andrey V. Kulikov graduated from the Perm Polytechnical Institute in 2006. He is the assistant of the Perm State Technical University, Chair of Measurement and Computer Systems for Aircrafts. He is developing adjustment methods of navigation systems.

МОДЕЛИРОВАНИЕ КАЛИБРОВКИ БЕСПЛАТФОРМЕННОЙ ИНЕРЦИАЛЬНОЙ НАВИГАЦИОННОЙ СИСТЕМЫ ПО ВЫХОДНЫМ СИГНАЛАМ

С.Г. Николаев¹, Ю.В. Ившина²

¹Пермский государственный технический университет
614990, Пермь, Комсомольский пр., 29, Россия
тел: +7 (342) 2-391-229

²ОАО Пермская научно-производственная приборостроительная компания
614990, Пермь, ул. 25 Октября, 106, Россия
ivshina@ppk.perm.ru, uivka@yandex.ru
тел: +7 (342) 2-400-620

***Ключевые слова:** калибровка, бесплатформенная инерциальная навигационная система, акселерометр, гироскоп, модель ошибок*

Abstract

The static mode of calibration of strapdown inertial navigating system is considered. There are two stages of the calibration. At the first stage according to the system data output the total errors of inertial sensor are determined, then at the second stage on the base of the total errors the calibration factors are computed.

Введение

Точность работы бесплатформенных инерциальных навигационных систем (БИНС) зависит от качества выполнения калибровки инерциальных измерителей. Целью калибровки является определение систематических ошибок инерциальных датчиков (калибровочных коэффициентов), которые затем учитываются в алгоритмах бортового вычислителя БИНС.

Традиционный способ калибровки предусматривает определение калибровочных коэффициентов с использованием сигналов, получаемых непосредственно с датчиков [1]. Различают два режима калибровки: статический и динамический [2]. При статическом режиме калибровки с помощью калибровочного наклонно-поворотного стола (НПС) задаются фиксированные положения (ориентации) БИНС. В этом режиме определяются все калибровочные коэффициенты акселерометров и большинство калибровочных коэффициентов гироскопов. При динамическом режиме выполняют вращения БИНС на НПС вокруг заданных осей, при этом определяются ошибки масштабных коэффициентов и углы перекосов осей чувствительности гироскопов.

Основным недостатком данного способа калибровки является высокий уровень шумов в сигналах инерциальных датчиков. Кроме того, в выходных сигналах гироскопов необходимо компенсировать составляющие угловой скорости суточного вращения Земли. Если гироскопы имеют не цифровой, а аналоговый выход, то возникает дополнительное требование: точное априорное знание масштабных коэффициентов гироскопов для компенсации составляющих угловой скорости суточного вращения Земли.

Доклад посвящен исследованию точностных характеристик альтернативного способа определения калибровочных коэффициентов инерциальных датчиков на примере определения ошибок масштабных коэффициентов гироскопов. Их определение производится в статическом режиме калибровки по выходным сигналам БИНС, шумовые погрешности которых существенно ниже, чем у выходных сигналов, получаемых непосредственно с датчиков. Исследование проведено методом численного моделирования с имитацией выходных данных БИНС, содержащих подлежащие калибровке погрешности инерциальных измерителей.

1 Постановка задачи

Необходимо оценить точность определения калибровочных коэффициентов датчиков по выходным сигналам БИНС в статическом режиме калибровки. Исследование будет проводиться методом имитационного моделирования на примере определения ошибок масштабных коэффициентов гироскопов.

Введем базовый навигационный трехгранник OX_wY_wZ_w с полусвободной ориентацией в азимуте и систему координат OXYZ, связанную с объектом. Положение связанной системы координат OXYZ относительно системы OX_wY_wZ_w определяется параметрами ориентации: ψ – угол рысканья, ϑ – угол тангажа, γ – угол крена.

Математическую модель ошибок БИНС с полусвободной ориентацией базового навигационного трехгранника в азимуте представим в виде системы дифференциальных уравнений:

$$\begin{aligned}
 \dot{\delta\gamma}_w &= \frac{1}{R} \delta V_{X_w} - \Omega \cdot \sin \varphi \cdot \delta \vartheta_w - \delta \omega_{Y_w}; \\
 \dot{\delta\psi}_w &= \Omega \cdot \cos \varphi \cdot \delta \vartheta_w - \delta \omega_{Z_w}; \\
 (1) \quad \dot{\delta\vartheta}_w &= -\frac{1}{R} \delta V_{Y_w} - \Omega \cdot \cos \varphi \cdot \delta \psi_w + \Omega \cdot \sin \varphi \cdot \delta \gamma_w - \delta \omega_{X_w}; \\
 \delta V_{X_w} &= -g \delta \gamma_w + 2 \cdot \Omega \cdot \sin \varphi \cdot \delta V_{Y_w} + \delta f_{X_w}; \\
 \delta V_{Y_w} &= g \delta \vartheta_w - 2 \cdot \Omega \cdot \sin \varphi \cdot \delta V_{X_w} + \delta f_{Y_w}.
 \end{aligned}$$

где R – радиус Земли; g – ускорение силы тяжести; $[\delta\gamma_w \ \delta\psi_w \ \delta\vartheta_w \ \delta V_{X_w} \ \delta V_{Y_w}]^T$ – вектор ошибок БИНС, принятый в качестве вектора измерений задачи калибровки (ошибки параметров ориентации $\delta\gamma_w$, $\delta\psi_w$, $\delta\vartheta_w$ определяются по разности показаний БИНС и НПС, ошибки определения составляющих линейной скорости δV_n , δV_e определяются по показаниям БИНС); $[\delta\omega_{X_w} \ \delta\omega_{Y_w} \ \delta\omega_{Z_w}]^T$ – вектор проекций суммарных ошибок гироскопов на оси базового навигационного трехгранника; $[\delta f_{X_w} \ \delta f_{Y_w}]^T$ – вектор проекций суммарных ошибок акселерометров на оси базового навигационного трехгранника.

Система (1) получена с учетом того, что БИНС установлена на неподвижном основании и в бортовые алгоритмы заводилось точное значение широты φ .

Приборное значение вектора угловой скорости $\omega^b + \delta\omega^b$ содержит проекции угловой скорости вращения Земли и инструментальные ошибки:

$$(2) \quad \omega^b + \delta\omega^b = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} + \begin{bmatrix} \omega_{nx} \\ \omega_{ny} \\ \omega_{nz} \end{bmatrix} + \begin{bmatrix} \Delta k_{dx} & 0 & 0 \\ 0 & \Delta k_{dy} & 0 \\ 0 & 0 & \Delta k_{dz} \end{bmatrix} \cdot \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} + \begin{bmatrix} n_{dx} \\ n_{dy} \\ n_{dz} \end{bmatrix},$$

где $[\omega_x \ \omega_y \ \omega_z]^T$ – проекции угловой скорости вращения Земли, зависящие от ориентации НПС; ω_{nx} , ω_{ny} , ω_{nz} – ошибки смещения нуля; Δk_{dx} , Δk_{dy} , Δk_{dz} – ошибки масштабных коэффициентов; $[n_{dx} \ n_{dy} \ n_{dz}]^T$ – вектор шумовых погрешностей.

Приборное значение вектора линейного ускорения $f^b + \delta f^b$ в проекциях на оси чувствительности акселерометров содержит проекции ускорения силы тяжести g и инструментальные ошибки:

$$(3) \quad f^b + \delta f^b = \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} + \begin{bmatrix} f_{nx} \\ f_{ny} \\ f_{nz} \end{bmatrix} + \begin{bmatrix} \Delta k_{ax} & 0 & 0 \\ 0 & \Delta k_{ay} & 0 \\ 0 & 0 & \Delta k_{az} \end{bmatrix} \cdot \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} + \begin{bmatrix} n_{ax} \\ n_{ay} \\ n_{az} \end{bmatrix},$$

где $[f_x f_y f_z]^T$ – проекции ускорения силы тяжести, зависящие от ориентации НПС; f_{nx}, f_{ny}, f_{nz} – ошибки смещения нуля; $\Delta k_{ax}, \Delta k_{ay}, \Delta k_{az}$ – ошибки масштабных коэффициентов; $[n_{ax} n_{ay} n_{az}]^T$ – вектор шумовых погрешностей.

Имитационное моделирование процесса калибровки будем выполнять с учетом следующих условий:

- уровни шумовых составляющих n_{di} в выходных сигналах гироскопов составляют от 1 до 30 °/ч^{3/2} при съеме информации с частотой 50 Гц;
- уровни шумовых составляющих n_{ai} в выходных сигналах акселерометров составляют от 10–4g до 10–3g м/с^{3/2} при съеме информации с частотой 50 Гц;
- шумовые составляющие задаются при помощи генератора случайных чисел, причем для исключения корреляции сигналов формируется отдельный массив для каждого датчика.

Статический режим калибровки предусматривает положение НПС с установленным на нем БИНС в определенных ориентациях. Бортовые алгоритмы накладывают ограничения в выборе ориентаций, т.к. углы γ и ϑ – не должны превышать значения $\pi/2$.

2 Описание алгоритма моделирования

Сформируем показания инерциальных измерителей с учетом принятых моделей ошибок и значений калибровочных коэффициентов. Далее имитируется работа бортовых алгоритмов БИНС с полусвободной ориентацией навигационного трехгранника в азимуте [3]. По выходным значениям БИНС при выбранной ориентации получаем вектор ошибок $X = [\delta\gamma_w \delta\psi_w \delta\vartheta_w \delta V_{Xw} \delta V_{Yw}]^T$ и решаем обратную задачу определения значений калибровочных коэффициентов.

Задача решается в два этапа. Решение обратной задачи на первом этапе сводится к решению некоторого операторного уравнения:

$$(4) \quad A \cdot y = X,$$

где A – некоторый математический известный оператор (интегральный, дифференциальный, алгебраический и т.д.); y – входная (искомая, неизвестная) функция; X – результат измерений или выходная функция.

Для нашего случая оператор A будет представляться единичной матрицей, а выходная функция X будет сформирована по измеренному вектору ошибок БИНС $[\delta\gamma_w \delta\psi_w \delta\vartheta_w \delta V_{Xw} \delta V_{Yw}]^T$ с использованием структуры математической модели системы (1).

На первом этапе, используя вектор ошибок X , по модели ошибок БИНС определяются каналные ошибки инерциальных измерителей $\delta\tilde{\omega}_{Xw}, \delta\tilde{\omega}_{Yw}, \delta\tilde{\omega}_{Zw}, \tilde{\delta}\tilde{\gamma}_{Xw}, \tilde{\delta}\tilde{\gamma}_{Yw}$, которые в свою очередь являются функциями инструментальных погрешностей инерциальных измерителей.

В нашем случае операторное уравнение (4) принимает вид:

$$(5) \quad E \cdot y = X,$$

где оператор A трансформируется в единичную матрицу E ; входная функция y представляется в виде вектора $[\delta\tilde{\omega}_{Xw} \delta\tilde{\omega}_{Yw} \delta\tilde{\omega}_{Zw} \tilde{\delta}\tilde{\gamma}_{Xw} \tilde{\delta}\tilde{\gamma}_{Yw}]^T$, а выходная функция X в виде:

$$(6) \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{bmatrix} = \begin{bmatrix} -\dot{\delta}\gamma_w + \frac{1}{R} \delta V_{Xw} - \Omega \cdot \sin \varphi \cdot \delta v_w \\ -\dot{\delta}\psi_w + \Omega \cdot \cos \varphi \cdot \delta v_w \\ -\dot{\delta}v_w - \frac{1}{R} \delta V_{Yw} - \Omega \cdot \cos \varphi \cdot \delta \psi_w + \Omega \cdot \sin \varphi \cdot \delta \gamma_w \\ \delta V_{Xw} + g \delta \gamma_w - 2 \cdot \Omega \cdot \sin \varphi \cdot \delta V_{Yw} \\ \delta V_{Yw} - g \delta v_w + 2 \cdot \Omega \cdot \sin \varphi \cdot \delta V_{Xw} \end{bmatrix}$$

В общем случае операторное уравнение (5) является дифференциальным. Для его алгебраизации поступим следующим образом. Из численного эксперимента следует, что составляющие вектора ошибок БИНС $[\delta\tilde{\omega}_{Xw} \delta\tilde{\omega}_{Yw} \delta\tilde{\omega}_{Zw} \tilde{\delta}f_{Xw} \tilde{\delta}f_{Yw}]^T$ представляют собой набор гладких, дифференцируемых функций. Для некоторых ограниченных отрезков времени представим выходную функцию X как суперпозицию среднего значения вектора ошибок и его производной.

В этом случае операторное уравнение (5) можно представить в виде алгебраических уравнений:

$$(7) \begin{cases} \delta\tilde{\omega}_{Yw}(t_i) = -\frac{\delta\gamma_w(t_i) - \delta\gamma_w(t_{i-1})}{T} + \frac{\delta V_{Xw}(t_i)}{R} - \Omega \cdot \sin \varphi \cdot \delta v_w(t_i); \\ \delta\tilde{\omega}_{Zw}(t_i) = -\frac{\delta\psi_w(t_i) - \delta\psi_w(t_{i-1})}{T} + \Omega \cdot \cos \varphi \cdot \delta v_w(t_i); \\ \delta\tilde{\omega}_{Xw}(t_i) = -\frac{\delta v_w(t_i) - \delta v_w(t_{i-1})}{T} - \frac{\delta V_{Yw}(t_i)}{R} - \Omega \cdot \cos \varphi \cdot \delta \psi_w(t_i) + \Omega \cdot \sin \varphi \cdot \delta \gamma_w(t_i); \\ \tilde{\delta}f_{Xw}(t_i) = \frac{\delta V_{Xw}(t_i) - \delta V_{Xw}(t_{i-1})}{T} + g \cdot \delta \gamma_w(t_i) - 2 \cdot \Omega \cdot \sin \varphi \cdot \delta V_{Yw}(t_i); \\ \tilde{\delta}f_{Yw}(t_i) = \frac{\delta V_{Yw}(t_i) - \delta V_{Yw}(t_{i-1})}{T} - g \cdot \delta v_w(t_i) + 2 \cdot \Omega \cdot \sin \varphi \cdot \delta V_{Xw}(t_i). \end{cases}$$

На рисунке 1 для примера представлен график канальной погрешности $\delta\tilde{\omega}_{Yw}$.

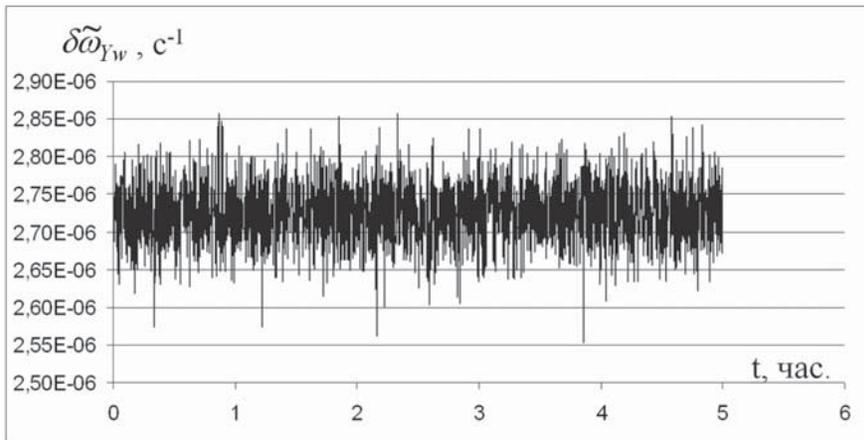


Рисунок 1 – График изменения канальной погрешности $\delta\tilde{\omega}_{Yw}$

Так как каналные ошибки БИНС $[\delta\tilde{\omega}_{Xw} \delta\tilde{\omega}_{Yw} \delta\tilde{\omega}_{Zw} \tilde{\delta}f_{Xw} \tilde{\delta}f_{Yw}]^T$ являются функциями инструментальных ошибок инерциальных измерителей, на втором этапе по традиционной технологии по каналным ошибкам определяются калибровочные коэффициенты инерциальных измерителей.

Для этого, используя определенные ранее значения канальных ошибок $\delta\tilde{\omega}_{x_w}, \delta\tilde{\omega}_{y_w}, \delta\tilde{\omega}_{z_w}, \tilde{f}_{x_w}, \tilde{f}_{y_w}$, составляется алгебраическая система уравнений для определения значений калибровочных коэффициентов. Фактически это тоже решение обратной задачи.

Выберем ориентации НПС так, как показано в таблице 1.

Таблица 1 – Ориентации наклонно-поворотного стола

№ п/п	Ориентация OXYZ		Ориентация OX _w Y _w Z _w
1	$\vartheta = 0^\circ$ $\psi = 0^\circ$ $\gamma = 0^\circ$		
2	$\vartheta = 0^\circ$ $\psi = 0^\circ$ $\gamma = 45^\circ$		
3	$\vartheta = 0^\circ$ $\psi = 0^\circ$ $\gamma = -45^\circ$		
4	$\vartheta = 0^\circ$ $\psi = 45^\circ$ $\gamma = 0^\circ$		
5	$\vartheta = 0^\circ$ $\psi = 180^\circ$ $\gamma = 0^\circ$		

Приведем в качестве примера определение масштабного коэффициента Δk_{dy} гироскопа, установленного по оси Y.

Для первой и пятой ориентаций НПС составим алгебраическую систему уравнений для определения Δk_{dy} .

$$(8) \begin{bmatrix} -1 & -\Omega \cos \varphi \\ 1 & -\Omega \cos \varphi \end{bmatrix} \cdot \begin{bmatrix} \omega_{ny} \\ \Delta k_{dy} \end{bmatrix} = \begin{bmatrix} \delta\omega_{y_w}(1) \\ \delta\omega_{y_w}(5) \end{bmatrix}$$

Используя (8) строим переопределенную алгебраическую систему, которую решаем методом наименьших квадратов и получаем первую итерацию Δk_{dy} .

Полученное значение Δk_{dy} заносим в бортовой вычислитель и повторяем процесс моделирования.

3 Результаты моделирования

Было проведено численное моделирование процесса калибровки и оценка точности предлагаемого метода на примере ошибки масштабного коэффициента Δk_{dy} . Условия моделирования соответствовали перечисленным во введении, а искомая ошибка составляла 3 % от номинального значения величины соответствующего масштабного коэффициента. Результаты моделирования приведены на рисунке 2 в виде графика изменения приведенной относительной погрешности определения Δk_{dy} в зависимости от шага итерации. Из графика следует, что процесс определения Δk_{dy} является сходящимся. После третьей итерации приведенная относительная ошибка Δk_{dy} находилась в «трубке» диаметром $4 \cdot 10^{-2}$ % от исходного заданного значения погрешности. Диаметр «трубки» в данном случае зависел от уровня шумов гироскопов и акселерометров.

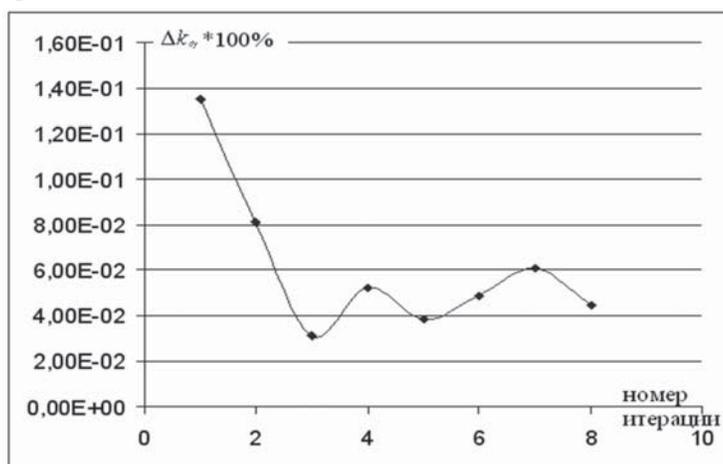


Рисунок 2 – Приведенная относительная погрешность определения масштабного коэффициента

Заключение

На основании результатов моделирования процесса определения ошибки масштабного коэффициента вертикального гироскопа по модели ошибок БИНС, можно сделать следующие выводы:

- моделирование подтвердило возможность калибровки БИНС по ее модели ошибок в статическом режиме;
- потенциальная точность определения ошибки масштабного коэффициента высокая, окончательное заключение о точности калибровки будет сделано по результатам натурального эксперимента;
- уровень шумов входящих в состав БИНС инерциальных измерителей влияет на точность определения ошибки масштабного коэффициента гироскопа.

Список литературы

- [1] Брозгуль Л.И. Динамически настраиваемые гироскопы: Модели погрешностей для систем навигации. – М.: Машиностроение. 1989. с. 105.
- [2] Titterton D.H., Weston J.L. Strapdown Inertial Navigation Technology - 2nd Edition. The Institution of Electrical Engineers, 2004. 255-257 p.
- [3] Salychev O. Inertial Systems in Navigation and Geophysics. Moscow, 1998. 352 p.

МЕТОДЫ ИССЛЕДОВАНИЯ ЧУВСТВИТЕЛЬНОСТИ АВТОКОЛЕБАНИЙ В РЕЛЕЙНЫХ СИСТЕМАХ УПРАВЛЕНИЯ К ИЗМЕНЕНИЮ ПАРАМЕТРОВ ОБЪЕКТА УПРАВЛЕНИЯ

С.В. Моржова, Н.В. Фалдин

Тульский Государственный университет
300600, Тула, пр. Ленина, 92, Россия
info@tsu.tula.ru
тел: +7 (4872) 35-34-44, факс: +7 (4872) 33-13-05

Ключевые слова: автоколебания, коэффициент чувствительности, релейная система, чувствительность, фазовый автограф

Abstract

The methods allowing to estimate relay systems periodic motions sensitivity to plant parameters variations are proposed.

Введение

Реальная техническая система является работоспособной, если она может нормально функционировать не только при номинальных значениях параметров, но и при некоторых их отклонениях, которые всегда наблюдаются на практике. Поэтому еще на этапе проектирования необходимо оценить и учесть с помощью теории чувствительности степень влияния изменений отдельных параметров объекта управления на различные характеристики системы.

Настоящая работа посвящена разработке методов исследования чувствительности релейных автоколебательных систем управления. Рассматриваются релейные системы с двухпозиционным релейным элементом и линейным объектом управления. В работе предлагается эффективный метод исследования чувствительности периодических движений в релейной системе к изменению параметров объекта управления. Приводится методика исследования чувствительности асимптотической орбитальной устойчивости автоколебаний.

1 Чувствительность автоколебаний

На рисунке 1 представлена типовая структурная схема релейной системы.

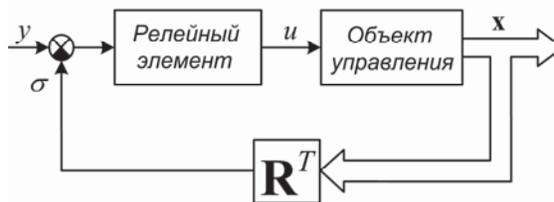


Рисунок 1 – Типовая структурная схема релейной системы

Движение системы задается уравнениями:

$$(1) \frac{dx}{dt} = \mathbf{C}(\alpha)\mathbf{x} + \mathbf{B}(\alpha)u,$$

$$(2) u = \Phi(y - \sigma(\mathbf{x}), b), \quad \sigma(\mathbf{x}) = \mathbf{R}^T \mathbf{x},$$

где \mathbf{x} – n -мерный вектор состояния, матрицы \mathbf{C} и \mathbf{B} имеют соответственно размерности $n \times n$ и $n \times 1$, α – некоторый изменяющийся параметр объекта управления, функция Φ задается статической характеристикой двухпозиционного релейного элемента (РЭ) (рисунок 2), $2b$ – ширина петли гистерезиса, \mathbf{R}^T – вектор-строка коэффициентов обратных связей, y – входной сигнал. На рисунке 3 представлен вид управляющего сигнала с выхода РЭ.

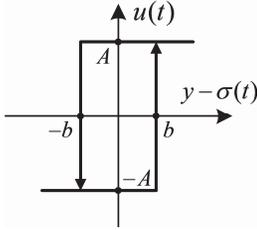


Рисунок 2 – Статическая характеристика двухпозиционного РЭ

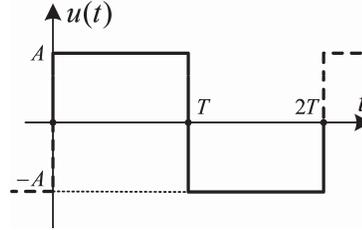


Рисунок 3 – Управляющий сигнал $u(t)$ с выхода РЭ

Будем предполагать, что исследование периодических движений выполняется методом фазового годографа [1].

В автономной ($y(t) \equiv 0$) релейной системе (1), (2) периодическое движение удобно задавать точкой $\mathbf{x}^*(T)$, соответствующей переключению РЭ с минуса на плюс. Вектор-функция $\mathbf{x}^*(T)$, $0 \leq T \leq \infty$, выделяет множество всех возможных периодических движений объекта управления (1) и называется фазовым годографом системы. Фазовый годограф позволяет очень просто определять возникающее в релейной системе периодическое движение [1].

Пусть в автономной релейной системе (1), (2) имеет место симметричное периодическое решение $\mathbf{x}(t)$ с периодом $2T$, задаваемое точкой $\mathbf{x}^*(T)$. Обозначим $\tilde{\mathbf{x}}(t) = \mathbf{x}(t) + \delta\mathbf{x}(t)$ «возмущенную» траекторию, причем «возмущение» вызвано малым изменением $\delta\alpha$ параметра α . Функция $\tilde{\mathbf{x}}(t)$ представляет собой периодическое движение с периодом $2(T + \delta T)$. Здесь δT – малое изменение полупериода. Траекторию $\tilde{\mathbf{x}}(t)$ будем считать близкой к траектории $\mathbf{x}(t)$ (см. рисунок 4). На траекториях $\mathbf{x}(t)$ и $\tilde{\mathbf{x}}(t)$ момент $t = 0$ совместим с моментом переключения РЭ с минуса на плюс. Переключение РЭ с плюса на минус на траектории $\mathbf{x}(t)$ происходит в момент T , а на траектории $\tilde{\mathbf{x}}(t)$ – в близкий момент времени $T + \delta T$.

Движение «возмущенной» системы задается уравнением

$$(3) \quad \frac{d\tilde{\mathbf{x}}}{dt} = \mathbf{C}(\alpha + \delta\alpha)\tilde{\mathbf{x}} + \mathbf{B}(\alpha + \delta\alpha)u,$$

$$(4) \quad u = \Phi(y - \sigma(\tilde{\mathbf{x}}), b), \quad \sigma(\tilde{\mathbf{x}}) = \mathbf{R}^T \tilde{\mathbf{x}}.$$

Из (3), опуская величины, имеющие порядок малости выше первого, и с учетом (1) получим уравнение в вариациях

$$(5) \quad \frac{d\delta\mathbf{x}}{dt} = \mathbf{C}(\alpha)\delta\mathbf{x} + \left[\frac{\partial\mathbf{C}(\alpha)}{\partial\alpha}\mathbf{x} + \frac{\partial\mathbf{B}(\alpha)}{\partial\alpha}u \right] \delta\alpha.$$

Поскольку рассматривается движение на полупериоде, то управление u полагается постоянным ($u = A$), $\frac{\partial\mathbf{C}(\alpha)}{\partial\alpha}$ и $\frac{\partial\mathbf{B}(\alpha)}{\partial\alpha}$ – матрицы с размерностью соответственно $n \times n$ и $n \times 1$.

Обозначим $\Delta\mathbf{x}(T) = \tilde{\mathbf{x}}(T + \delta T) - \mathbf{x}(T)$. С точностью до величин порядка малости не выше первого

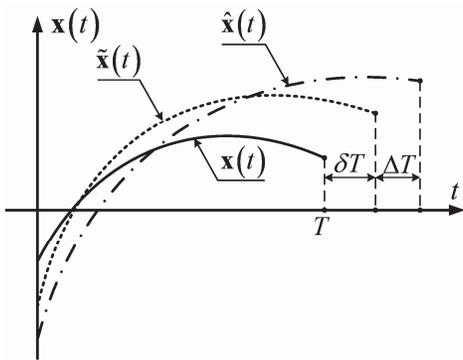


Рисунок 4 – Вид траекторий $x(t)$, $\tilde{x}(t)$ и $\hat{x}(t)$

$$(6) \quad \Delta x(T) = \dot{x}^-(T)\delta T + \delta x^-(T).$$

Обозначим $V(T)$ нормированную фундаментальную матрицу решений уравнения (5), а $r(T)$ – решение уравнения (5) при нулевых начальных условиях и $\delta\alpha = 1$. Тогда

$$(7) \quad \delta x^-(T) = V(T)\delta x(0) + r(T)\delta\alpha.$$

Подставляя (7) в (6), найдем

$$(8) \quad \Delta x(T) = V(T)\delta x(0) + r(T)\delta\alpha + \dot{x}^-(T)\delta T.$$

Из условий переключения РЭ для «возмущенной» и «невозмущенной» траекторий следует

$$(9) \quad R^T \Delta x(T) = 0.$$

Из (8) и (9) получим

$$(10) \quad \delta T = -\frac{R^T (V(T)\delta x(0) + r(T)\delta\alpha)}{R^T \dot{x}^-(T)}.$$

С учетом последнего выражения равенство (8) запишется в виде

$$(11) \quad \Delta x(T) = \left[I - \frac{\dot{x}^-(T)R^T}{R^T \dot{x}^-(T)} \right] (V(T)\delta x(0) + r(T)\delta\alpha).$$

Если имеет место периодическое движение, то $\Delta x(T) = -\delta x(0)$, следовательно

$$(12) \quad \delta x(0) = -\left[I + V(T) - \frac{\dot{x}^-(T)R^T V(T)}{R^T \dot{x}^-(T)} \right]^{-1} \left[I - \frac{\dot{x}^-(T)R^T}{R^T \dot{x}^-(T)} \right] r(T)\delta\alpha.$$

Таким образом, «возмущенное» периодическое движение задается фазовой точкой $x^*(T) + \delta x(0)$. При этом с учетом (12) и (10) полупериод изменится на величину

$$(13) \quad \delta T = K_\alpha^T \delta\alpha,$$

где

$$K_\alpha^T = \frac{R^T \left(V(T) \left[I + V(T) - \frac{\dot{x}^-(T)R^T V(T)}{R^T \dot{x}^-(T)} \right]^{-1} \left[I - \frac{\dot{x}^-(T)R^T}{R^T \dot{x}^-(T)} \right] - I \right) r(T)}{R^T \dot{x}^-(T)}$$

коэффициент чувствительности полупериода автоколебаний к параметру α .

Обратимся к чувствительности устойчивости периодических движений. Для этого, помимо траектории $\tilde{x}(t)$, необходимо рассмотреть траекторию $\hat{x}(t) = \tilde{x}(t) + \delta\tilde{x}(t)$, представляющую собой результат «возмущения» траектории $\tilde{x}(t)$ малым изменением начального вектора состояния $\tilde{x}(0)$. Так же, как и $\tilde{x}(t)$, $\hat{x}(t)$ будем считать близкой к $x(t)$ (см. рисунок 4). Момент $t = 0$ на ней совместим с моментом переключения РЭ с минуса на плюс. Переключение РЭ с плюса на минус на траектории $\hat{x}(t)$ происходит в близкий к $T + \delta T$ момент времени $T + \delta T + \Delta T$.

Движение $\hat{x}(t)$ задается уравнением

$$(14) \quad \frac{d\hat{\mathbf{x}}}{dt} = \mathbf{C}(\alpha + \delta\alpha)\hat{\mathbf{x}} + \mathbf{B}(\alpha + \delta\alpha)u.$$

Из сравнения траекторий (3) и (14) несложно получить уравнение в вариациях (15)

$$(15) \quad \frac{d\delta\tilde{\mathbf{x}}}{dt} = \mathbf{C}(\alpha + \delta\alpha)\delta\tilde{\mathbf{x}}.$$

Из (15), пренебрегая величинами порядка малости выше первого, получим

$$(16) \quad \frac{d\delta\tilde{\mathbf{x}}}{dt} = \left(\mathbf{C}(\alpha) + \frac{\partial\mathbf{C}(\alpha)}{\partial\alpha}\delta\alpha \right) \delta\tilde{\mathbf{x}}.$$

Рассмотрим функцию $\Delta\tilde{\mathbf{x}}(T + \delta T) = \hat{\mathbf{x}}(T + \delta T + \Delta T) - \tilde{\mathbf{x}}(T + \delta T)$. С точностью до величин порядка малости не выше первого

$$(17) \quad \Delta\tilde{\mathbf{x}}(T + \delta T) = \dot{\mathbf{x}}^-(T)\Delta T + \delta\tilde{\mathbf{x}}^-(T + \delta T).$$

Обозначим $\mathbf{V}^*(t) = \exp\left(\left(\mathbf{C}(\alpha) + \frac{\partial\mathbf{C}(\alpha)}{\partial\alpha}\delta\alpha\right)t\right)$ нормированную фундаментальную матрицу решений уравнения (16). Тогда, очевидно,

$$(18) \quad \delta\tilde{\mathbf{x}}^-(T + \delta T) = \mathbf{V}^*(T + \delta T)\delta\tilde{\mathbf{x}}(0).$$

Подставляя (18) в (17), найдем

$$(19) \quad \Delta\tilde{\mathbf{x}}(T + \delta T) = \mathbf{V}^*(T + \delta T)\delta\tilde{\mathbf{x}}(0) + \dot{\mathbf{x}}^-(T)\Delta T.$$

Запишем теперь условие переключения РЭ на траектории $\hat{\mathbf{x}}(t)$ в момент $T + \delta T + \Delta T$

$$\mathbf{R}^T(\hat{\mathbf{x}}(T + \delta T + \Delta T)) = \mathbf{R}^T(\tilde{\mathbf{x}}(T + \delta T) + \Delta\tilde{\mathbf{x}}(T + \delta T)) = b.$$

Так как $\mathbf{R}^T\tilde{\mathbf{x}}(T + \delta T) = b$, из последнего равенства и (19) получим

$$(20) \quad \mathbf{R}^T\Delta\tilde{\mathbf{x}}(T + \delta T) = \mathbf{R}^T\mathbf{V}^*(T + \delta T)\delta\tilde{\mathbf{x}}(0) + \mathbf{R}^T\dot{\mathbf{x}}^-(T)\Delta T = 0.$$

Таким образом,

$$(21) \quad \Delta T = -\frac{\mathbf{R}^T\mathbf{V}^*(T + \delta T)\delta\tilde{\mathbf{x}}(0)}{\mathbf{R}^T\dot{\mathbf{x}}^-(T)}.$$

Принимая во внимание (13), (19) и (21), легко заметить, что

$$(22) \quad \Delta\tilde{\mathbf{x}}(T + K_\alpha^T\delta\alpha) = \left(\mathbf{I} - \frac{\dot{\mathbf{x}}^-(T)\mathbf{R}^T}{\mathbf{R}^T\dot{\mathbf{x}}^-(T)} \right) \mathbf{V}^*(T + K_\alpha^T\delta\alpha)\delta\tilde{\mathbf{x}}(0),$$

где \mathbf{I} – единичная матрица.

Назовем

$$(23) \quad \mathbf{G}(\alpha + \delta\alpha) = \left(\mathbf{I} - \frac{\dot{\mathbf{x}}^-(T)\mathbf{R}^T}{\mathbf{R}^T\dot{\mathbf{x}}^-(T)} \right) \mathbf{V}^*(T + K_\alpha^T\delta\alpha)$$

матрицей устойчивости «возмущенной» системы. Далее, перепишем равенство (22) в виде

$$\Delta\tilde{\mathbf{x}}(T + K_\alpha^T\delta\alpha) = \mathbf{G}(\alpha + \delta\alpha)\delta\tilde{\mathbf{x}}(0).$$

Очевидно, что для момента $t = 2T$ с учетом симметрии

$$\Delta \tilde{\mathbf{x}}(2(T + K_\alpha^T \delta \alpha)) = \mathbf{G}(\alpha + \delta \alpha) \Delta \tilde{\mathbf{x}}(T + K_\alpha^T \delta \alpha).$$

Вообще, имеет место равенство

$$(24) \quad \Delta \tilde{\mathbf{x}}((k+1)(T + K_\alpha^T \delta \alpha)) = \mathbf{G}(\alpha + \delta \alpha) \Delta \tilde{\mathbf{x}}(k(T + K_\alpha^T \delta \alpha)).$$

Очевидно, «возмущенное» периодическое решение $\tilde{\mathbf{x}}(t)$ является асимптотически орбитально устойчивым, если $\Delta \tilde{\mathbf{x}}(k(T + K_\alpha^T \delta \alpha)) \rightarrow 0$ при $k \rightarrow \infty$. Из теории линейных разностных систем известно, что тривиальное решение $\Delta \tilde{\mathbf{x}}(k(T + K_\alpha^T \delta \alpha)) \equiv 0$ уравнения (24) устойчиво, если собственные числа матрицы $\mathbf{G}(\alpha + \delta \alpha)$ по модулю меньше 1. Итак, пусть $\lambda_1, \lambda_2, \dots, \lambda_m$ – собственные числа матрицы $\mathbf{G}(\alpha + \delta \alpha)$. Периодическое решение $\tilde{\mathbf{x}}(t)$ релейной системы (3), (4) асимптотически орбитально устойчиво, если имеют место неравенства

$$(25) \quad |\lambda_j| < 1, j = \overline{1, m}.$$

С учетом ранее введенных обозначений перепишем

$$\mathbf{G}(\alpha + \delta \alpha) = \left(\mathbf{I} - \frac{\dot{\mathbf{x}}^-(T + K_\alpha^T \delta \alpha) \mathbf{R}^T}{\mathbf{R}^T \dot{\mathbf{x}}^-(T + K_\alpha^T \delta \alpha)} \right) e^{\left(\mathbf{C}(\alpha) + \frac{\partial \mathbf{C}(\alpha)}{\partial \alpha} \delta \alpha \right) (T + K_\alpha^T \delta \alpha)}.$$

Принимая во внимание величины порядка малости не выше первого, последнее выражение можно представить в виде

$$(26) \quad \mathbf{G}(\alpha + \delta \alpha) = \left(\mathbf{I} - \frac{(\dot{\mathbf{x}}^-(T) + \mathbf{K}_\alpha^{\dot{\mathbf{x}}^-(T)} \delta \alpha) \mathbf{R}^T}{\mathbf{R}^T (\dot{\mathbf{x}}^-(T) + \mathbf{K}_\alpha^{\dot{\mathbf{x}}^-(T)} \delta \alpha)} \right) e^{\left(\mathbf{C}(\alpha) T + \left(\frac{\partial \mathbf{C}(\alpha)}{\partial \alpha} T + \mathbf{C}(\alpha) K_\alpha^T \right) \delta \alpha \right)}.$$

Здесь

$$\mathbf{K}_\alpha^{\dot{\mathbf{x}}^-(T)} = -\mathbf{C}(\alpha) \mathbf{K}_\alpha^{\dot{\mathbf{x}}(0)} - \frac{\partial \mathbf{C}(\alpha)}{\partial \alpha} \mathbf{x}^*(T) + \frac{\partial \mathbf{B}(\alpha)}{\partial \alpha} A$$

– коэффициент чувствительности вектора скорости $\dot{\mathbf{x}}^-(T)$ к изменению параметра α .

С учетом величин порядка малости не выше первого

$$(27) \quad e^{\left(\mathbf{C}(\alpha) T + \left(\frac{\partial \mathbf{C}(\alpha)}{\partial \alpha} T + \mathbf{C}(\alpha) K_\alpha^T \right) \delta \alpha \right)} = e^{\mathbf{C}(\alpha) T} + \Theta \delta \alpha,$$

где

$$\Theta = \sum_{i=0}^{\infty} \frac{\sum_{j=0}^i (e^{\mathbf{C}(\alpha) T})^{i-j} \left(\frac{\partial \mathbf{C}(\alpha)}{\partial \alpha} + \mathbf{C}(\alpha) K_\alpha^T \right) (e^{\mathbf{C}(\alpha) T})^j}{(i+1)!}.$$

С учетом (27) равенство (26) примет вид

$$\mathbf{G}(\alpha + \delta \alpha) = \left(\mathbf{I} - \frac{(\dot{\mathbf{x}}^-(T) + \mathbf{K}_\alpha^{\dot{\mathbf{x}}^-(T)} \delta \alpha) \mathbf{R}^T}{\mathbf{R}^T (\dot{\mathbf{x}}^-(T) + \mathbf{K}_\alpha^{\dot{\mathbf{x}}^-(T)} \delta \alpha)} \right) (e^{\mathbf{C}(\alpha) T} + \Theta \delta \alpha).$$

Определим производную $\frac{\partial \mathbf{G}(\alpha)}{\partial \alpha} = \lim_{\delta \alpha \rightarrow 0} \frac{\mathbf{G}(\alpha + \delta \alpha) - \mathbf{G}(\alpha)}{\delta \alpha}$. Здесь $\mathbf{G}(\alpha)$ – матрица устойчивости системы (1), (2). Итак,

$$(28) \quad \frac{\partial \mathbf{G}(\alpha)}{\partial \alpha} = \left(\frac{\dot{\mathbf{x}}^-(T) \mathbf{R}^T (\mathbf{R}^T \mathbf{K}_\alpha^{\dot{\mathbf{x}}^-(T)} - \mathbf{K}_\alpha^{\dot{\mathbf{x}}^-(T)} \mathbf{R}^T (\mathbf{R}^T \dot{\mathbf{x}}^-(T)))}{\mathbf{R}^T \dot{\mathbf{x}}^-(T) \mathbf{R}^T \dot{\mathbf{x}}^-(T)} \right) e^{\mathbf{C}(\alpha) T} + \left(\mathbf{I} - \frac{\dot{\mathbf{x}}^-(T) \mathbf{R}^T}{\mathbf{R}^T \dot{\mathbf{x}}^-(T)} \right) \Theta.$$

Таким образом, матрица устойчивости для возмущенной системы определяется равенством

$$(29) \quad \mathbf{G}(\alpha + \delta\alpha) = \mathbf{G}(\alpha) + \frac{\partial \mathbf{G}(\alpha)}{\partial \alpha} \delta\alpha.$$

Пусть характеристическое уравнение матрицы устойчивости «невозмущенной» системы имеет вид

$$(30) \quad a_n \lambda_0^n + a_{n-1} \lambda_0^{n-1} + \dots + a_1 \lambda_0 + a_0 = 0,$$

где $a_k (k = \overline{0, n})$ – постоянные коэффициенты, λ_0 – номинальное собственное число. Анализируя зависимость (26), можно сделать вывод, что характеристическое уравнение матрицы $\mathbf{G}(\alpha + \delta\alpha)$ имеет следующую структуру

$$(31) \quad (a_n + v_n \delta\alpha)(\lambda_0 + \delta\lambda)^n + (a_{n-1} + v_{n-1} \delta\alpha)(\lambda_0 + \delta\lambda)^{n-1} + \dots + (a_1 + v_1 \delta\alpha)(\lambda_0 + \delta\lambda) + (a_0 + v_0 \delta\alpha) = 0.$$

Здесь $v_i (i = \overline{0, n})$ – постоянные коэффициенты, $\delta\lambda$ – вариация номинального собственного числа. Принимая во внимание величины, имеющие порядок малости не выше первого, из (30) и (31) получим равенство

$$n a_n \lambda_0^{n-1} \delta\lambda + v_n \lambda_0^n \delta\alpha + (n-1) a_{n-1} \lambda_0^{n-2} \delta\lambda + v_{n-1} \lambda_0^{n-1} \delta\alpha + \dots + a_1 \delta\lambda + v_1 \lambda_0 \delta\alpha + v_0 \delta\alpha = 0.$$

Отсюда следует, что

$$(32) \quad \delta\lambda = K_\alpha^{\lambda_0} \delta\alpha,$$

где

$$(33) \quad K_\alpha^{\lambda_0} = - \sum_{i=0}^n v_i \lambda_0^i / \sum_{k=1}^n k a_k \lambda_0^{k-1}.$$

Подставляя различные номинальные собственные числа в (32), легко найти их приращения, обусловленные изменением параметра α . Величину $K_\alpha^{\lambda_0}$ будем называть коэффициентом чувствительности номинального собственного числа λ_0 матрицы устойчивости $\mathbf{G}(\alpha)$ к изменению параметра α . Собственные числа матрицы $\mathbf{G}(\alpha + \delta\alpha)$ при этом задаются зависимостью

$$(34) \quad \lambda = \lambda_0 + K_\alpha^{\lambda_0} \delta\alpha.$$

На практике для определения коэффициентов v_i и a_k целесообразно воспользоваться функциями пакета символьных вычислений Symbolic Math Toolbox системы Matlab2008a.

Заключение

Разработанные методы позволяют весьма просто анализировать чувствительность релейной системы к изменению параметров объекта управления, что, в свою очередь, делает возможным выполнять синтез и оптимизацию системы, задаваясь ограничениями на ее чувствительность. Здесь следует иметь в виду, что если при оптимизации системы не контролировать ее чувствительность, то можно получить систему, которая (из-за высокой чувствительности) оказывается неработоспособной.

Список литературы

- [1] Методы классической и современной теории автоматического управления. Том 1. Математические модели, динамические характеристики и анализ систем автоматического управления. Под ред. К.А. Пупкова, Н.Д. Егупова. М.: Изд-во МГТУ им. Н.Э. Баумана, 2004. – 656 с.

BASIC PRINCIPLES OF INTELLECTUAL COMPUTER CALCULATORS CONSTRUCTION FOR REAL-TIME SIMULATION OF MANUFACTURING PROCESSES AND CONTROL

S.A. Kizhaev, A.K. Bulkhin, V.F. Klyuchnikov

Closed Joint Stock Company "Samara Cable Company"
9, Ul. Kabelnaya, 443022 Samara, Russia
kijaev@samaracable.ru; post-office@samaracable.ru
tel: +7 (846) 2282-455, fax: +7 (846) 333-27-70

Key words: gauges of direct and indirect measurement, intellectual computer calculators, methods, analytical methods, programs, adjustment, adaptation

Abstract

Предлагается альтернатива для датчиков прямого измерения в виде датчиков косвенного измерения, построенных на базе интеллектуальных компьютерных вычислителей параметров в сложных технологических процессах.

Introduction

With comprehensive automation of the manufacturing processes (MP) the important matter as regards the applied control theory is the matter connected with the basic principles of construction of gauges of indirect calculation, the due attention to which is not given so far [1,2].

In case of nonstop production – metallurgical production, papermaking, cable manufacturing, textile manufacturing, glasswork, etc. – the direct measurements of process parameters refer to the direct measuring methods.

With intensive development of manufacturing processes automation in cable industry in 1970–1980 there was lack of gauges for measuring manufacturing processes parameters both in this country and abroad.

Therefore in parallel with any automation the respective gauges were being developed for measuring the required parameters. Such way of solving the matter resulted in significant protraction of the rates of development of manufacturing processes automation.

The gauges which were developed were extremely expensive. At present in Russia the gauges of parameters to be measured are mainly purchased abroad, they are high-priced and need expensive maintenance.

1 Task Description

The gauges of direct measurement of parameters have mainly one benefit – high-precision measurement.

However, they have a lot of shortcomings – large size, extremely high price, criticality to temperature conditions of the equipment with heating zones, impossibility of installation in desired places, complexity, expensive maintenance, etc.

The required gauges are often not available at all, but automation of manufacturing processes should be carried out in any case.

About 200 process parameters are not measured but should be measured: molten viscosity of polymers; specific heat capacities of process materials; porosity coefficient; lay length of cable films applied by winding method; primary and secondary electric parameters of cables and wires, and others.

It is possible to use the method of calculation for measuring such parameters.

2 Solution Algorithm

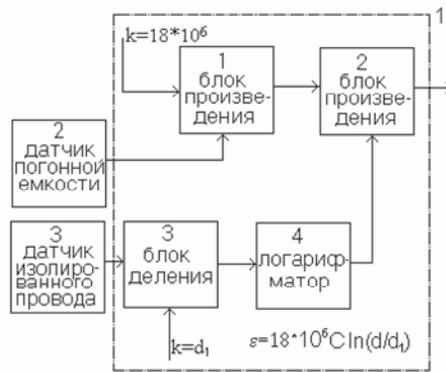
2.1 Basic Principles of Construction of Gauges of Indirect Measurement of Parameters

Formerly the design of gauges of indirect measurement was based on shared circuit elements (microchips) (Figure 1-2).

Later on it was based on analog single-chip microcomputers (Figure 3) [3]. Both options have disadvantages - they are difficult to be transformed.

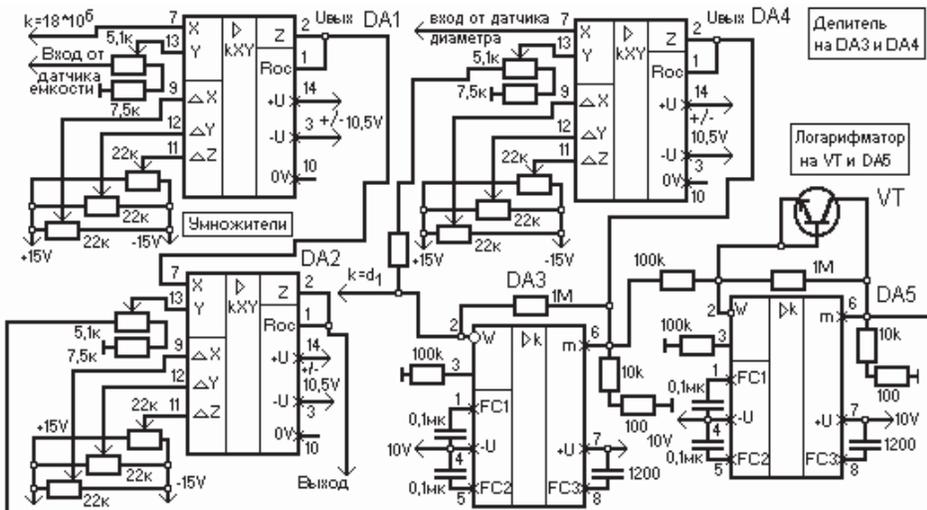
At present they can be designed on the basis of reprogrammable controllers with programmed calculation.

The calculation program can be recorded in the controller which carries out the control of the complicated manufacturing process.



1 – product unit; 2 – product unit; 2 – gauge of capacity per unit of length; 3 – gauge of insulated wire; 3 – divider; 4 – logarithmator

Figure 1 – Functional Arrangement of Computer Calculator of Dielectric Constant



DA1– DA2, D4 – signal multipliers K525 IC, DA3, DA5 – amplifiers 140 УД 2А

Figure 2 – Circuit Diagram of Computer Calculator of Telephone Wire Dielectric Constant

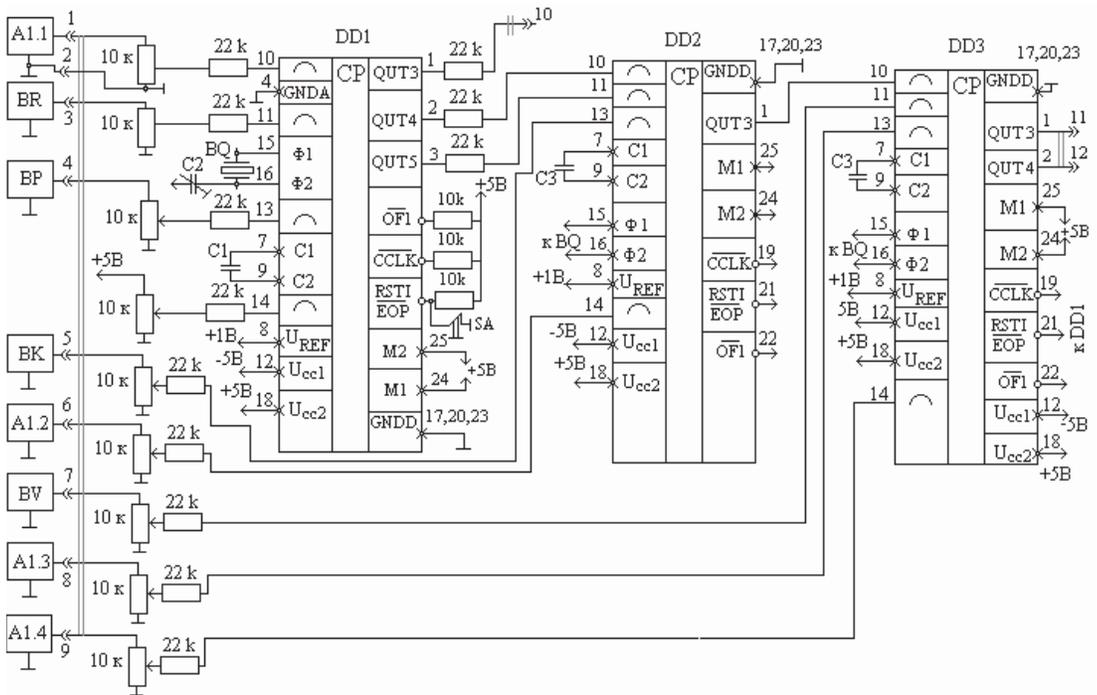


Figure 3 – Circuit Diagram of Computer Calculator of Gauge for Measuring Polymer Viscosity, Insulation Diameter, and Shrinkage Factor in Cooling Trough

In Figure 3: DD1– DD3 – analog single-chip microcomputers KM 1813 BE1 with analog inputs and outputs; BQ – quartz-crystal resonator for 6.67 MHz, start toggle switch of ПТ -57 type; C1– C3 – capacitors K73-9, 1000 pF; C2 – adjustable capacitor КТЧ–27-25 В-04/2; A1.1 – A1.4 – gauge of the extruder electric motor power, gauge of the extruder barrel zones heating, gauge of diameter of half-finished cable product, and gauge of diameter of insulation of the cable produced, respectively; BR, BV – gauges of extruder screw rotation speed and of cable speed; BP, BK – gauges of polymer pressure in press head and of polymer hot melt temperature.

Basic principles of construction of gauges of indirect measurement – intellectual reprogrammable computer calculators (IC): analytical methods of description of complicated manufacturing processes using physical, chemical, and biological laws. For this purpose it is necessary to study these processes properly.

In [3, 4] these basic principles are based on 6 methods of description of manufacturing process objects in industrial production:

- 1) *Directly on the basis of heat-balance and mass-balance equations*, when the required output value is directly included in the said equation (mass of materials processed, overall dimensions, temperature, heat capacity, electric parameters, processing time).
- 2) *By means of graphical interpretation* for parameters not included in the balance equations either directly or indirectly (beating rate of paper stock or insulation porosity in extruders).
- 3) *By means of mathematical transformation*, when the required output parameter is not included in the description of the balance equation, too, but it is connected mathematically with one of the values (elongation of wire, porosity of impregnated paper insulation);
- 4) *By means of analogs* (extrusion process and cable winding process);
- 5) *Based on physical laws* (process of cable drying);
- 6) For more complicated processes it necessary to *use combined method*.

In Table 1 you can see the ontology of methods of application of basic principles of analytical description of complicated technical processes applied by the authors in cable production, and references to the cited literature where their application is described in detail.

Table 1 – Ontology of Methods of Application of Laws of Conservation of Mass and Energy and Similarity Rule When Describing MP Objects, & References to Cited Literature

No.	Description of methods of application	Description of computable parameters	References to cited literature
1.	Direct calculation from heat-balance and mass-balance equations	a) mass of materials processed, b) processing temperature, c) heat capacity, d) electric parameters (power, current and voltage), e) processing time	[3, 4,5]
2.	Mathematical transformations from heat-balance and mass-balance equations	a) elongation of conductors during their annealing in furnaces or in the course of drawing, b) porosity of paper porous and polymeric insulation	[6,7]
3.	Based on graphical interpretation of mass transfer process	a) porosity of insulation, b) tape overlap pitch when applying coverings, c) beating rate of paper stock	[5,8, 9]
4.	Based on similarity in geometry, masses, moment of inertia	a) tape overlap pitch, b) thickness of bitumen applied, c) diameter of conductor insulation	[10,11,12]
5.	Based on physical laws	a) time of cable drying, b) time of furnace annealing, c) impedance	[3,8,11,12]
6.	Combined application	a) diameter of aluminum sheath, b) parameters of extrusion process, c) impedance	[10,13],

2.2 Application of Analytical Methods of Description

As an example please find below the application of analytical method of description based on heat-balance and mass-balance equations when the required output value is included directly in the said equation for MP of polymer extrusion (material-balance equation) when manufacturing wires/cables.

The process of polymer extrusion consists in loading granules of polymer to the heated extruder from where the polymer is extruded by means of screw and applied to the running wire or cable.

It is recorded that insulation mass extruded by screw from the barrel is determined statically in unit time when getting in the extruder head:

$$(1) \quad m_e = \frac{\pi}{4} \gamma (D_b^2 - D_{scr}^2) \frac{l}{t} = \frac{\pi}{4} \gamma (D_b^2 - D_{scr}^2) \cdot v_1 = \frac{\pi}{4} (D_b^2 - D_{scr}^2) \cdot \omega \cdot R \cdot \gamma,$$

where γ , l – its density and length of extrusion section, D_b, D_{scr} – diameters of barrel and screw,

v_1, ω – speed of mass and angular speed of the extruder screw rotation,

$R = D_{scr} / 2$ – radius of area of the mass extrusion from the extruder head.

Mass of insulation applied to the wire/cable is determined in unit time when coming out of the head:

$$(2) \quad m_h = \frac{\pi}{4} \gamma (D^2 - d^2) v$$

where D, d – diameter of wire/cable over insulation and initial diameter of wire/cable respectively;
 v – speed of wire/cable running.

Based on law of conservation of mass the equation is set up statically:

$$(3) \quad m_e = m_h = \frac{\pi}{4}(D_y^2 - D_{uni}^2) \cdot \omega \cdot R \cdot \gamma = \frac{\pi}{4} \gamma (D^2 - d^2) v$$

The expression for diameter of wire/cable is determined statically out of (3) after the insulation is applied:

$$(4) \quad D = \sqrt{k_o \frac{\omega}{v} + d^2},$$

where $k_o = (D_b^2 - D_{scr}^2) R$ – coefficient of proportionality.

2.3 Programs for Calculating Parameters

Upon derivation of coupling equations of output value (4) required for indirect measurement and control it is necessary to translate them to calculation programs for controller or industrial computer carrying out control of MP. It's the authors' opinion that these programs should possess the following performance capabilities:

- readjustment during the MP running with some other materials or modes used;
- generality for calculating parameters on similar equipment;
- adaptability, i.e. capability of adjustment from gauges of direct measurement.

Readjustment is carried out by the line operator according to the recipes given by the Process Engineer.

Generality is ensured by changing the equipment characteristics and is carried out by the Equipment Engineer.

Adaptability is downloaded in the program by IC designer and is ensured automatically by MP automatic control system.

Please find below the devices of indirect measurement working on principle of IC program calculation. It is necessary to note that IC of technological parameters can be used either directly or in prediction mode.

In case the gauges of direct measurement are not available, or there is no possibility for their installation because of the dimensions or environmental conditions, the IC of indirect measurement are used.

When the gauges of direct measurement are available, intellectual computer calculators-observers may be used as predictors. In the third option, in case the gauges of direct measurement are very expensive or the delivery term is too long, the intellectual devices of indirect measurement may be also used.

So IC is a device designed on the principle of mathematical relationship calculation of output value connecting input parameters which are measurable. Formally it is a virtual device which calculates the required value according to the program developed.

Therefore along with mathematical description the programming operation is also important.

By **IC tuning coefficient** k_{ic} we shall mean coefficient of proportionality, appeared in the coupling equation on which basis it functions. Here two options are possible:

- for coupling equations with explicit coefficient of proportionality by which we shall mean natural tuning coefficient (for equation (5) – k_o),
- for coupling equations with implicit coefficient of proportionality new coefficients of proportionality are introduced artificially.

In general form it is possible to specify for IC instantaneous values:

$$(5) A_{ic_i}(t) = k_{t ic_i}(t) \cdot f\{A_i(t)\},$$

where $A_{ic_i}(t)$ – effective value at the output of the intellectual observer,

$f\{A_i(t)\}$ – coupling equation between IC output and input values.

From (5) we get:

$$(6) k_{t ic_i}(t) = \frac{A_{ic_i}(t)}{f\{A_i(t)\}}$$

For IC having analogs in the form of gauges of direct measurement we introduce the notion of **adaptation coefficient**:

$$(7) k_{a_i}(t) = \left| \frac{A_{ic_i}(t)}{A_{e_i}(t)} \right|$$

where $A_{e_i}(t)$ – effective values at the output of the gauge of direct measurement. Ideally in case of signals coincidence the adaptation coefficient $k_{a_i}(t) \rightarrow 1$.

Let us consider the connection between the tuning and adaptation coefficients. Having solved jointly (6 and 7) we get:

$$(8) k_{t ic_i}(t) = \frac{k_{a_i}(t) A_{e_i}(t)}{f\{A_i(t)\}}.$$

Let us determine by way of example the connection of IC adaptation and tuning coefficients for equation (4), having solved it with regard to (8) for instantaneous values of parameters introduced in it:

$$(9) k_{t ic_i} = k_{o_i}(t) \frac{\{ [|k_{a_i}(t)| D_{e_i}(t)]^2 - d^2 \} v_i(t)}{\omega_i(t)}.$$

So it is possible to adapt parameters of IC programmatically by tuning it to fit the changeable operation conditions of the MP equipment.

Manufacturing processes computation programs may be developed both for master controller and industrial computer which carry out MP control, and for local one meant only for calculation of manufacturing parameters.

Table 2 displays the program for calculating the cable/wire diameter executed on the base of Simatic-300 controller, which carries out the control of complicated manufacturing process in ME-160 line.

Table 2 – Program for Calculation Insulation Diameter in extrusion lines [14]

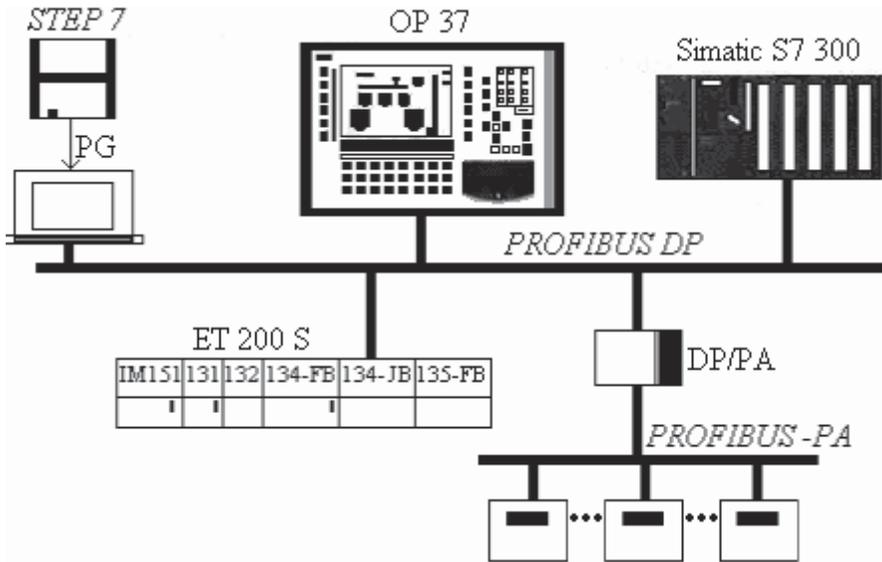
No.	Type of operation	Value	Comments
1	L	ω	loading ω
2	L	v	loading v
3	/	R	result ω/v
4	L	k_o	loading k_o
5	X	R	result $k_o \omega/v$
6	T	DB1.DB D _o	transfer $k_o \omega/v$
7	L	d	loading d
8	SQR	d	result d^2
9	L	DB1.DB D _o	loading $k_o \omega/v$

The end of table 2

10	+	R	result $d^2 + k_o \omega / v$
11	T	MD 100	transfer $d^2 + k_o \omega / v$
12	SQRT	MD 100	result $D = \sqrt{d^2 + k_o \omega / v}$
13	T	MD 104	transfer $D = \sqrt{D_o^2 + k_o \omega / v}$

2.4 Unified Functional Diagram of Intellectual Computer Calculator

Figure 4 displays the proved functional diagram of manufacturing parameter IC on local controller. The programs are executed in Step 7 language for Simatic S7 controllers, but they can be executed with any controller.



STEP7 – programming language, OP 37 – operator’s panel, Simatic S7 – controller, PG – programmer, ET 200 S – input- output station, DP/PA – coupler

Figure 4 – Unified Functional Diagram of Manufacturing Parameters IC

3 Economics

The program for calculating, for example, viscosity of polymers during extrusion, have made it possible to introduce the measurement of the parameter, which was not measured before, but which is connected with the products quality through insulation breakdowns, and as a result the quantity of such breakdowns can be reduced by (3.8–4.5) %. The program of calculating elongation of copper/aluminum wires during annealing has made it possible to reduce their stickiness by (6.3–7.1) %.

Conclusion

The offered method/basic principles of indirect measurements of parameters using general physical properties of the objects of different nature with preservation of energy, mass, linear momentum make it possible to obtain the required results when gauges of direct measurement are not available. This method of measuring parameters also makes it possible to attract the methods of MP

modeling based on computer control of complicated technical systems with the purpose of adaptation of the obtained results to the changeable operating conditions of these systems.

Literature

- [1] Author's Certificate 875206 USSR, IPC³ G01 B 7/04. Device for Measuring Parameters of Conductors Winding up / S.A. Kizhaev – No. 2804142/18 – 28; Application 30.07.79; issued 23.10.81, Bulletin No. 39. – 6 p.: il.
- [2] Vydrin V.I., Fedosienko A.S. Automation of Rolling Process.– M.: Metallurgy, 1984.–386 p.
- [3] Author's Certificate 1837266 USSR, IPC⁵ G 05 D 22/02. Control System of Moisture Insulation of Conductors / S.A. Kizhaev. – No. 4668700/24; Application 28.03.89; issued. 30.08.93, Bulletin No. 32. – 8 p.: il.
- [4] Kizhaev S.A. Comprehensive Automation of Impregnated Paper Plant. – Samara: Samara University, 2003. – 421 p.
- [5] Kizhaev S.A. Automation of Drying Cables with Paper Insulation // Automation in Industry. – 2004. – No. 2. – P. 3–5.
- [6] Kizhaev S.A. Automation of Annealing Process of Wire in Furnaces // Automation in Industry. – 2004. – No. 3. – P. 11–13.
- [7] Patent 2022321 RF, IPC⁵ G 05 D 11/13. System Adjustment of Porosity of Impregnated Paper Insulation of Cable Group / S.A. Kizhaev. – No. 4758115/24; Application 13.11.89; issued 30.10.94, Bulletin No. 20. – 6 p.: il.
- [8] Kizhaev S.A. Intellectual Systems of Calculation in the Process of Extrusion in Cable Industry // Gauges and Systems. – 2004. – No. 4. – P. 8–14.
- [9] Kizhaev S.A. Analytical Methods of Synthesis of Automatic Control Systems of Manufacturing Processes. – Samara: NTC, 2006. – 97 p.
- [10] [10] Bulkhin A.K., Kidyayev V.F., Kizhaev S.A. Automation of the Process of Cable Armouring // Cables and Wires. – 2003. – No. 1. – P. 21–23.
- [11] Kizhaev S.A. Gauges of Process Parameters in Cable Industry: interacademic collection of scientific works "Gauges of Measuring Systems, Monitor and Control". – Penza : PPI, 1991. – P. 103–107.
- [12] Kizhaev S.A. Mathematical Models and Methods of Formalization in the Process of Automation of Manufacturing Processes for Their Integration in CALS-Technologies // Materials of the 2nd All-Russia Research and Practice Conference "Computer Integration of Production and IPI Technologies". – Orenburg : IPK OGU, 2005. – P. 85–95.
- [13] Grodnev I.I., Frolov P.A. Coaxial Communication Cables. – Svyaz, 1970.– 312 p.
- [14] Certificate 2086143. Program for Calculating Wire/Cable Diameter / S.A. Kizhaev. – No. 2008613145, Application 09.08.2008, registered Bulletin "Programs for Computers. Data Bases. Integrated Circuit Technologies, No. 3, 2008.

Biography

- 1) Кизаев Станислав Алексеевич – канд. техн. наук, Заслуженный изобретатель Российской Федерации, лауреат губернской премии в области науки и техники.
- 2) Бульхин Анвар Кашафович – канд. техн. наук, профессор Самарского государственного технического университета, Заслуженный машиностроитель Российской Федерации, лауреат губернской премии в области науки и техники.
- 3) Ключников Валерий Федорович – канд. техн. наук, доцент Самарского аэрокосмического университета им. Королева.

РАЗРАБОТКА МЕТОДА ОЦЕНКИ АДЕКВАТНОСТИ СИНТЕЗИРОВАННЫХ МОДЕЛЕЙ РЕАЛЬНЫМ ОБЪЕКТАМ В СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМАХ, ИСПОЛЬЗУЮЩИХ РАЗНЫЕ ВИДЫ ЭНЕРГИИ

С.А. Кижяев, А.К. Бульхин, В.Ф. Ключников

Закрытое акционерное общество «Самарская кабельная компания»
443022, Самара, ул. Кабельная, 9, Россия
kijaev@samaracable.ru; post-office@samaracable.ru
тел: +7 (846) 2282-455, факс: +7 (846) 333-27-70

Ключевые слова: оценка, адекватность, модель, погрешность, комплексная автоматизация, технологические процессы, планирование экспериментов, пассивные и активные эксперименты, контроллер, коррекция, датчики косвенного измерения, коэффициент влияния

Abstract

The method of an estimation of adequacy of models to the real objects, constructed on the experiments spent by means of computer programs directly in an operating time of the equipment with adaptation of their parameters is proved.

Введение

При комплексной автоматизации технологических процессов (ТП) важным с точки зрения прикладной теории управления является метод оценки адекватности синтезированных моделей реальным объектам для ТП, использующих разные виды энергии и движения. В настоящее время используются старые методы оценки, построенные на пассивных и активных способах, которые сложны и занимают много времени на эксперименты, или уменьшают машинное время оборудования.

Авторы при автоматизации ТП исследовали и применяли все существующие способы и методы оценки адекватности моделей. В основном в инженерной практике применялись экспериментальные методы, которые позволяют получить простые по виду описания объектов, но часто не подходящие для многомерных многосвязных процессов.

Основным недостатком экспериментальных методов является неприемлемо большое время для постановки экспериментов [1,2], особенно сейчас в условиях рыночной экономики.

В работе [3] описание объектов управления (ОУ) производилось экспериментально, частично активными, частично пассивными методами, например, для смесителя, сетчатого барабана, гладильного устройства. Активные или пассивные методы экспериментов определялись временем работы оборудования и сроками проектирования, но в обоих случаях необходимо было тщательно планировать их проведение. Поэтому и применялись старые проверенные методы оценки адекватности моделей объектам.

1 Постановка задачи

Однако из-за недостатка времени, отводимого на комплексную автоматизацию ТП, приходится снова вернуться к построению математических моделей аналитическим способом на основе физических законов их действия.

В кабельной промышленности, как показывает опыт авторов, применяются в основном законы сохранения энергии, массы и количества движения [3,4, 5]. В связи с этим необходимо разработать и новые методы оценки адекватности моделей объектам управления, использую-

щих разные виды энергии, и произвести их онтологию. Нужно уменьшить время на эксперименты и производить их прямо во время работы оборудования.

1.1 Оценка точности существовавшими способами для тепловых процессов

Технологический процесс – наложение алюминиевых сварных защитных оболочек

Для этого ТП получена статическая зависимость на основе закона сохранения массы для ширины зоны сварки a [3]:

$$(1) \quad a = \frac{P\eta l}{v c_m \gamma h l \Delta T} - \frac{\gamma \pi (D_{об}^2 - D_k^2) l}{4 c_m \gamma h l \Delta T} = k_1 \frac{P}{v h l \Delta T} - k_2 \frac{(D_{об}^2 - D_k^2)}{h \Delta T},$$

где $k_1 = \frac{\eta}{c_m \gamma}$, $k_2 = \frac{\pi}{c_m \gamma}$ – коэффициенты пропорциональности, c_m – удельная теплоемкость металла, P – мощность, необходимая для сварки при скорости v на длине участка l , η – КПД сварки токами высокой частоты, γ – плотность металла оболочки кабеля, h – толщина ленты для сварки оболочки, $D_{об}, D_k$ – диаметры металлической оболочки и сердечника кабеля.

Оценка точности выражения (1) производилась методом активного эксперимента. Результаты эксперимента представлены в таблице 1.

Таблица 1 – Оценка точности модели по зависимости (1)

№ п/п	Наименование параметра	Относительная погрешность
1.	мощность сварки P , кВт	4,8 %
2.	толщина ленты h , мм	5,38 %
3.	ширина ленты l , мм	6,67 %
4.	скорость сварки v , м / мин	6,76 %

Делается вывод об относительной погрешности модели ε_i по Y_{mi} – выходному ее параметру. Ее значение определяется $\varepsilon_i = (Y_{ii} - Y_i) / Y_i$, где Y_i – тот же параметр, имеющий место в реальном объекте.

Исследование зависимости (1) велось также статистическим методом по зависимости [3]:

$$(2) \quad a = b_1 P - b_2 T - b_3 h - b_4 v + b_5 D_k,$$

где $b_1 = 0,000076$ см/кВт; $b_2 = 0,000036$ см/град; $b_3 = 0,01$; $b_4 = 0,01$ мин/см; $b_5 = 0,02$.

Полученная относительная погрешность в этом случае не превысила 4,76 % (таблица 2).

Таблица 2 – Оценка точности модели по зависимости (2)

№ п/п	Наименование коэффициента	Относительная погрешность
1.	b_1	3,56 %
2.	b_2	2,78 %
3.	b_3	3,47 %
4.	b_4	4,76 %

Для технологического процесса экструзии получены следующие передаточные функции:

$$(3) \quad W_1(p) = \frac{(D^2 - d^2)(p)}{\omega(p)} = \frac{k_1}{T_1(p) + 1};$$

$$(4) \quad W_2(p) = \frac{(D-d)(p)}{\omega(p)} = \frac{k_2}{T_1(p)+1};$$

$$(5) \quad W_3(p) = \frac{(D^2 - d^2)(p)}{v(p)} = -\frac{k_3}{T_2(p)+1};$$

$$(6) \quad W_4(p) = \frac{(D-d)(p)}{v(p)} = -\frac{k_4}{T_2(p)+1}.$$

При этом оценку точности соответствия модели объекту необходимо осуществлять на трех этапах:

- 1) оценка модели объекта в статике, например, для зависимости (1);
- 2) оценка непосредственно сравнением характеристик модели и объекта путем активного эксперимента для зависимостей (3-6);
- 3) оценка сравнением характеристик внедренной системы ее заданным характеристикам по оценкам близости параметров реального и спроектированного регулятора.

В исследованиях использовались первый и третий этапы, как более простые, а третий этап необходим в любом случае для доводки системы. С появлением компьютерных систем моделирования появилась простая возможность производить промежуточное моделирование для сравнения точности модели на всех этапах синтеза систем автоматического управления.

Планируется проведение экспериментов для построения переходные характеристики для объектов управления, описанных зависимостями (3 – 6). При этом оно осуществляется следующим образом. Предполагая, что при загрузке $G = 60$ кг полимера, скорости изолирования $v = 2$ м /с и при угловой скорости $\omega = 2$ с⁻¹ экструдер МЕ-60 можно считать линейным и стационарным.

При этом для канала «угловая скорость экструзии – сечение изоляции» необходимо:

- 1) перед проведением экспериментов изолировать систему от случайных воздействий в виде количества загружаемой полимерной массы, которая готовится на прецизионных цифровых весах с точностью 0,01% и изменения скорости изолирования;
- 2) входное воздействие – угловая скорость экструзии – принять $\omega = 2$ с⁻¹ (100 % воздействия в виде единичного сигнала);
- 3) количество опытов определить как $10 > 2^3 + 1$, где $n = 3$ – число входов многомерного объекта управления экструдера (угловая и линейная скорости и температура T);
- 4) отключить регулятор диаметра;
- 5) в других каналах управления наоборот системы стабилизации температуры полимера, расхода охлаждающей воды и скорости изолирования настраиваются на высокую точность стабилизации указанных параметров;
- 6) для удешевления опытов они производятся по понедельникам при запуске каждой из экструзионных линий в работу.

Аналогично, предполагая, что при скорости изолирования $v = 2$ м/с $G = 60$ кг полимера при скорости изолирования $v = 2$ м /с и при угловой скорости $\omega = 2$ с⁻¹ экструдер МЕ-160 можно считать линейным и стационарным.

При этом для канала «линейная скорость экструзии – сечение изоляции» необходимо:

- 1) перед проведением экспериментов изолировать систему от случайных воздействий в виде количества загружаемой полимерной массы, которая готовится на прецизионных цифровых весах с точностью 0,01% и изменения угловой скорости экструдера;
- 2) входное воздействие – линейная скорость экструзии – принять $v = 2$ м /с (100 % воздействия в виде единичного сигнала);

- 3) количество опытов определить как $10 > 2^3 + 1$, где $n = 3$ – число входов многомерного объекта управления экструдера (угловая и линейная скорости и температура T);
- 4) отключить регулятор диаметра;
- 5) в других каналах управления наоборот системы стабилизации температуры полимера, расхода охлаждающей воды и угловой скорости экструзии настраиваются на высокую точность стабилизации указанных параметров;
- 6) для удешевления опытов они также проводятся по понедельникам при запуске каждой из экструзионных линий в работу.

После этого снимаются графики переходных процессов, на основании которых определяется относительная погрешности проведенных экспериментов.

В таблице 3 приведены относительные погрешности моделирования на основе законов сохранения массы и энергии по выражениям (3-6).

Эксперименты проводились при следующих параметрах:

длина шнека $L = 1,2 \text{ м}$; диаметр шнека $D_{\text{шн}} = 60 \text{ мм}$; диаметр цилиндра $D_{\text{ц}} = 63 \text{ мм}$; мощность двигателя экструдера $P_{\text{дв}} = 32 \text{ кВт}$; мощность нагрева экструдера $P_{\text{н}} = 12 \text{ кВт}$; время нагрева массы $t_{\text{н}} = 3600 \text{ с}$; массы шнека и цилиндра $m_{\text{ц}} + m_{\text{ш}} = 521 \text{ кг}$; удельная теплоемкость полимера

$c_n = 1050 \frac{\text{Вт} \cdot \text{с}}{\text{кг} \cdot \text{град}}$; теплоемкость металла шнека и цилиндра экструдера $c = 460 \frac{\text{Вт} \cdot \text{с}}{\text{кг} \cdot \text{град}}$;

плотность полимера $\gamma_n = 1400 \frac{\text{кг}}{\text{м}^3}$; диаметр изоляции провода $D_{\text{max}} = 6 \text{ мм}$; диаметр заготовки

$d = 1 \text{ мм}$; угловая скорость шнека экструдера $\omega = 2 \text{ 1/с}$; линейная скорость изолирования $v = 2 \text{ м/с}$; температура переработки полимера $T = 150 \text{ град С}$; коэффициенты пропорциональности, определяемые конструкцией экструдера $k_0 = 5833 \text{ мм}^3$; $k_6 = 437,5 \text{ мм}^2 \text{ с град}$.

Таблица 3 – Оценка точности модели по зависимостям (3-6)

№ п. п.	Передаточная функция	Относительная погрешность
1.	$W_{\omega 1}(p) = \frac{(D^2 - d^2)(p)}{\omega(p)} = \frac{2,92}{0,43p + 1}$	4,42% по коэффициенту передачи, 2,23 % по постоянной времени
2.	$W_{v1}(p) = \frac{(D^2 - d^2)(p)}{v(p)} = \frac{17,5878}{0,087p + 1}$	5,57% по коэффициенту передачи, 2,56 % по постоянной времени
3.	$W_{T1}(p) = \frac{(D^2 - d^2)(p)}{T(p)} = \frac{0,219}{145,32p + 1}$	6,12% по коэффициенту передачи, 1,89 % по постоянной времени
4.	$W_{\omega 2}(p) = \frac{(D - d)(p)}{\omega(p)} = \frac{0,21}{0,38p + 1}$	3.84% по коэффициенту передачи, 2,01% по постоянной времени
5.	$W_{v2}(p) = \frac{(D - d)(p)}{v(p)} = \frac{0,419}{0,088p + 1}$	5.48% по коэффициенту передачи, 3,01% по постоянной времени
6.	$W_{T2}(p) = \frac{(D - d)(p)}{T(p)} = \frac{0,033}{136,23p + 1}$	6.04% по коэффициенту передачи, 1,09% по постоянной времени

1.2 Оценка точности существовавшими способами для процессов передачи веществ вентиляторами и насосами

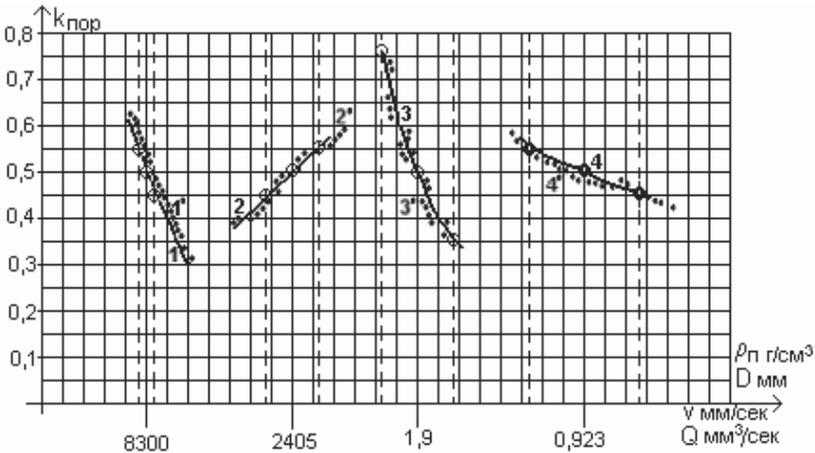
ТП изготовления проводов, кабелей с полимерной изоляцией на экструзионных линиях

В [7] приведена зависимость коэффициента пористости $k_{пор}$ полимерной изоляции при ее вспенивании газом (азотом) на экструзионных прессах:

$$(7) \quad k_{пор} = k_o \frac{Q}{v(D^2 - d^2)} \cdot \frac{\rho_z}{\rho_n},$$

где k_o , D , d , Q , v , ρ_z , ρ_n – соответственно коэффициент пропорциональности; диаметр провода по изоляции, начальный диаметр провода, расход газа, скорость изолирования, плотность газа и полимера.

Оценим точность этого выражения, полученного на основании графической интерпретации закона сохранения массы полимера при наложении на металлический проводник. Оценка точности производилась методом активного эксперимента и представлена в графиках на рисунке 1, где кривые с индексами 1–4 получены аналитически на основании зависимости (7), а кривые с индексами (1¹–4¹) – это экспериментальные кривые, полученные на действующем оборудовании.



(1,1¹) – скорость изолирования v , (2,2¹) – расход газа Q ,
 (3,3¹) – диаметр изоляции D , (4,4¹) – плотность полимера ρ_n

Рисунок 1 – Графики зависимостей коэффициента пористости

Оценка точности выражения (7) представлена в таблице 4.

Таблица 4 – Оценка точности модели по зависимости (7)

№ п/п	Наименование параметра	Относительная погрешность
1.	диаметр изоляции провода D , мм	6,28 %
2.	расход массы Q , мм ³ / с	2,31 %
3.	плотность полимера ρ_n , Г / см ³	5,761 %
4.	скорость изолирования v , мм / с	2,823 %

1.3 Анализ оценки адекватности полученных моделей

Анализируя полученные результаты, можно сделать вывод, что метод аналитического описания ТП на основе применения физических законов функционирования объектов в сложных технических системах позволяет с достаточной точностью для практического уровня производить их алгоритмизацию.

Однако эта точность не соответствует сегодняшнему уровню требований.

После проведенного анализа и реального применения методов оценки точности моделей и датчиков косвенного измерения (ИВ) технологических параметров с помощью старых методов авторы в таблице 5 изложили свое видение на возможность их применения.

Таблица 5 – Виды экспериментов по оценке точности моделей

№ п. п	Вид эксперимента	Тип ТП	Используемый метод	Вид воздействия
1.	Активный	1.1.Тепловой	Метод переходных характеристик	Единичный
		1.2.Передача твердых, жидких и газообразных веществ вентиляторами и насосами		Единичный
		1.3.Передача механического движения вращением		Единичный при мощности привода $P \leq 15$ кВт; 0,5 единичного при $15 \leq P \leq 35$ кВт; 0,2 единичного при $35 \leq P \leq 100$ кВт
2.	Пассивный	Передача механического движения вращением	Статистические методы	При мощности привода $P \geq 100$ кВт

При этом результаты не удовлетворили исследователей, т.к. это было фактически применение старых методов к современному оборудованию.

2 Алгоритм решения

Поэтому была поставлена задача разработать новый метод оценки моделей и ИВ, который бы одновременно использовал базу аналитического описания моделей (ИВ) и их коррекцию в результате работы без остановки процесса качественного изготовления продукции. Предлагается использовать контроллер с автоматической или ручной коррекцией. В исследованиях об основах построения ИВ отмечалось, как производится адаптация моделей (ИВ) по результатам показаний датчика прямого измерения данного технологического параметра.

При отсутствии датчика прямого измерения необходим другой вариант коррекции параметров модели. Рассмотрим это на примере оценки адекватности модели ТП экструзии. Для нее получена формула связи входных и выходного параметра – диаметра провода, кабеля.

$$(8) D = \sqrt{D_o^2 + k^1 \omega / \nu T_p},$$

где D , D_o , ω , v , T_p – соответственно диаметры кабеля и его заготовки, угловая частота вращения шнека экструдера, скорость изолирования и температура расплава изоляции, k^1 – коэффициент пропорциональности.

При этом возможны 2 варианта коррекции модели.

2.1 Вариант 1

В нем прямо во время работы в контроллере по программе вычисляется значение диаметра на выходе экструдера, а через определенные интервалы времени при измеряемых значениях ω , v , T_p вычисляется коэффициент пропорциональности k^1 . Затем его значение подставляется в формулу (4.8) и вычисляется скорректированное значение диаметра D .

В общем виде можно записать для начального вычисления значения диаметра в момент времени t_i :

$$(9) D_i(t_i) = \sqrt{D_o^2 + k_i^1(t_i)\omega_i(t_i)/v_i(t_i)T_{pi}(t_i)}$$

Затем автоматически вычисляется коэффициент пропорциональности в момент времени t_{i+1} :

$$(10) (k_{i+1}(t) = \frac{[D_{i+1}^2(t) - D_o^2] \cdot v_{i+1}(t) \cdot T_{pi+1}(t)}{\omega_{i+1}(t)}$$

После чего вычисляется новое скорректированное значение диаметра $D_{i+1}(t_{i+1})$ провода, кабеля.

Оптимальный интервал времени вычисления должен быть равен самой меньшей постоянной времени изменения входных параметров ω , v , T_p .

2.2 Вариант 2

Второй вариант оценки адекватности модели методом активного эксперимента на базе контроллера и ее коррекции применяется не во время работы оборудования.

Для определения влияния каждого входного параметра на выходной общий коэффициент пропорциональности k^1 заменяется коэффициентом пропорциональности:

$$(11) k^1 = \frac{k_\omega}{k_v \cdot k_{T_p}}$$

где коэффициенты k_ω , k_v и k_{T_p} назовем **коэффициентами влияния входных величин на выходные**.

Производится изменение, например, входной величины с ω_i на $\omega_i + \Delta\omega$ и рассчитывается по программе величина на выходе:

$$(12) D_i(t_i) = \sqrt{D_o^2 + k_\omega\omega_i(t_i)/k_v v_i(t_i) \cdot k_{T_p} T_{pi}(t_i)}$$

при неизменных входных параметрах v и T_p .

Из этого выражения рассчитывается коэффициент влияния величины входа модели $\omega_i(t)$ на выходной параметр модели $D_i(t)$.

$$(13) k_\omega = \frac{[D_i^2(t) - D_o^2] \cdot k_v \cdot v_i \cdot k_{T_p} \cdot T_{pi}}{(\omega_i + \Delta\omega)}$$

Аналогично определяются коэффициенты влияния других входных параметров. Следовательно, в таблице 5 появляется пункт 3 с предложенным методом эксперимента (таблица 6).

3 Экономическая часть

Предложенный метод оценки адекватности модели реальному объекту, проверенный на различных ТП, позволяет повысить точность измерений до (1,5–2) % и уменьшить время на эксперименты на (16–18) %, что повышает производительность на (6,8–7,4) %.

Таблица 6 – Вид экспериментов по оценке точности моделей

№ п. п	Вид эксперимента	Тип ТП	Используемый метод	Вид воздействия
3.	активно-пассивный с настройкой характеристик (2 варианта)	все виды	предложенный метод с подстройкой коэффициентов передачи	во время работы – подстройка, во время наладки – активный эксперимент

Заключение

Такой метод может применяться для непрерывных ТП других производств, он дает возможность не только произвести оценку точности моделей и датчиков косвенного измерения (ИВ) параметров ТП с объектами, использующих разные виды энергии, на трех этапах, но и производить подстройку их характеристик [8, 9].

Список литературы

- [1] Власов-Власюк, О.Б. Экспериментальные методы в автоматике. – М.: Машиностроение, 1969. – 412с.
- [2] Балакирев, В.С., Дудников Е.Г., Цирлин А.М. Экспериментальное определение динамических характеристик промышленных объектов. – М.: Энергия, 1967. – 232с.
- [3] Кижаяев С.А. Комплексная автоматизация бумагомассного агрегата. – Самара: Самарский университет, 2003. – 421с.
- [4] Кижаяев С.А. Математические модели и методы формализации при автоматизации технологических процессов для их интеграции в CALS-технологии // материалы 2 Всероссийской науч. - практ. конф. «Компьютерная интеграция производства и ИПИ-технологии» – Оренбург: ИПК ОГУ, 2005. – С. 85 – 95.
- [5] Кижаяев С.А. Аналитический метод синтеза современных систем автоматического управления в кабельной промышленности // материалы II Международной научной конференции «Аналитическая теория автоматического управления и ее приложения». – Саратов: СГТУ, 2005. – С. 100–102.
- [6] Кижаяев С.А. Метод синтеза сложных технических систем на основе принципов декомпозиции, децентрализации и законов сохранения материи и движения С.А. Кижаяев // Приборы и системы. Управление, контроль, диагностика. – Часть 2. – 2006. – № 6. – С. 7–12.
- [7] Кижаяев С.А. Интеллектуальные системы измерения в процессе экструзии в кабельной промышленности / С.А. Кижаяев // Датчики и системы. – 2004. – № 4. – С. 8–14.
- [8] Свидельство 2086143. Программа для вычисления диаметра провода (кабеля) / С.А. Кижаяев. – № 2008613145, заявл.09.08.2008, зарегистр. Бюл. «Программы для ЭВМ. Базы данных. Топологии интегральных микросхем», № 3, 2008.
- [9] Свидельство 2009615297. Программа для вычисления вязкости полимера (свинца) при экструзии / С.А. Кижаяев. – № 2009613132, заявл. 22.06.2009, зарегистр. Бюл. «Программы для ЭВМ. Базы данных. Топологии интегральных микросхем», № 3, 2009.

ПРАКТИЧЕСКИЕ ПРИМЕРЫ РЕАЛИЗАЦИИ МЕТОДОВ ДЕКОМПОЗИЦИИ И ДЕЦЕНТРАЛИЗАЦИИ В СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМАХ

С.А. Кижяев, А.К. Бульхин, В.Ф. Ключников

Закрытое акционерное общество «Самарская кабельная компания»

443022, Самара, ул. Кабельная, 9, Россия

kijaev@samaracable.ru; post-office@samaracable.ru

тел: +7 (846) 2282-455, факс: +7 (846) 333-27-70

***Ключевые слова:** централизованная система управления, алгоритмизация объектов, принципы декомпозиции, децентрализации и агрегирования, унификация схем, онтология и конфигурирование, автоматизированные системы управления*

Abstract

Schemes of realisation of decomposition and the decentralization, executed on the basis of the offered methods of the unified construction functional and block diagrammes for objects of difficult technological processes in industrial production are resulted.

Введение

Ранее технологический процесс (ТП) управлялся одной системой управления. При этом это были несложные процессы управления, как правило, с одномерными объектами [1]. Затем технологические процессы усложнялись, становились многомерными [2], но системы управления строились как централизованные, с использованием одной ЭВМ [3].

Это приводило к разработке сложнейших программ управления, труднейшей алгоритмизации технологического процесса, снижению его надежности и увеличению сроков внедрения. После этого стали ТП декомпозировать на ряд подпроцессов, что позволило упростить алгоритмизацию объектов, входящих в технологический процесс. Однако это не привело к упрощению управления, т.к. оно осталось централизованным со всеми ему присущими недостатками. Даже в настоящее время объекты управления в сложных ТП не декомпозируются в полном объеме [4], что приводит к трудным расчетам, не доступным на инженерном уровне.

Вопросу онтологии (формализации) и конфигурирования в построении функциональных и структурных схем в литературе не уделяется должного внимания. В производстве он многими проектировщиками и исследователями просто игнорируется, и техническая часть проекта начинается прямо с разработки принципиальных схем.

Многие инженеры на производстве разработку проектов ведут по-разному не унифицировано, а значит, их часто трудно переложить на программы САПР. Это приводит к значительным потерям времени на проектирование.

1 Постановка задачи

Поэтому необходимо было разработать метод построения функциональных и структурных схем систем управления ТП, который базируется на принципах декомпозиции ТП и децентрализации систем управления (СУ). Работы в этом направлении велись еще с 1976 – 1977 г.г. [5], но не закончены и в настоящее время [6-8]. Не произведена онтология (формализация) принципов декомпозиции и децентрализации для ТП, хотя это сделано для информационных и логистических процессов.

Как кажется авторам, это позволило бы производить комплексный синтез более простых СУ подобъектами, которые постепенно объединялись бы в более сложные, вплоть до полной автоматизации всего ТП, даже невероятной конфигурации. А сама СУ децентрализуется.

Такое построение автоматизированной системы управления (АСУ) ТП в кабельной и других отраслях промышленности с непрерывным производством позволило бы значительно сократить время на синтез системы, создало бы удобства в работе, повысило бы надежность технологического оборудования. Вот здесь и оказывается, что необходимо пользоваться положениями системного анализа из общей теории систем. Научные принципы концептуализации основ онтологии и конфигурирования в построении функциональных и структурных схем на базе декомпозиции, децентрализации и агрегирования и экономические аспекты приведены в [9]. В данной работе приведем практические примеры реализации в построении функциональных схем АСУ ТП.

2 Алгоритм решения

2.1 Декомпозиция объектов и технологических процессов

Декомпозиция – это научный метод исследования сложных ТП (объектов, целей), заменяемых исследованием отдельных элементов. Под декомпозицией будем понимать разделение (расщепление) сложной задачи (цели, объекта, ТП) на более простые компоненты.

На рисунке 1 изображена технологическая схема бумагомассного агрегата, предназначенного для наложения бумагопористой изоляции на провода телефонного кабеля.

На ней изображены: 1 – катушки с медными проводниками, 2 – смеситель бумажной массы, 3 – сетчатый барабан для наложения бумажной массы на медные проводники, 4 – прессовая часть для отжима влаги, 5 – гладильное устройство для придания бумажной изоляции концентрической формы, 6 – красильное устройство, 7 – печь сушки для формирования изоляции по влажности и диаметру изоляции, 7а – устройство вентиляции (вытяжки), 8 – тяговое устройство, 9 – крутильно-намоточной устройство. На рисунке 2 приведена функциональная схема декомпозиции ТП наложения бумагопористой изоляции на провода телефонного кабеля. Он расчленен на 7 подобъектов, которые являются самостоятельными подпроцессами.

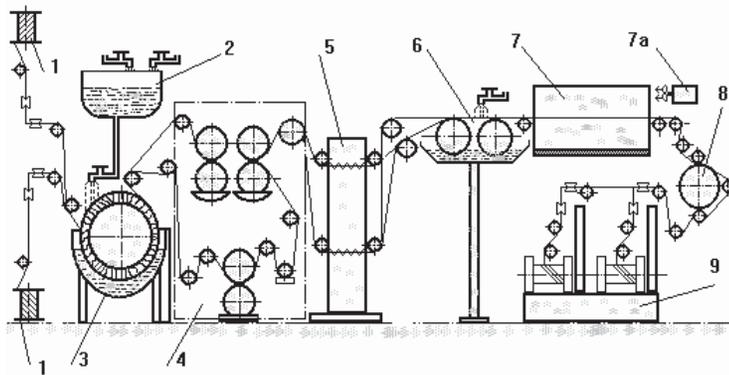


Рисунок 1 – Технологическая схема процесса наложения изоляции на бумагомассном агрегате

На нем обозначено: $Q_w(t), Q_b(t)$ – расходы воды и бумажной массы; $c(t), m_b(t)$ – концентрация и масса бумаги; $s(t), v(t)$ – степень помола массы и скорость ее наложения; $F(t), n(t)$ – усилие отжима влаги и число оборотов гладильного устройства; $P(t), P_v(t)$ – мощности печи сушки и вентиляции; $W_{нб}(p), W_{см}(p), W_{сб}(p), W_{np}(p), W_{зл}(p), W_n(p), W_{км}(p)$ – передаточные функции соответственно напорного бака воды, смесителя массы, сетчатого барабана

для наложения изоляции, пресса для отжима влаги, печи для формирования изоляции и конической мельницы для размола бумажной массы

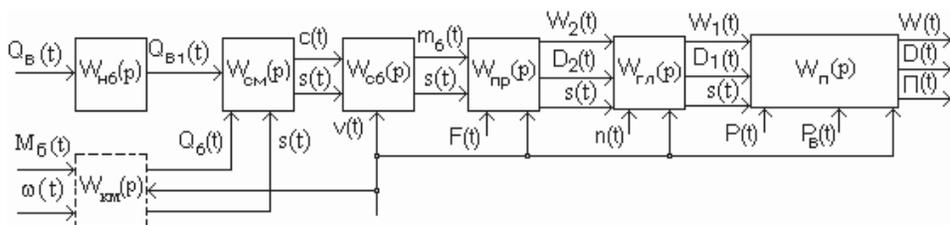


Рисунок 2 – Функциональная схема декомпозированного объекта управления (без возмущающих воздействий)

На рисунке 3 приведена схема процесса наложения алюминиевой защитной оболочки на гидравлическом пресса типа П 6043.

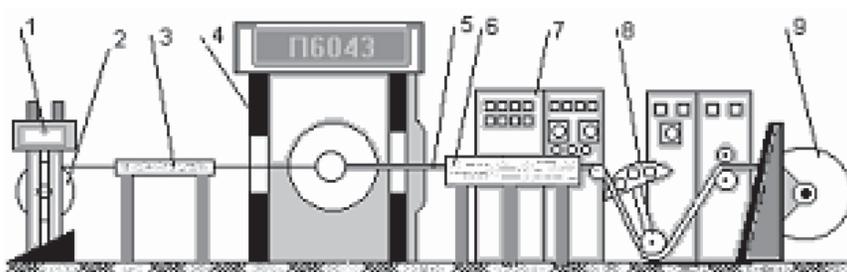


Рисунок 3 – Технологическая схема процесса

На нем обозначено: 1 – отдающее устройство, 2 – барабан с заготовкой кабеля, 3 – направляющий стол для исключения провисания кабеля, 4 – собственно гидравлический пресс, 5 – кабель в наложенной оболочке, 6 – ванна для охлаждения, 7 – система управления, 8 – компенсатор, 9 – приемный барабан с изготовленным кабелем

На рисунке 4 представлен вариант удачной искусственной декомпозиции объекта – пресса.

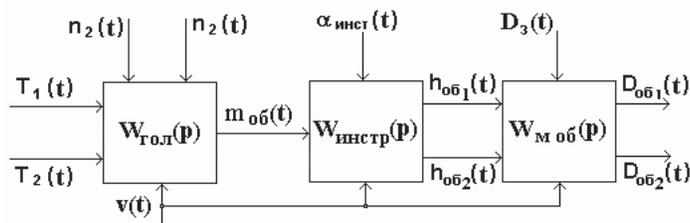


Рисунок 4 – Функциональная схема искусственной декомпозиции объекта – пресса для наложения алюминиевых оболочек методом выдавливания

На нем обозначено: $W_{гол}(p)$, $W_{инст}(p)$, $W_{м\ об}(p)$ – МПФ соответственно головки пресса, инструмента (дорн-матрица) и механизма образования диаметра оболочки; $m_{об}(t)$ – масса выдавленного алюминия оболочки; $h_{об1}(t)$, $h_{об2}(t)$ – толщины оболочки кабеля с правой и левой стороны; $D_3(t)$, $D_{об1}(t)$, $D_{об2}(t)$ – диаметры заготовки кабеля и диаметры оболочки кабеля с разных сторон; $\alpha_{инст}(t)$ – угол управления ключом инструмента перемещения дорна.

На рисунке 5 показана в качестве примера технологическая схема наложения бронепроводов на бронемашине.

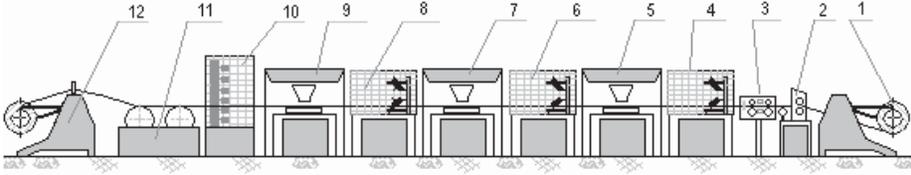


Рисунок 5 – Технологическая схема процесса наложения покрытов на бронемашине

На рисунке 5 приведено: 1– отдающее устройство, 2– правильное устройство, 3 – измеритель длины и диаметра кабеля, 4, 6 – бумагообмотчик, 5, 7, 9 – ванны с битумом, 8 – лентообмотчик, 10 – джутообмотчик, 11 – меловая ванна, 12 – приемное устройство

На рисунке 6 представлен вариант схемы декомпозиции этого процесса.

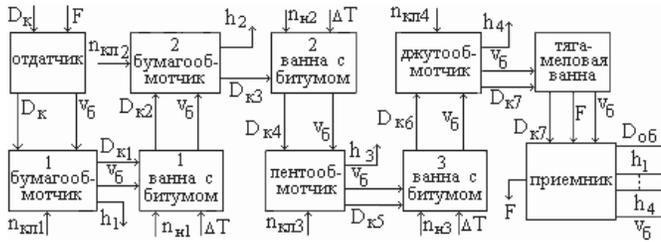


Рисунок. 6– Функциональная схема декомпозиции ТП наложения покрытов на бронемашине

На рисунке. 6 обозначено: D_k – диаметр заготовки кабеля; $D_{к1} - D_{к7}$ – диаметры кабеля на выходе соответствующих узлов после наложения лент, битума и джутовой нити; F – натяжение кабеля; $n_{кл1} - n_{кл4}$ – обороты клетей бумагообмотчиков, лентообмотчика и джутообмотчика; $n_{н1} - n_{н3}$ – обороты насосов подачи битума; v_b – скорость движения кабеля, ΔT – температура битума, $h_1 - h_4$ – шаги перекрытия лент и нити.

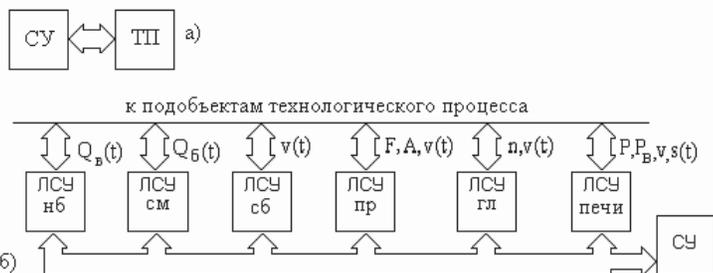
В результате исследований на основании формализации принципов декомпозиции разработана методика.

2.2 Децентрализация систем управления технологическими процессами

Децентрализация СУ может быть проведена по входным воздействиям или по выходным величинам. В первом случае упрощается структура составляющих подсистем, во втором случае – повышается их надежность. Примем децентрализацию СУ ТП наложения изоляции провода телефонного кабеля по промежуточному варианту: по выходным параметрам и по входному воздействию – скорости изолирования, как одному из возмущающих воздействий, которое необходимо отслеживать.

Для процесса наложения бумагомассной изоляции на провода телефонного кабеля (к рисунку 2) функциональная схема децентрализованной СУ будет выглядеть следующим образом на рисунке (7). На рисунках 8 – 11 приведены функциональные схемы подсистем управления диаметром изоляции проводов, управления влажностью изоляции, пористостью изоляции, скоростью изолирования. На них обозначено: R - D, R - D1, R -D2, R - m_б, R - c – регуляторы соответственно диаметра изоляции на выходе печи сушки; на ее входе, на входе гладильного узла, массы и ее концентрации; D-D, D-D1, D-D2, D-m_б, D - c – датчики измерения соответствующих технологических величин; ЛСУ – локальные СУ соответствующими технологически-

ми параметрами; R-W; R-W1; R-W2; R-F – регуляторы соответственно влажности изоляции на выходе печи сушки, перед печью, перед гладильным устройством и усилия отжима влаги; D-W, D-W1, D-W2, D-F – датчики измеряемых величин; R-П, R-N, R-s – регуляторы пористости, числа проводов и степени помола; D-N, D-П, D-s – датчики числа проводов, пористости изоляции и степени помола; D-ЧВ – датчик частоты вращения; D-A – датчик натяжения провода; R-A – регулятор натяжения.



а) централизованная система управления, б) система управления с децентрализацией на локальные системы управления отдельными подобъектами технологического процесса

Рисунок 7 – Функциональные схемы СУ ТП наложения изоляции на провода:

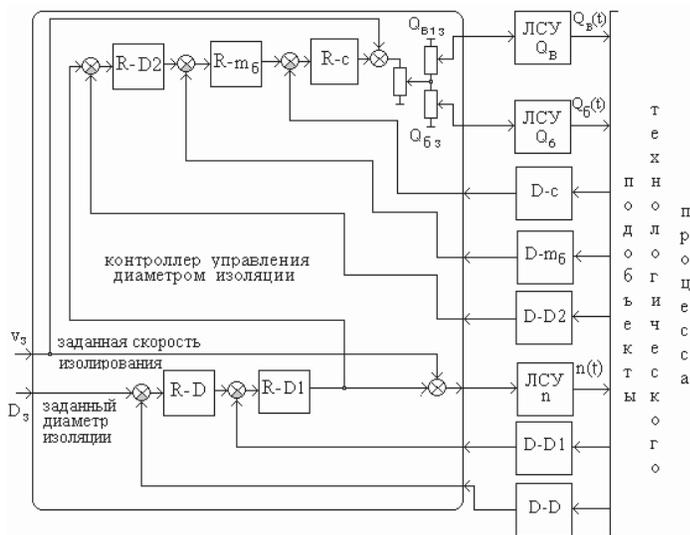


Рисунок 8 – Функциональная схема подсистемы управления диаметром изоляции проводов телефонного кабеля

Для ТП наложения алюминиевых оболочек на кабели можно произвести децентрализацию схемы управления процессом по рисунку 3 (рисунок 12).

При этом функциональные схемы ЛСУ подобъектами пресса П6043 будут иметь вид, представленный на рисунках 13 – 16. На рисунках 13 – 16 обозначено:

ПН – регулируемый преобразователь напряжения; ДН – датчик напряжения; R-T – регулятор температуры нагрева слитков, ЧП – частотный преобразователь напряжения; ДЧ – датчик частоты вращения электродвигателей; Дп – датчик числа оборотов насоса и тяги; Дv – датчик скорости движения кабеля; R-n – регулятор оборотов насосов; R-v – регулятор скорости; Д α – датчик угла поворота ключа механизма управления инструментом.

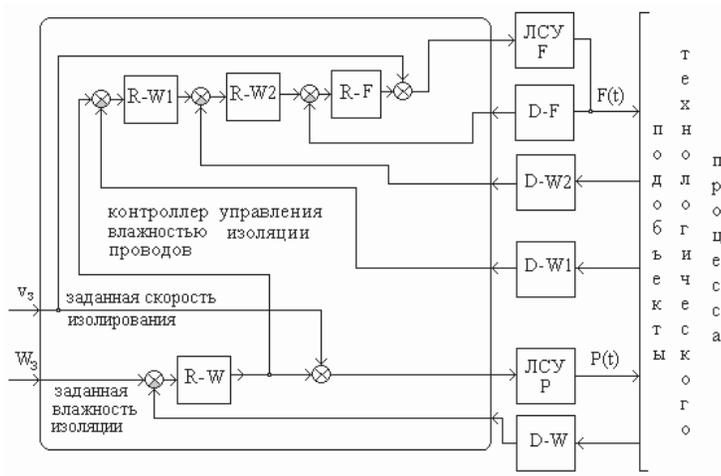


Рисунок 9 – Функциональная схема подсистемы управления влажностью изоляции

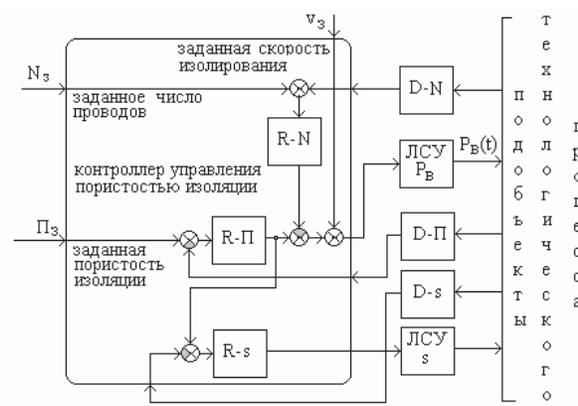


Рисунок 10 – Функциональная схема подсистемы управления пористостью изоляции

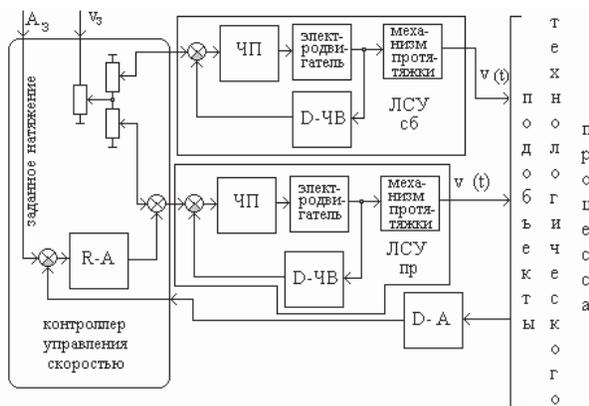
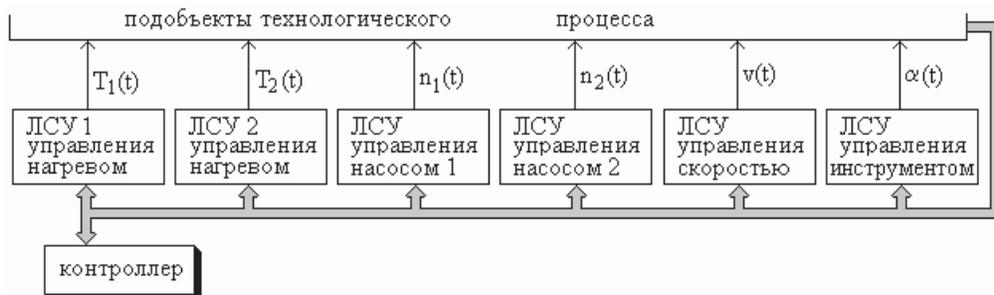


Рисунок 11 – Схема подсистемы управления скоростью изолирования проводов



ЛСУ–локальные системы управления подбъектами процесса

Рисунок 12 – Функциональная схема децентрализованной системы управления процессом наложения алюминиевых оболочек на кабели на гидравлическом прессе П6043

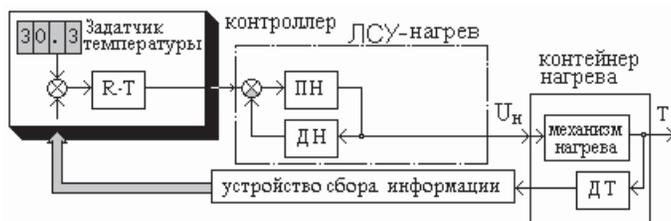


Рисунок 13 – Функциональная схема управления подбъектом процесса нагрева алюминиевых слитков на прессе П6043

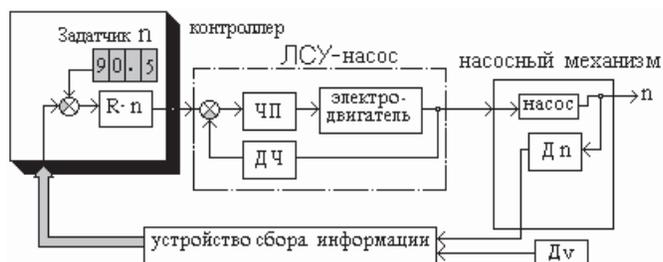


Рисунок 14 – Функциональная схема управления подсистемой насосных механизмов

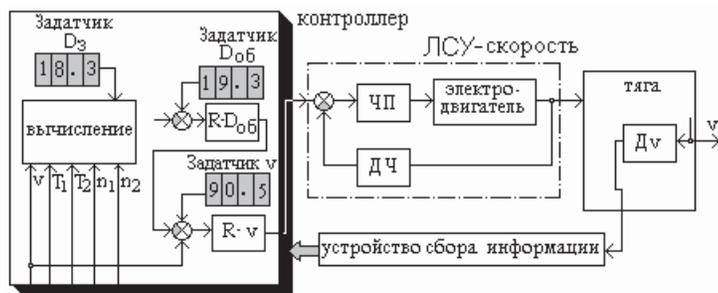


Рисунок 15 – Функциональная схема управления подсистемой управления скоростью

Разработана методика децентрализации.

2.3 Декомпозиция и децентрализация в управлении технологическими процессами

Теперь покажем как основы онтологии и конфигурирования в построении функциональных и структурных схем действуют в целом, т.е. при построении АСУ на основе принципов декомпозиции ТП на подобъекты с децентрализацией системы на подсистемы.

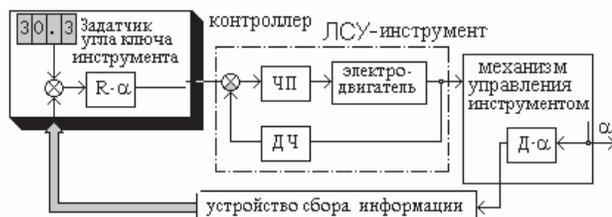


Рисунок 16 – Функциональная схема управления подсистемой механизма инструмента

ТП наложения покровов на кабели на бронемашинах

На рисунке 17 приведена функциональная схема декомпозиции ТП, управляемого децентрализованной АСУ ТП БМ.

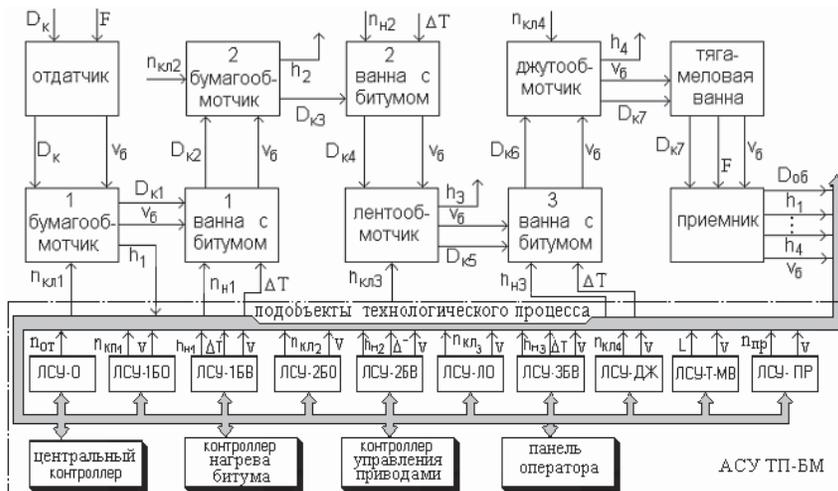


Рисунок 17 – Функциональная схема АСУ ТП-БМ

ТП наложения алюминиевых оболочек на гидравлических прессах

На рисунке 18 приведена функциональная схема декомпозиции ТП наложения алюминиевых оболочек на кабели, управляемого децентрализованной АСУ ТП-АО.

Разработана методика онтологии и конфигурирования при построении функциональных и структурных схем.

3 Экономическая часть

Предлагаемые основы в конфигурации, формализации и унификации в построении функциональных и структурных схем на принципах декомпозиции ТП, децентрализации СУ и агрегирования с учетом иерархического управления и поставленной цели позволили снизить уровень субъективизма, повысить уровень унификации и формализации проектов, уменьшить в среднем на (18–20) % стоимость проектов, на (23–29) % время наладки и обслуживания в зависимости от сложности объектов ТП и заложить основы для построения схем в пакетах САПР.

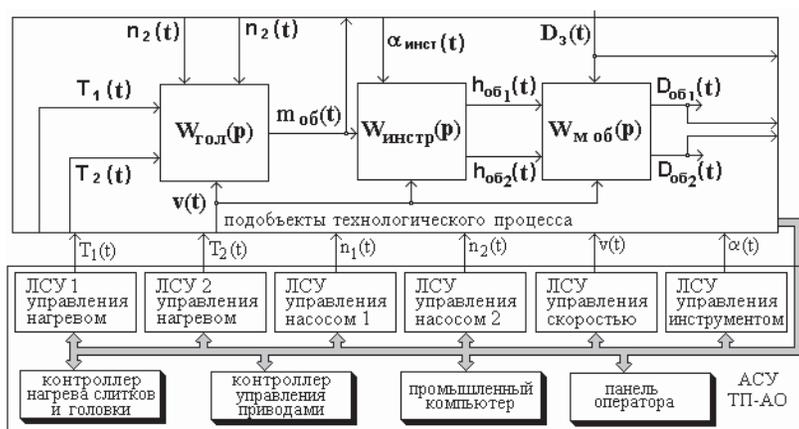


Рисунок 18 – Функциональная схема декомпозиции ТП , управляемого АСУ ТП-АО

Заключение

Применение принципов декомпозиции ТП и децентрализации СУ позволит производить их комплексную автоматизацию на инженерном уровне и создать банк для проектирования и расчетов АСУ ТП с применением программных комплексов.

Список литературы

- [1] Безносов, Б.Л. и др. Автоматизация и механизация производства кабельных изделий / Б.Л. Безносов и др. – М. – Л.: Энергия, 1967. – 328с.
- [2] Кижаяев С.А. Аналитические методы синтеза систем автоматического управления технологическими процессами / С.А. Кижаяев. – Самара : НТЦ, 2006. – 97с.
- [3] Рей У. Методы управления технологическими процессами /У. Рей. – М.: Мир, 1983. – 365с.
- [4] Авраменко, И.Е. Структурный синтез модели системы регулирования паровой турбоустановки / И.Е. Авраменко, Т.Ф. Жеребкина // Приборы и системы. Управление, контроль, диагностика. – 2005. – № 6. – С. 8–10.
- [5] Кухтенко А.И. О теории управления сложными объектами: труды 4-го Всесоюзного совещания по автоматическому управлению «Управление производством». – М.: Наука, 1978. – 369 с.
- [6] Гранкин Б.К., Козлов В.В., Лысенко И.В. Принципы декомпозиции сложных объектов в проектных исследованиях // Мехатроника, автоматизация, управление.– 2008.– № 6.– С.2-6.
- [7] Мирошник И.В. Теория автоматического управления. Линейные системы. СПб.: Питер, 2005. – 336 с.
- [8] Ефанов В.Н., Шевяков Е.Н., Лысенко И.В. Согласованно-децентрализованное управление сложными техническими объектами // Мехатроника, автоматизация, управление. – 2008. – № 6.– С.12-17.
- [9] Кижаяев, С.А. Принципы декомпозиции, децентрализации и агрегирования при управлении сложными технологическими процессами: сборник трудов XI Международной конференции «Проблемы управления и моделирования в сложных системах ПУМСС-2009» / Кижаяев С.А. – Самара: ИПУ СС РАН, 2009. – С.220–224.

Biography

1. Kizhaev Stanislav Alekseevich – Cand.Tech.Sci., The deserved inventor of the Russian Federation, The chief of electronics.
2. Bulhin Anwar Kashapovich – Cand.Tech.Sci., Professor SamGtUtion, The chairman of board of directors.
3. Kluchnikov Valeri Fedorovich – Cand.Econ.Sci., Senior lecturer it The Samara state space, university Koroleva, the general director.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ЧЕТЫРЕХФАЗНОГО РАСТРОВОГО СКП С КОЛЬЦЕВОЙ СЧИТЫВАЮЩЕЙ СИСТЕМОЙ И АКСИАЛЬНЫМ ИЗЛУЧАТЕЛЕМ

М.А. Габидулин

Московский государственный институт радиотехники, электроники и автоматики
(технический университет)
119454, Москва, пр. Вернадского, 78, Россия
gabidulin_ma@mail.ru
тел: +7 (495) 434-73-74

Ключевые слова: математическая модель, четырехфазный растровый синусно-косинусный преобразователь, кольцевая считывающая система, аксиальный излучатель

Abstract

The model of function of transformation, technique of optimum synthesis of interface of radial rasters and the analysis of influence of is constructive-technological errors on accuracy of the sine-cosine converter are given.

Введение

Создание цифровых преобразователей угловых перемещений, используемых в качестве информационно-измерительных элементов различных автоматических систем, с повышенными метрологическими и эксплуатационными характеристиками является одной из актуальных и технически сложных задач [1, 2]. Особенно актуально затронутая проблема стоит при разработке высокоточных преобразователей со встроенной конструкцией, что характерно для информационно-силовых модулей цифровых электроприводов, оптических, геодезических и других приборов [3].

Массовое их применение обуславливает широкий диапазон предъявляемых требований и, как следствие, многообразие принципов построения. Однако для прецизионных измерений наибольшее применение получили фотоэлектрические и индуктивные цифровые преобразователи перемещений, точность которых обеспечивается повышением электрической редукиции, прецизионности технологии изготовления, в первую очередь, растровых оптических и индуктивных шкал, интегральным съемом с них измерительной информации и применением интерполяции. Из них наибольшими универсальностью и возможностью совершенствования обладают фотоэлектрические цифровые преобразователи угла (ЦПУ), построенные с использованием растровых синусно-косинусных преобразователей (СКП) в качестве датчиков перемещений.

В них задача повышения статической и динамической точности электронных узлов при приемлемых аппаратных затратах решается относительно просто выбором коэффициентов электрической редукиции и интерполяции, унификацией и переводом их на микроэлектронную основу. Однако повышение точности в целом представляет собой сложную научно-техническую задачу из-за значительных инструментальных погрешностей растровых СКП [2,4].

Постановка задачи

Основным источником погрешностей растровых синусно-косинусных преобразователей являются погрешности изготовления и монтажа оптико-механических элементов, главным образом, растровых сопряжений, приводящих к искажениям функциональных зависимостей синусного и косинусного сигналов и выражающиеся в их несинусоидальности, переменных от

периода к периоду постоянных составляющих, неравенстве амплитуд, неортогональности пространственных фаз и появлении крупнопериодной фазовой погрешности.

Непосредственными причинами погрешностей являются отклонения закона модуляции потока лучистой энергии в функции угла от синусоидального, радиальные эксцентриситеты растров и неравномерность зазора между ними, погрешности геометрии растров (вариации положения и ширины штрихов).

Анализ показывает [2, 4], что существующие методики не позволяют обосновать требования к допускам на погрешности растровых сопряжений, гарантирующие заданную точность, и разработать модели пространственных сигналов, учитывающие погрешности от основных причин и дополнительные погрешности, вызываемые совместным их проявлением с другими причинами и определяющие реальную эффективность способов повышения точности.

Оценки указанных погрешностей требуют уточнений, так как получены на модели функции пропускания сопряжения линейных растров, справедливой в первом приближении только для секторных сопряжений, расположенных непосредственно перед излучателем. Не произведены сравнительные исследования усредняющих свойств секторных и бессекторных растровых сопряжений с учетом неравномерности яркости площадки излучателя, неоднородности чувствительности площадки приемников и оптической плотности штрихов растров.

Для решения затронутых вопросов поставлена задача разработки модели растрового СКП с кольцевой считывающей системой и аксиально расположенным источником и исследования его погрешностей.

Модель функции пропускания растрового сопряжения СКП

С целью выявления факторов, определяющих форму выходного сигнала растрового сопряжения и возможностей управления ею, разработана математическая модель функции пропускания сопряжения радиально-нецентральных растров для случая круглого аксиально расположенного источника излучения, имеющая вид

$$\tau_{pc} = \tau_1 \tau_2 + 2\tau_1 \tau_2 \sum_{p=1}^{\infty} \left(\prod_{i=1}^4 Sa(p\pi\tau_i) \right) Ja(p\pi\tau_5) \cos p(N_n \alpha + \alpha_0),$$

где τ_1 и τ_2 – относительная ширина штриха подвижного и неподвижного растров; τ_3 , τ_4 , τ_5 – параметры, учитывающие соответственно относительную разность периодов растров, взаимный наклон их штрихов, апертуру источника излучения; N_n – число штрихов подвижного растра (коэффициент редукции); α – величина преобразуемого угла; α_0 – постоянный сдвиг, определяемый конструктивными параметрами; p – номер высшей пространственной гармоники; $Sa(x) = \frac{\sin x}{x}$, $Ja(x) = \frac{2J_1(x)}{x}$, $J_1(x)$ – бесселева функция 1-го порядка 1-го рода.

Полученное выражение показывает, что применение растров с разным числом штрихов, скоса направления штрихов, а также использование источника излучения конечных размеров позволяет за счет возникновения эффектов нониуса и муара и влияния апертуры источника получить дополнительные коэффициенты подавления высших пространственных гармоник функции пропускания (соответственно множители $Sa(p\pi\tau_3)$, $Sa(p\pi\tau_4)$ и $Ja(p\pi\tau_5)$).

Фильтрующие свойства растрового сопряжения эквивалентны пятизвенному фильтру нижних пространственных частот, параметры звеньев которого определяются величинами τ_1, \dots, τ_5 . Варьируя ими и пользуясь тем, что функции $Sa(p\pi\tau_{1, \dots, 4})$ при $p\pi\tau_{1, \dots, 4} = k\pi$ ($k=1, 2, \dots$) обращаются в нуль, а также учитывая, что при равенстве значений $p\pi\tau_5$ ненулевым корням бесселевой функции $J_1(p\pi\tau_5)$ функция $Ja(p\pi\tau_5)$ также обращается в нуль, можно обеспечить эффективное многократное подавление гармоник (вплоть до $p = 10$ включительно), обеспечи-

вающее синусоидальную форму сигнала растрового сопряжения в функции перемещения. Так, при выборе $\tau_1 = \tau_2 = 0,5$, $\tau_3 = 0,333$, $\tau_4 = 0,2$ и $\tau_5 = 0,173$ достигается пренебрежимо малый относительный уровень искажений формы, не превышающий 0,04%.

Выбор $\tau_1 = \tau_2 = 0,5$ обеспечивает надежное двукратное подавление наиболее трудно фильтруемой второй гармоники и максимизацию амплитуды основной гармоники, а также полное подавление остальных четных гармоник и значительное подавление нечетных гармоник. Дополнительное подавление четных гармоник обеспечивает дифференциальная обработка сигналов в СКП. Выбор $\tau_3 = \frac{1}{3}$, $\tau_4 = \frac{1}{5}$, $\tau_5 = \frac{1}{7}$ обеспечивает подавление третьей, пятой и седьмой гармоник.

При этом учтено, что звено $Ja(p\pi\tau_5)$ целесообразно настраивать на подавление наивысшей из спектра значимых нечетных гармоник, так как параметр τ_5 пропорционален величине зазора между растрами и, следовательно, не может быть высокостабильным из-за влияния неточности установки и нестабильности ее величины. В силу этого звено настраивается на подавление гармоники наименьшей амплитуды, уже сильно ослабленной предыдущими звеньями, с тем, чтобы вариации зазора не приводили к заметным искажениям формы сигнала растрового сопряжения. Кроме того, это необходимо еще и для того, чтобы указанные вариации не приводили к заметному изменению амплитуды основной гармоники функции пропускания.

Показано, что вариация амплитуды при наихудшем сочетании параметров приведет к фазовой погрешности порядка 0,007 эл. рад, а при правильном конструировании СКП, обеспечивающем идентичность параметров синусно-косинусных сигналов, вообще не проявится. Совместное влияние вариации амплитуды и иных погрешностей элементов растрового СКП (например, угловой погрешности штрихов) дает составляющие фазовой погрешности, имеющие второй порядок малости, величина которых не превышает десятитысячных долей радиана.

Следовательно, анализ функции пропускания при наличии конструктивно-технологических погрешностей, учитывающий составляющие фазовой погрешности величиной порядка 10^{-3} эл. рад и выше, может быть проведен без учета влияния вариации зазора на амплитуду основной составляющей функции пропускания. Данный вывод позволяет считать произведение передаточных функций звеньев фильтра по первой гармонике постоянным, не зависящим от вариации зазора и при дальнейшем анализе рассматривать простейшее обтюрационное сопряжение растров, для которого $Sa(p\pi\tau_3) = Sa(p\pi\tau_4) = Ja(p\pi\tau_5) = 1$, а учет фильтрующих свойств этих звеньев может быть произведен простым умножением амплитуд гармонических составляющих на произведение:

$$K_{\phi} = Sa(p\pi\tau_3)Sa(p\pi\tau_4)Ja(p\pi\tau_5).$$

При неблагоприятном сочетании параметров (когда величина τ_5 не мала и нельзя пренебречь вариацией амплитуды) вариации зазора, приводящие к амплитудной модуляции гармоник, могут быть учтены введением эквивалентной модуляции прозрачности растров.

Таким образом, рассмотрен синтез сопряжения радиальных растров с аксиально расположенным источником с точки зрения обеспечения наилучшей синусоидальности и стабильности амплитуды сигнала растрового сопряжения как пятизвенного фильтра нижних пространственных частот.

Для сопряжения радиально-центрального подвижного и радиально-нецентрального неподвижного растров [5,6], что обеспечивает устранение влияния на величину τ_4 вариации зазора второго порядка малости, а также смещения муаровой картины при эксцентриситете подвижного растра, получены соотношения, позволяющие выбрать параметры растрового сопряжения, обеспечивающие наиболее эффективное подавление высших пространственных гармоник

$$N_n = N_n - \frac{\pi\tau_3}{\beta}, r_n = \tau_4 \frac{R_0^2 w_n}{2\Delta R}, r_n = \frac{x_{J1} R_0 (L-h)}{k_5 N_n h},$$

где $w_n = \frac{2\pi}{N_n}$, N_n – угловой шаг и число штрихов неподвижного растра на угле 2π ; ΔR и β – радиальная и угловая полуширина диафрагмы в виде сектора узкого кольца со средним радиусом R_0 ; r_n – радиус начальной окружности смещения радиально-нецентрального неподвижного растра; R_0 – средний радиус рабочей зоны растров; L – расстояние между плоскостями источника излучения и неподвижного растра; h – зазор между растрами; x_{J1} – наименьший ненулевой корень $J_1(x)$; k_5 – номер главной гармоники, на подавление которой настраивается пятое звено фильтра. Однако если параметры настроены неточно или используются не все звенья фильтра, то величина искажений может возрасти на 1-2 порядка и привести к значительным погрешностям СКП и ЦПУ в целом.

Величина N_n для секторного растрового сопряжения может выбираться любой, а для бессекторного должна быть только целым числом. Поэтому нониусный эффект при полном интегральном усреднении по всей окружности растров не может быть использован и, следовательно, $\tau_3 = 0$.

Для анализа влияния конструктивно-технологических погрешностей достаточно рассмотреть обтюрационное растровое сопряжение с последующим учетом дополнительного коэффициента подавления фильтра нижних пространственных частот.

Анализ влияния на функцию пропускания конструктивно-технологических погрешностей

Для случая обтюрационного растрового сопряжения произведен анализ влияния конструктивно-технологических погрешностей СКП (положения и ширины штрихов растров, монтажа элементов СКП, модуляции прозрачности растров, совместного действия наиболее значимых факторов) на функцию пропускания [3,7]. Погрешности задавались аналитическими разложениями в виде рядов Фурье, а формулы получены с точностью до вторых порядков малости включительно.

Произведенный для каждого из перечисленных влияющих факторов анализ позволил синтезировать выражение для обобщенной функции пропускания растрового сопряжения в виде “идеальной” составляющей, составляющих первого порядка малости, имеющих пространственную частоту гармоник влияющих факторов, а также комбинационных составляющих второго порядка малости.

Установлено, что по отношению ко всем влияющим факторам считающаяся диафрагма в первом приближении является пространственно-частотным фильтром, модуль передаточной функции которого определяется выражением

$$K(l) = Sa(l\beta),$$

где l – номер пространственной гармоники влияющего фактора, β – угловая полуширина диафрагмы.

Получено выражение для коэффициента интегрального усреднения K_n растрового сопряжения, показывающего, во сколько раз снижается среднеквадратическое отклонение фазовой погрешности функции пропускания в угловых единицах по сравнению со среднеквадратическим отклонением углового положения штрихов растров:

$$K_n = \sqrt{\frac{\xi N_n}{1-\xi}},$$

где ξ – относительная ширина диафрагмы.

Анализ функции пропускания показал, что бессекторное растровое сопряжение в первом приближении осуществляет подавление всех составляющих погрешности, однако вторые порядки малости остаются неподдавленными.

Погрешности монтажа элементов растрового сопряжения приводят к возникновению первой гармоники угловой погрешности функции пропускания на обороте, вызываемой, во-первых, совместным влиянием эксцентриситета e_3 установки источника излучения и перекоса θ_1 подвижного растра, и во-вторых, эксцентриситетами подвижного растра и источника излучения. Причем эксцентриситет источника излучения совместно с неперпендикулярностью подвижного растра оси вращения приводит к появлению погрешности, не устраняемой бессекторным растровым сопряжением, которая имеет вид первой гармоники с амплитудой

$$\Delta\alpha_m = \frac{e_3\theta_1}{2L}.$$

Особенностью рассматриваемой кольцевой считывающей системы с аксиально расположенным источником излучения является пониженная чувствительность к вариации зазора между растрами по сравнению с другими структурами считывающих систем, также использующими расходящийся поток излучения, но с не аксиально расположенными источниками.

Анализ модуляции ширины штрихов выявил, что для достижения минимума ее влияния на амплитуду функции пропускания целесообразно выбирать $\tau_1 = \tau_2 = 0,5$.

Учет совместного влияния модуляции прозрачности и положения штрихов растров показал, что действие данных факторов не позволяет полностью подавить угловую погрешность функции пропускания, вызванную модуляцией положения штрихов, которая в случае бессекторного сопряжения может быть оценена формулой

$$\Delta\alpha \leq E_{m1}G_{m2} + E_{m2}G_{m1},$$

где G_{m1} , G_{m2} – среднеквадратическое значение погрешности углового положения штрихов подвижного, неподвижного растров соответственно; E_{m1} , E_{m2} – среднеквадратическое значение функции относительного отклонения прозрачности подвижного, неподвижного растров соответственно от ее среднего значения. Использование эквивалентной функции модуляции прозрачности позволяет учесть также неравномерность освещенности, чувствительности фотоприемников, зазора между растрами и ширины штрихов.

Формирование выходных синусно-косинусных сигналов СКП

Выражение для выходного сигнала фотоприемника, воспринимающего поток, прошедший через растровое сопряжение r -ой фазы,

$$I_r = \Phi_0 S_{\text{пн}} \tau_{\text{рс}r},$$

где Φ_0 – поток излучения, облучающий растровое сопряжение; $S_{\text{пн}}$ – интегральная чувствительность фотоприемника, повторяет с точностью до коэффициента пропорциональности выражение для функции пропускания $\tau_{\text{рс}r}$. Такой сигнал не может быть непосредственно использован для аналого-цифрового преобразования, поскольку содержит постоянную составляющую и, следовательно, не соответствует формату СКП. Наличие постоянной составляющей делает точность последующей обработки зависимой от абсолютной величины сигнала, а, стало быть, и от мощности излучения источника, что снижает точность преобразования.

Простейшим способом ее компенсации является простое вычитание из сигнала фотоприемника постоянного сигнала, подобранного по уровню равным среднему значению постоянной составляющей. Однако это не устраняет источника погрешности, поскольку настройка становится зависимой от мощности излучения источника, а постоянная составляющая в общем случае колеблется при изменении преобразуемого угла.

Другим способом является вычитание постоянного сигнала, полученного с помощью дополнительного фотоприемника, размещенного рядом с источником излучения. Способ устраняет влияние изменения мощности источника излучения. Однако флуктуации постоянной составляющей по-прежнему влияют на погрешность преобразования.

Наибольшее распространение получил способ, основанный на формировании дифференциального сигнала вычитанием двух сигналов, образованных с помощью двух растровых сопряжений максимальной степени идентичности, причем основные составляющие сигналов сдвинуты друг относительно друга по фазе на угол π . Этим обеспечивается устранение не только постоянной составляющей, в том числе переменной по обороту, но и дополнительно устранение или снижение остальных четных высших пространственных гармоник.

Таким образом, с точки зрения достижения наивысшей точности для синтеза синусного и косинусного сигналов СКП наиболее приемлема четырехфазная система растровых сопряжений, обеспечивающая формирование сигналов, имеющих взаимный сдвиг по основной гармо-

нике, равный $\frac{\pi}{2}(r-1)$, где $r = \overline{1,4}$ – номер растрового сопряжения.

С учетом обобщенной модели функции пропускания растрового сопряжения разработана математическая модель четырехфазного растрового СКП с аксиально расположенным источником излучения, позволяющая учесть влияние конструктивно-технологических погрешностей и оценить возможности их дополнительного подавления в выходных сигналах СКП:

$$\bar{I}_s = \frac{(1 + \delta_4)I_4 - (1 + \delta_2)I_2}{2\Phi_0 S_{\text{пн}}} = \frac{(1 + \delta_4)\tau_{\text{pc}4} - (1 + \delta_2)\tau_{\text{pc}2}}{2},$$

$$\bar{I}_c = \frac{(1 + \delta_1)I_1 - (1 + \delta_3)I_3}{2\Phi_0 S_{\text{пн}}} = \frac{(1 + \delta_1)\tau_{\text{pc}1} - (1 + \delta_3)\tau_{\text{pc}3}}{2},$$

где $\bar{I}_{s,c}$ – нормированные разности фототоков приемников, обеспечивающие формирование синусного и косинусного сигналов; $\delta_{1,\dots,4}$ – относительные отклонения коэффициентов передачи дифференциальных схем по соответствующим входам, учитывающие также отклонения амплитуд составляющих $\tau_{\text{pc}r}$ под действием иных факторов.

Список литературы

- [1] Домрачев В.Г. и др. Схемотехника цифровых преобразователей перемещений: Справоч. пособ./ В.Г. Домрачев, В.Р. Матвеевский, Ю.С. Смирнов. – М.: Энергоатомиздат, 1987.
- [2] Высокоточные преобразователи угловых перемещений / Э.Н. Асиновский, А.А.Ахметжанов, М.А. Габидулин и др.; Под общ. ред. А.А.Ахметжанова. – М.: Энергоатомиздат, 1986.
- [3] Габидулин М.А. Синтез фотоэлектрического цифрового преобразователя угла и скорости, встраиваемого в исполнительный электродвигатель // Мехатроника, автоматизация, управление. 2008, №5. – С. 23-27.
- [4] Габидулин М.А. Методы повышения точности накапливающих фотоэлектрических ЦПУ // Оптические сканирующие устройства и измерительные приборы на их основе: Сб. тез. докл. 5 Всесоюзн. совещ. – Барнаул: АлтПИ, 1990, ч. II, – С. 188-190.
- [5] Фотоэлектрические преобразователи информации / Л.Н. Преснухин, В.Ф.Шаньгин, С.А.Майоров, И.В.Меськин. Под ред. Л.Н. Преснухина. – М.: Машиностроение, 1974.
- [6] Габидулин М.А. Исследование функций пропускания растровых сопряжений растровых датчиков линейных и угловых перемещений методом рядов Фурье // Техника преобразования информации. Сб. трудов. – М.: МИРЭА, 1975, вып. 83. – С. 24-43.
- [7] Габидулин М.А., Лейбович И.Д. Исследование погрешностей четырехфазных растровых СКП // Оптические сканирующие устройства и измерительные приборы на их основе: Сб. тез. докл. 5 Всесоюзн. совещ. – Барнаул: АлтПИ, 1990, ч. II, – С. 184-185.

МЕХАНИЗМ НЕЧЕТКОГО ЛОГИЧЕСКОГО ВЫВОДА В УПРАВЛЕНИИ ЖЕЛЕЗНОДОРОЖНЫМ ПЕРЕЕЗДОМ

В.А. Буцких, Н.Н. Васин

Поволжский государственный университет телекоммуникаций и информатики
443010, Самара, Л.Толстого, 23, Россия
brainvictor@rambler.ru, vasin@psati.ru

Ключевые слова: лингвистическая переменная, участок приближения, состояние переезда, функция принадлежности

Abstract

The aim of study is to describe railway crossing control system which takes into account the speed of the train and condition of railways, using the apparatus of fuzzy logic.

Введение

Система управления переездом обеспечивает безопасность его функционирования. При вступлении поезда на участок приближения система управления включает светофоры, звуковую сигнализацию и закрывает автоматический шлагбаум. Длина участка приближения зависит от возможной скорости движения поездов и составляет, как правило 1500 ÷ 2000 м. При высокой скорости движения поезда 60 ÷ 120 км/ч время ожидания (от момента вступления состава на участок приближения до прибытия на переезд) составляет от 45 секунд до 2 минут. Однако при низкой скорости движения 10 ÷ 15 км/ч время ожидания у закрытого шлагбаума превышает 10 ÷ 15 минут, что вызывает нервозность водителей транспортных средств и, как следствие, ведет к нарушениям требований безопасности: проезд на запрещенный сигнал светофора, объезд закрытого шлагбаума. Результатом являются аварии с тяжелыми последствиями.

Для решения данной проблемы предлагается создать систему управления переездной сигнализацией адаптивной к скорости, ускорению поезда и состоянию железнодорожного переезда. Внедрение подобной системы позволяет существенно сократить время ожидания транспортных средств у закрытого шлагбаума.

На практике информация о состоянии, необходимая для принятия решения о закрытии переезда является неточной, неполной, некачественной, поскольку определяется параметрами окружающей среды и качеством подъездных путей. Подобная информация необходима, чтобы транспортное средство, находящееся на переезде, успело освободить его до прибытия состава на железнодорожный переезд. Удобным инструментом построения системы с нечетко заданными параметрами, представления экспертных знаний, выработки управляющих решений и оценки полученных результатов является лингвистическая переменная и ее значения (нечеткие переменные и нечеткие множества).

Логико-лингвистическое описание систем позволяет описывать процесс в терминах, близких и понятных специалистам в той области деятельности, для которой разрабатывается система. Так переменная Состояние, определяющая качество дорожного покрытия на переезде, будет являться лингвистической. Ее значения будут задаваться в виде термов: хорошее, выше среднего, среднее, ниже среднего, плохое. Поскольку переменная Состояние оказывает существенное влияние на момент принятия решения о закрытии переезда и является лингвистической, то и остальные переменные такие как Ускорение и Скорость удобно представлять в нечетком виде. Так переменная скорость будет являться лингвистической, если ее значения

будут не числовыми (1, 2, 3, ..., 100), а заданными в виде терм, например: низкая, ниже средней, средняя, выше средней, высокая.

Структурная схема системы

Общая схема системы представлена на рисунке 1. Логически она состоит из трех основных модулей:

- Модуль оценки скорости и ускорения поезда – собирает данные о скорости, ускорении состава на всем участке приближения к переезду.
- Система принятия решений – обрабатывает данные полученные от датчиков скорости и ускорения поезда, а также принимает информацию от дежурного по станции о состоянии переезда. На основе этой информации принимается решение о блокировании переезда.
- Исполнительные механизмы – блокируют переезд при достижении поездом определенной координаты.



Рисунок 1 – Общая схема системы управления переездом

Модуль оценки скорости и ускорения поезда – собирает данные о скорости, ускорении состава на всем участке приближения к переезду. Он состоит из датчиков скорости состава расположенных на расстоянии 200 ÷ 500 метров вдоль всего участка приближения. Датчик скорости поезда представляет собой кластер состоящих из двух магнитных педалей расположенных на расстоянии $l = 1$ друг от друга. При этом скорость движения $V = l/T$, где T - время прохождения колеса вагона над кластером. Ускорение можно определить как изменение скорости ΔV за некий промежуток времени Δt . Обмен данными между датчиками и системой принятия решения происходит по шине RS-485 с использованием протокола Modbus.

Система принятия решений представляет собой микроконтроллер с процессором архитектуры ARM и интерфейсами обеспечивающими взаимодействие с RS-485, исполнительными механизмами, и системой связи с дежурным по станции.

Использование механизма нечеткого логического вывода для численной оценки координаты закрытия переезда

Важным преимуществом систем, построенных на основе теории нечетких множеств, является большая гибкость и адекватность реальному миру, а также более быстрое, по сравнению с традиционными моделями, получение окончательного результата через специфическое построение и простоту используемых нечетких операций. Для успешного применения математических методов при принятии решения о закрытии переезда необходимо средство, которое учитывает суждения специалистов-экспертов.

Несмотря на известную аналогию с методами теории вероятности существенное отличие метода теории нечетких множеств заключается в том, что неопределенность связана не со случайностью, а с имеющимися неточностями и размытостями [2].

Следует подчеркнуть, что количество терм лингвистических переменных должно быть адекватным задаче и не превышать 5-7 [1]. Значение базовой переменной Состояние определяется дежурным по станции на основании его знаний о состоянии переезда и задается непосредственно в виде числа в диапазоне 0÷10. При этом наиболее плохому состоянию соответствует число 0, а хорошему число 10. На рисунке 2 изображена функция принадлежности лингвистической переменной Состояние.

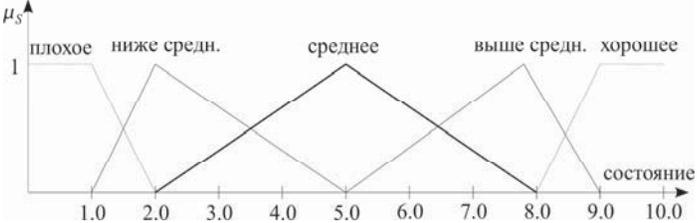


Рисунок 2 – Функция принадлежности нечеткого множества Состояние

Функции принадлежности для переменных Скорость и Ускорение представлены на рисунке 3.

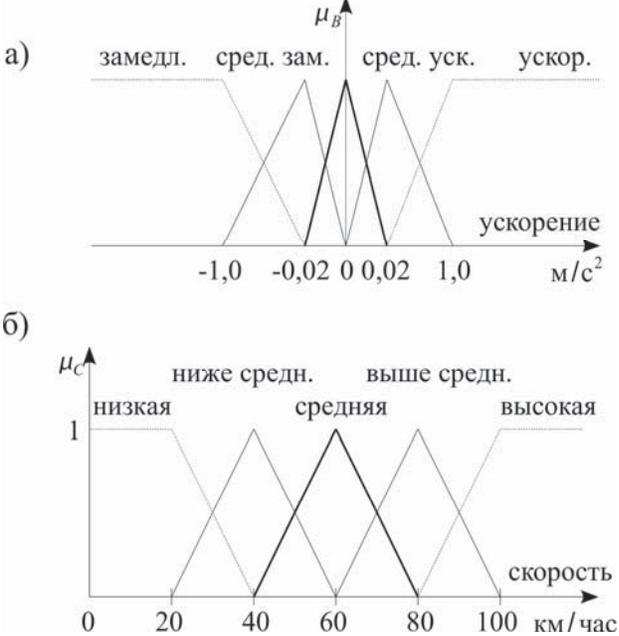


Рисунок 3 – Функции принадлежности нечетких множеств: Скорости и Ускорения

В процессе моделирования Скорость V , Ускорение a , Состояние S являются входными лингвистическими переменными системы принятия решений. На выходе системы будет Координата Z , показывающая координату участка приближения, при достижении которой составом должно производиться закрытие шлагбаума. Функция принадлежности лингвистической переменной Координата изображена на рисунке 4.

Анализ влияния скорости, ускорения и состояния переезда производится с помощью композиционного правила вывода, когда результат получается на основе нечеткого вывода на правилах. Совокупность этих правил, хранимых в базе знаний системы принятия решений, определяет взаимосвязь входных и выходной лингвистических переменных, т.е. образует не-

четкое отношение $R((X, Y, S), Z)$ [3]. Функция принадлежности выходной переменной Координата имеет следующий вид:

$$(1) \mu_z(z) = \bigvee_{x \in X} \bigvee_{y \in Y} \bigvee_{s \in S} (\mu_A(x) \wedge \mu_B(y) \wedge \mu_S(s) \wedge \mu_R(x, y, s, z)).$$

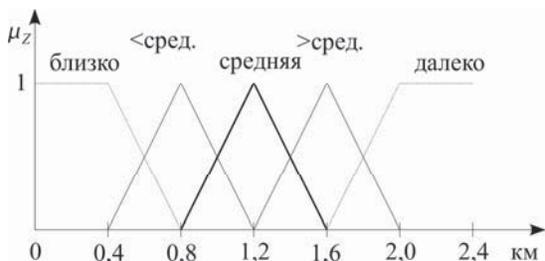


Рисунок 4 – Функция принадлежности лингвистической переменной Координата

Как можно заметить база знаний, содержащая правила выработанные экспертами имеет трехмерную структуру. Осью x в ней является термы лингвистической переменной Скорость, осью y Ускорение, осью z Состояние. Базу данных можно представить в виде совокупности таблиц, которые представляют слои трехмерной структуры. Имя слоя будет задаваться термом переменной Состояние. В качестве примера приведен слой, соответствующий среднему состоянию переезда (таблица 1).

Таблица 1

Ускорение, a	Скорость, V				
	низкая	ниже средней	средняя	выше средней	высокая
замедление	близко	близко	<сред.	средняя	>сред.
сред. зам.	близко	<сред.	<сред.	средняя	далеко
нулевое	близко	<сред.	<сред.	>сред.	далеко
сред. уск.	<сред.	средняя	средняя	далеко	далеко
ускорение	средняя	>сред.	>сред.	далеко	далеко

Пример

Используя функции принадлежности лингвистических переменных (рисунки 2-4), правила функционирования заданные экспертами, можно определить выходное значение переменной Координата. Процесс решения изображен на рисунке 5.



Рисунок 5 – Схема нечеткого логического вывода

Вывод для трех лингвистических переменных с последующим процессом дефазификации представляет собой весьма сложную вычислительную задачу. Для ее упрощения применяется математический пакет моделирования MATLAB. Применение MATLAB позволяет построить трехмерные графики зависимости значения базовой переменной координаты от изменения входных переменных. Графики зависимости изображены на рисунках 6-8

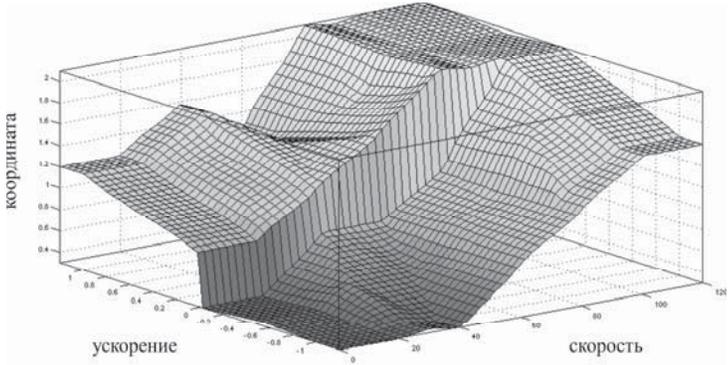


Рисунок 6 – Зависимость координаты от скорости и ускорения

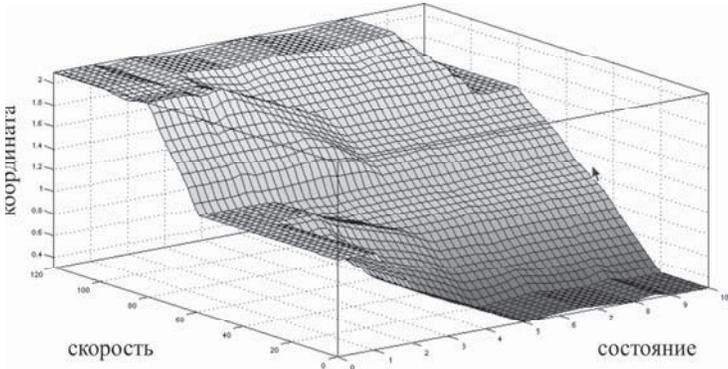


Рисунок 7 – Зависимость координаты от скорости и состояния

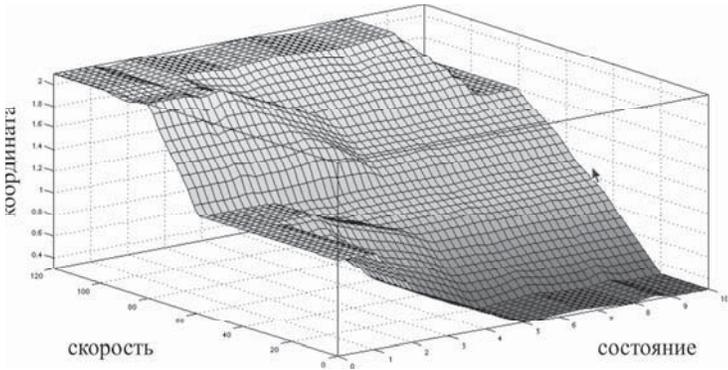


Рисунок 8 – Зависимость расстояния от ускорения и состояния

Примеры расчетов для некоторых численных значений Скорости, Ускорения, Состояния и Координаты представлены в таблице 2.

Таблица 2

Входные переменные			Результат
Скорость V , км/ч	Ускорение a , м/с ²	Состояние S	Координата Z , км
10	0,5	4	1,0
40	0,03	5	1,21
60	-0,13	3	0,8
90	0,17	2	2,05
120	-0,01	7	1,85

Список литературы

- [1] Cornelius T. Leondes Fuzzy Theory Systems: Techniques and Applications // Academic Press, New York, 1999 p. 3-25.
- [2] Гузаиров М.Б., Машкина И.В. Интеллектуальная поддержка принятия решений по управлению защитой информации в критически важных сегментах информационных систем // Приложение к журналу «Информационные технологии» №7/2008.
- [3] Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений // Мир 1976 с. 113-126.



Reform of Regional Governance
Реформирование государственного
управления в регионе

СРАВНИТЕЛЬНЫЙ АНАЛИЗ РАЗВИТИЯ ФЕДЕРАЛЬНЫХ ОКРУГОВ РОССИИ. ЧИСЛЕННОСТЬ НАСЕЛЕНИЯ, РОЖДАЕМОСТЬ И СМЕРТНОСТЬ

В.Д. Фурасов, Д.В. Фурасов

Государственный университет управления
109542, Москва, Рязанский проспект, 99, Россия
furasov@mail.ru
тел: +7 (495) 371-70-88

Ключевые слова: динамические модели, интервальные эволюционные индексы, таблицы о ранах, демографические показатели развития

Abstract

A set of problems dealing with the construction of non-stationary dynamic models under limited availability of statistical information is considered. With the help of synthesized models some evolutionary indices of dynamic processes are determined. Principal ideas are illustrated by examples of dynamical analysis of growth of federal districts.

Введение

По своей направленности данная работа примыкает к [1-3]. В значительной мере она мотивирована указом Президента РФ от 19.01.2010 N 82, в соответствии с которым из Южного федерального выделен Северо-Кавказский федеральный округ. В его состав вошли Республика Дагестан, Республика Ингушетия, Кабардино-Балкарская Республика, Карачаево-Черкесская Республика, Республика Северная Осетия - Алания, Чеченская Республика и Ставропольский край. В Южном федеральном округе остались Республика Адыгея, Республика Калмыкия, Краснодарский край, Астраханская, Волгоградская и Ростовская области.

Изменение числа федеральных округов потребовало внесения корректив в проводимые исследования. В данной работе с единых позиций рассматриваются вопросы построения нестационарных динамических моделей и сравнительного анализа численности, рождаемости и смертности населения восьми федеральных округов Российской Федерации.

1 Общие соотношения

Пусть k - дискретное время, $[k_0, k_1]$ - интервал наблюдения, $k_0 < k_1$, и при любом рассматриваемом k исследуемый процесс социально-экономического развития $y(k)$, $k \in [k_0, k_1]$, удовлетворяет условию $y(k) > 0$. В соответствии с [2,3] каждому такому процессу при любом положительном целом θ , выбираемом из условия $\theta \leq (k_1 - k_0)$, вместе с моделью

$$x(k+1) = A(k)x[k], \quad k \in [k_0 + \theta - 1, k_1 - 1], \text{ где}$$

$$A(k) = y(k+1)y^T[k] / \|y[k]\|^2, \quad x[k] = [x(k), \dots, x(k - \theta + 1)]^T,$$

может быть поставлен в соответствие индекс развития

$$ir^\theta(k) = \sum_{\alpha=1}^{\theta} a_\alpha(k).$$

Введение индекса развития позволило формализовать понятия спада ($ir^\theta(k) < 1$) и подъема ($ir^\theta(k) > 1$) в момент времени k . Индекс развития $ir^\theta(k_1 - 1)$, отвечающий максимальному значению $\theta = k_1 - k_0$, называется эволюционным и обозначается символом $ei(k_0, k_1)$. Знание $ei(k_0, k_1)$ позволяет говорить об эволюционном спаде (эволюционном подъеме) на интервале наблюдения $[k_0, k_1]$, если $ei(k_0, k_1) < 1$ ($ei(k_0, k_1) > 1$).

Предположим теперь, что равенство $N(\theta + 1) = k_1 - k_0 + 1$ имеет место при выбранном θ и некотором целом $N \geq 1$. В этом случае, положив $t_h = k_0 + (h-1)(\theta + 1)$, моделируемому процессу $y(k)$, $k \in [k_0, k_1]$, можно поставить в соответствие последовательность интервальных эволюционных индексов

$$ei(t_h, t_h + \theta) = \sum_{\alpha=1}^{\theta} a_{\alpha}(t_h + \theta - 1), \quad h = 1, \dots, N.$$

Введение такой последовательности позволяет говорить об интервальном эволюционном спаде – спаде на промежутке времени $[t_h, t_h + \theta]$, если $ei(t_h, t_h + \theta) < 1$, и – о подъеме, если $ei(t_h, t_h + \theta) > 1$. При этом $t_1 = k_0$ и $t_N = k_1 - \theta$ при любом $N \geq 1$.

Приведем также следующее определение. Пусть $N \geq 2$. Социально-экономический процесс $y(k)$, $k \in [k_0, k_1]$, называется θ -растущим (θ -убывающим или θ -снижающимся) на интервале $k \in [k_0, k_1]$, если $ei(k_0, k_0 + \theta) < ei(k_1 - \theta, k_1)$ ($ei(k_0, k_0 + \theta) > ei(k_1 - \theta, k_1)$). Если $N \geq 3$ и при любом $h = 1, \dots, N - 1$ выполняется неравенство $ei(t_h, t_h + \theta) < ei(t_{h+1}, t_{h+1} + \theta)$ ($ei(t_h, t_h + \theta) > ei(t_{h+1}, t_{h+1} + \theta)$), то процесс $y(k)$, $k \in [k_0, k_1]$, называется монотонно θ -растущим (монотонно θ -убывающим или монотонно θ -снижающимся) на интервале наблюдения $k \in [k_0, k_1]$.

Важная роль в компьютерной визуализации проводимых исследований принадлежит интервальным эволюционным индексам

$$ei(t_h, k + 1) = \sum_{\alpha=1}^{k+1-t_h} a_{\alpha}(k),$$

порождаемым моделями с расширяющимися пространствами состояний

$$x(k + 1) = A(k)x[k], \quad k \in [t_h, t_h + \theta - 1], \text{ где}$$

$$A(k) = y(k + 1)y^T[k] / \|y[k]\|^2, \quad x[k] = [x(k_0), \dots, x(k)]^T.$$

Вместе с индексами $ei(t_h, k + 1)$ в рассмотрение естественным образом вводятся интервальные таблицы о рангах – таблицы эволюционных индексов, в которых каждый процесс (показатель) развития и, следовательно, регион, занимает место в соответствии со своим интервальным эволюционным индексом $ei(t_h, t_{h+1})$.

В основе сравнительного анализа численности, рождаемости и смертности населения федеральных округов лежит статистика 1996 – 2007 гг., содержащаяся в [1]. В принятых обозначениях $k_0 = 1996$, $k_1 = 2007$ и в роли θ , таким образом, может выступать любое из чисел 2, 3, 5. Условию $N(\theta + 1) = k_1 - k_0 + 1$ при этом можно удовлетворить, положив $\theta = 3$, $N = 3$. Далее при рассмотрении показателей развития используются следующие обозначения: РФ – Российская Федерация, ЦФО – Центральный федеральный округ, СЗФО – Северо-Западный федеральный округ, ЮФО – Южный федеральный округ, СКФО – Северо-Кавказский феде-

ральный округ, ПФО – Приволжский федеральный округ, УФО – Уральский федеральный округ, СФО – Сибирский федеральный округ, ДФО – Дальневосточный федеральный округ.

Графики эволюционных индексов Российской Федерации на рисунках, посвященных сравнительному анализу федеральных округов, выделяются кружками. При сравнительном анализе субъектов федерации, образующих тот или иной федеральный округ, кружками обозначается эволюционный индекс этого округа.

2 Численность населения

Эволюционные индексы численности населения и таблицы о рангах восьми федеральных округов показаны на рисунке 1.

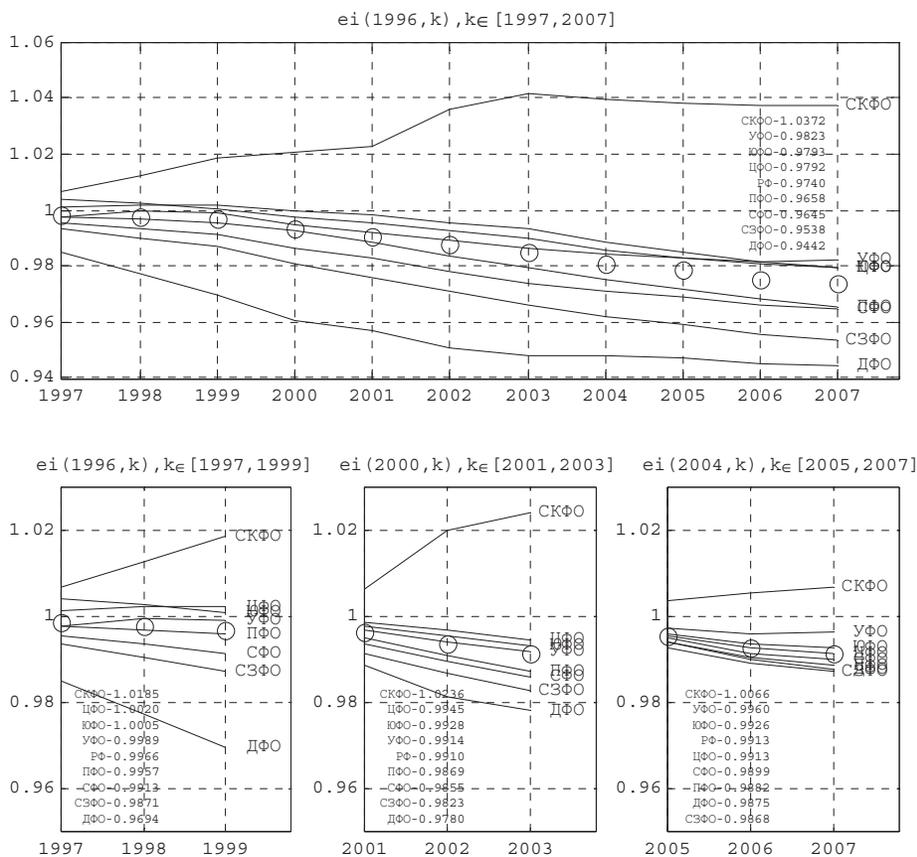


Рисунок 1 – Эволюционные индексы численности населения федеральных округов Российской Федерации

Сравнение графиков и таблиц, приведенных на рисунке 1, с соответствующими графиками и таблицами о рангах семи федеральных округов [3] показывает, что Северо-Кавказский округ в восьмерке занял место Южного федерального округа в составе 7. Возглавляя на рис. 1, все таблицы о рангах, он удовлетворяет как всем условиям интервального эволюционного подъема, так и условию $ei(1996, 2007) > 1$. Численность населения Северо-Кавказского округа θ -растет в 1996-2003 гг. и θ -убывает в 1996-2007 и 2000-2007 гг.

Южный федеральный округ на рис.1 условию эволюционного подъема вместе с Северо-Кавказским округом удовлетворяет только в 1996-1999 гг. Численность населения Южного федерального округа, монотонно θ - убывая на рассматриваемом интервале наблюдения, θ -убывает в 1996-2003 гг. и (вместе с численностью Северо-Кавказского округа) в 2000-2007 гг.

Заметим, что в 1996-2003 гг. вместе с численностью населения Северо-Кавказского федерального округа θ -растет численность населения Дальневосточного федерального округа, вместе с численностью населения Южного федерального округа θ -убывает численность населения Центрального, Северо-Западного, Приволжского, Уральского и Сибирского федеральных округов. Вместе с численностью населения Северо-Кавказского и Южного федеральных округов в 2000-2007 гг. θ -убывает численность населения Центрального федерального округа, в 1996-2007 гг. – численность населения Центрального, Северо-Западного, Приволжского, Уральского и Сибирского федеральных округов. Вместе с численностью населения Южного федерального округа в 1996-2007 гг. монотонно θ -убывает численность населения Центрального федерального округа.

Численность населения Российской Федерации на рисунке 1 θ -убывает в 1996-2003, θ -растет в 2000-2007 и θ -убывает в 1996-2007 гг.

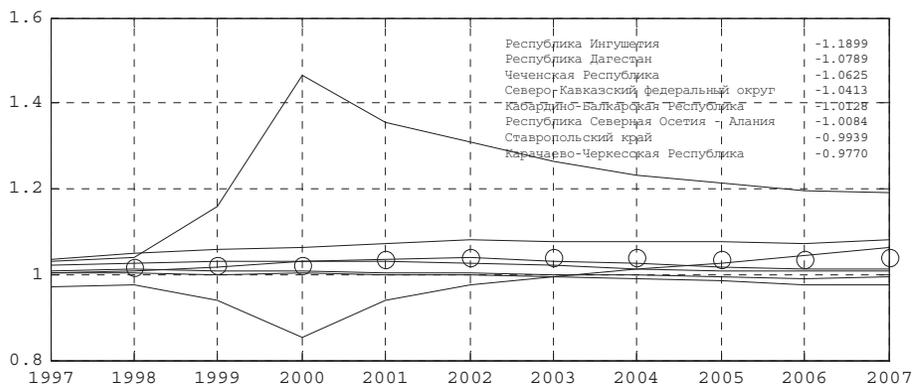


Рисунок 2 – Численность населения. Северо-Кавказский федеральный округ

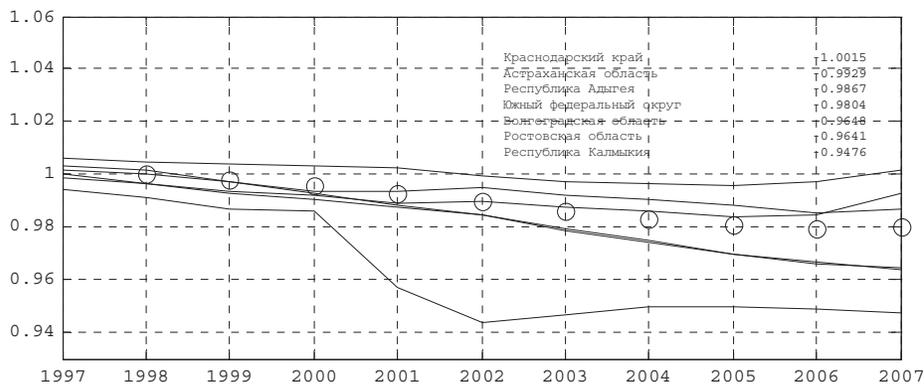


Рисунок 3 – Численность населения. Южный федеральный округ

Для лучшего понимания, если не особенностей динамики развития, то, во всяком случае, имеющихся статистических данных, эволюционные индексы численности населения Северо-Западного и Южного федеральных округов показаны на рисунках 2, 3.

3 Рождаемость населения

Все эволюционные индексы рождаемости населения (и в том числе – эволюционные индексы Северо-Кавказского и Южного федеральных округов), показанные на рисунке 4, удовлетворяют условиям интервального спада $ei(1996,1999) < 1$ и подъема $ei(2000,2003) > 1$, $ei(2004,2007) > 1$. Удовлетворяют они и условию подъема $ei(1996,2007) > 1$.

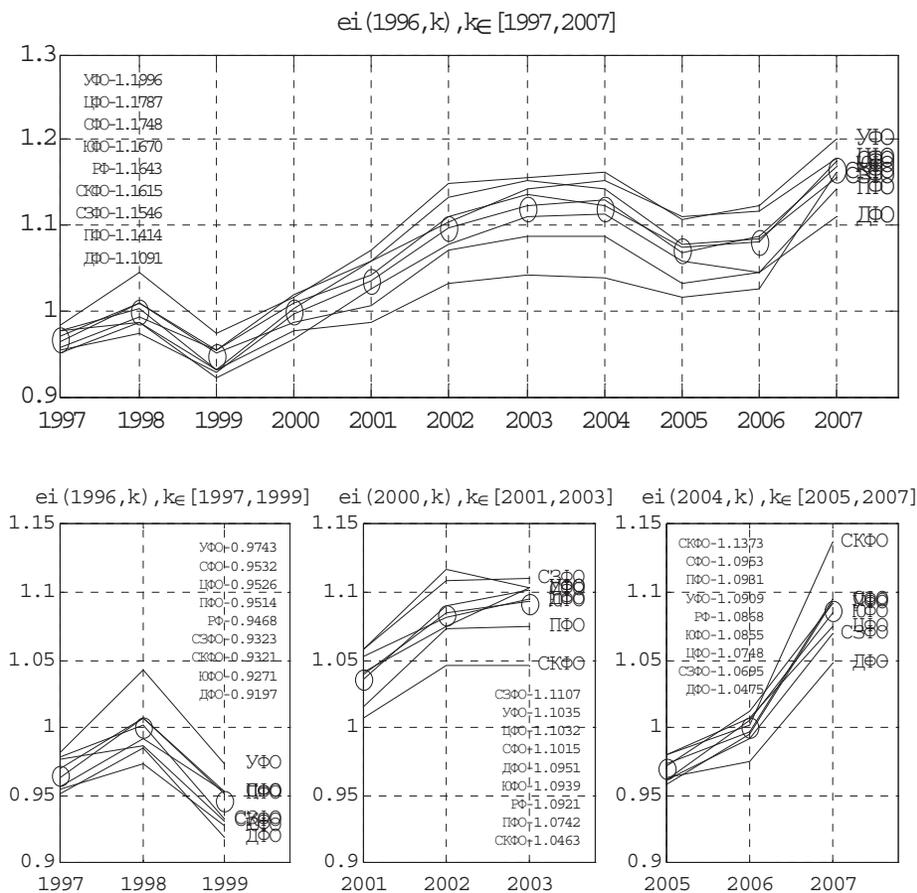


Рисунок 4 – Эволюционные индексы рождаемости населения федеральных округов Российской Федерации

Из неравенств интервального спада $ei(1996,1999) < 1$ и подъема $ei(2000,2003) > 1$, справедливых для всех федеральных округов следует θ -рост рождаемости населения всех федеральных округов в 1996-2003 гг. Точно так же θ -рост рождаемости населения всех федеральных округов в 1996-2007 гг. следует из неравенств интервального спада $ei(1996,1999) < 1$ и подъема $ei(2004,2007) > 1$. При этом в 2000-2007 гг. вместе с рождаемостью населения Северо-

ро-Кавказского федерального округа θ -растет рождаемость населения Приволжского федерального округа, а вместе с рождаемостью населения Южного федерального округа θ -снижается рождаемость населения Центрального, Северо-Западного, Уральского, Сибирского и Дальневосточного федеральных округов.

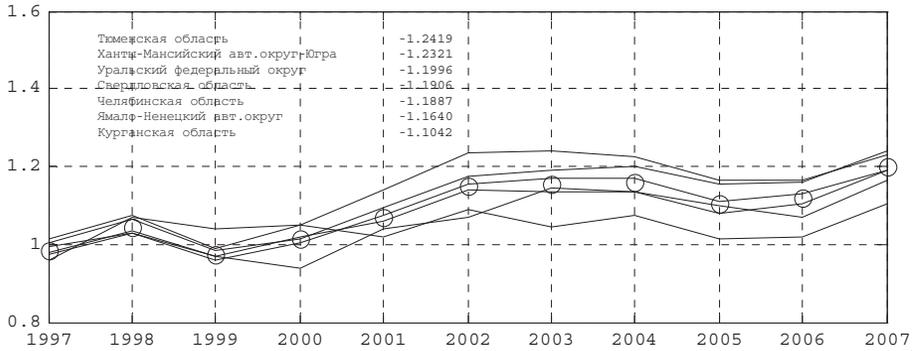


Рисунок 5 – Рождаемость населения. Уральский федеральный округ

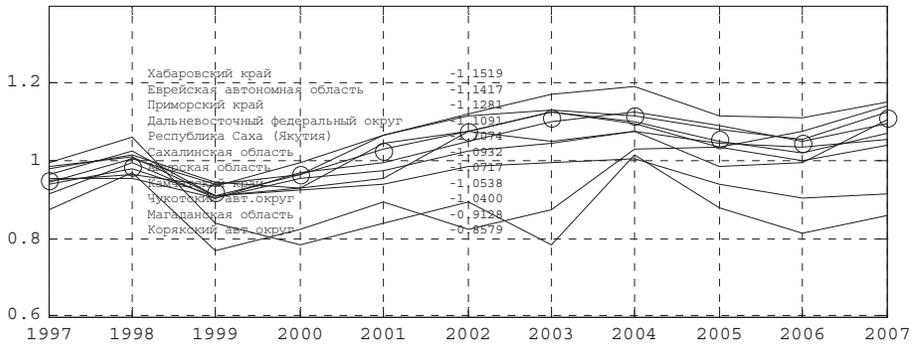


Рисунок 6 – Рождаемость населения. Дальневосточный федеральный округ

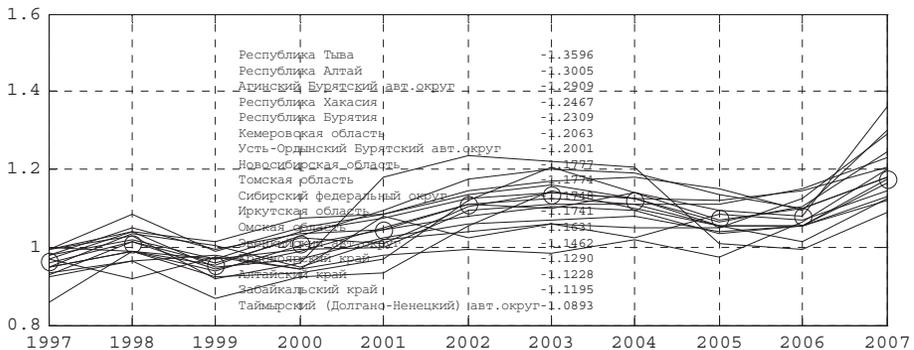


Рисунок 7 – Рождаемость населения. Сибирский федеральный округ

Особого внимания на рисунке 4 заслуживают локальные максимум и минимум, достигаемые эволюционными индексами $ei(1996, k)$, $k \in [1997, 2007]$, в 1998 г. и в 2005, 2006 гг. Дополнительный свет на выявленные закономерности динамики статистических данных рождаемости населения проливают рисунки 5-7, посвященные эволюционным индексам рождаемости населения субъектов федерации, образующих Уральский, Дальневосточный и Сибирский федеральные округа России.

4 Смертность населения

В отличие от эволюционных индексов рождаемости все эволюционные индексы смертности населения на рисунке 8 удовлетворяют условиям интервального подъема $ei(1996, 1999) > 1$, $ei(2000, 2003) > 1$ и спада $ei(2004, 2007) < 1$. На смену неравенству $ei(1996, 2007) > 1$ пришло условие эволюционного спада $ei(1996, 2007) < 1$.

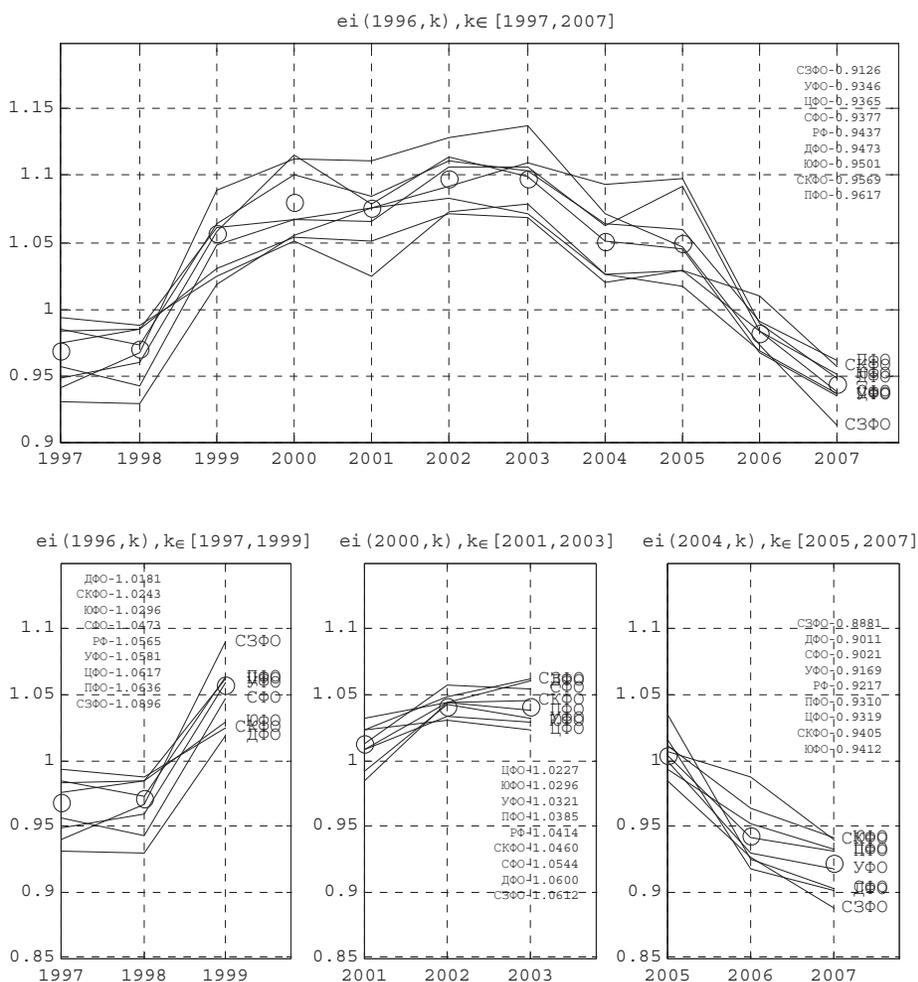


Рисунок 8 – Эволюционные индексы смертности населения федеральных округов Российской Федерации

Из неравенств $ei(1996,1999) > 1$, $ei(2000,2003) > 1$ и $ei(2004,2007) < 1$ следует θ -снижение смертности населения всех федеральных округов как в 2000-2007 гг., так и в 1996-2007 гг. В 1996-2003 гг. вместе со смертностью населения Северо-Кавказского федерального округа θ -растет смертность населения Сибирского и Дальневосточного федеральных округов, а вместе со смертностью населения Южного федерального округа θ -убывает смертность населения Центрального, Северо-Западного, Приволжского и Уральского федеральных округов.

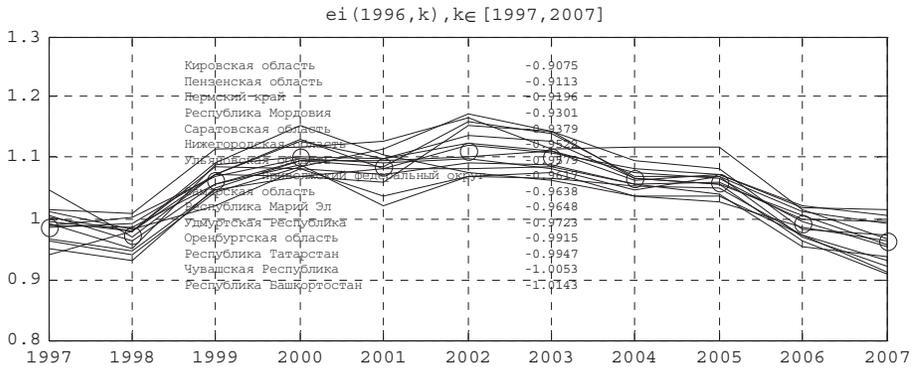


Рисунок 9 – Смертность населения. Приволжский федеральный округ

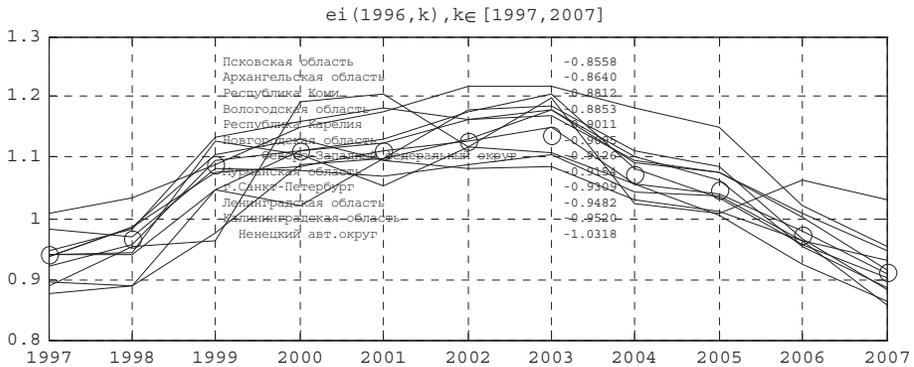


Рисунок 10 – Смертность населения. Северо-Западный федеральный округ

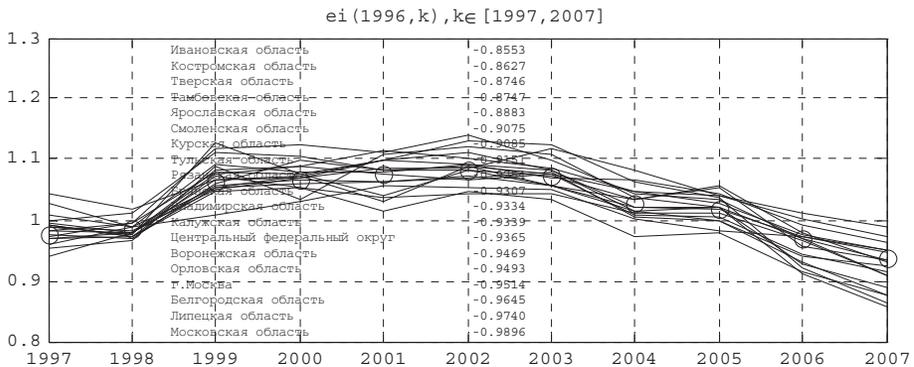


Рисунок 11 – Смертность населения. Центральный федеральный округ

Особого внимания заслуживают субъекты федерации, эволюционные индексы смертности населения которых удовлетворяют условию эволюционного подъема $ei(1996,2007) > 1$. Вместе с Чувашской Республикой (1.0053), Республикой Башкортостан (1.0143) и Ненецким автономным округом (1.0318), «засветившимися» на рис. 9- 11, к ним относятся: Ямало-Ненецкий автономный округ (1.0119), Ханты-Мансийский автономный округ - Югра (1.0606), Омская область (1.0222), Эвенкийский автономный округ (1.0362).

Заключение

В работе рассмотрены вопросы сравнительного анализа процессов развития в условиях ограниченной реально доступной статистической информации. С использованием нестационарных динамических моделей с расширяющимися пространствами состояний определены интервальные эволюционные индексы. Сформулированы условия интервального подъема и спада. С единых позиций рассмотрены особенности динамики демографических показателей развития – численности, рождаемости и смертности населения существующих с 19.01.2010 восьми федеральных округов Российской Федерации.

Благодарности

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, проект № 07—06—00224.

Список литературы

- [1] Фурасов В.Д., Фурасов Д.В. Сравнительный анализ развития регионов России. Численность населения и валовой региональный продукт// Проблемы управления и моделирования в сложных системах: Труды X Международной конференции (Самара, 23 – 25 июня 2008). Самара: Самарский научный центр РАН. 2008. С. 323-330.
- [2] Фурасов В.Д., Фурасов Д.В. Сравнительный анализ развития федеральных округов России. Модели с расширяющимися пространствами состояний, цепные и базисные эволюционные индексы. Проблемы управления и моделирования в сложных системах: Труды X Международной конференции (Самара, 22 – 24 июня 2009). Самара: Самарский научный центр РАН. 2009. С. 397 – 404.
- [3] Фурасов В.Д., Фурасов Д.В. Динамический анализ социально-экономических процессов. Интервальные эволюционные индексы. Третья Международная конференция «Системный анализ и информационные технологии» САИТ-2009 (14-18 сентября 2009 г., Звенигород, Россия): Труды конференции. М., 2009. С. 361 – 370.
- [4] Федеральная служба государственной статистики (www.gks.ru/scripts/db_inet/dbinet.cgi 2009).

ФЕНОМЕНОЛОГИЧЕСКОЕ ВРЕМЯ В ПРОЦЕССАХ ПРИНЯТИЯ РЕШЕНИЙ

В.А. Виттих

Институт проблем управления сложными системами РАН
443020 г. Самара, ул. Садовая, 61
vittikh@iccs.ru
тел: +7 (846) - 332-39-27, факс: +7 (846) - 333-27-70

Ключевые слова: феноменологическое время, принятие решений, сознание, предметный мир человека, онтологическая модель ситуации, датируемость

Abstract

In the article is validated the necessity of phenomenological time usage in course of decision-making, associated with human time perception and emotional experience, that precedes the application of "objective" physical time.

Введение

Мы часто говорим, что «опоздали с принятием решения», что «не хватило времени для принятия обоснованного решения», что «решение принимать преждевременно» и т.п., не задумываясь о том, имеем ли мы в виду «объективное» время, измеряемое с помощью часовых механизмов, или иное – «субъективное». Ведь в зависимости от условий, в которых мы оказались, и нашего психологического состояния, течение времени может осознаваться нами по-разному. Если мы ожидаем окончания хирургической операции, время тянется медленно, а концерт, на котором исполняются любимые музыкальные произведения, пролетает незаметно. И наша готовность выделить своё время тоже может быть различной: на одно у нас «всегда есть время», а на другое - «никогда нет времени». В процессах принятия решений мы можем ориентироваться именно на это «внутреннее» время, прежде чем использовать те или иные приборы, измеряющие время. Иными словами, представления о времени могут быть разными.

И. Ньютон, с его идеей «абсолютного времени», опираясь на определение Аристотеля времени как меры движения, утверждал, что его ход равномерен и синхронен по всем точкам пространства и ни от чего не зависит. Поскольку уравнения классической механики были нечувствительными к направлению времени, в них ничего не изменялось, если время начинало течь в обратную сторону, что противоречило действительности. Открытие законов термодинамики объяснило существование «стрелы времени», указывающей, что время течёт только в одном направлении. После подтверждения модели расширяющейся Вселенной стали говорить о космологической «стреле времени», а созданная А. Эйнштейном теория относительности освободила физику от «абсолютного времени» [1].

Все эти взгляды на понятие «время» формировались в рамках парадигмы классического естествознания, которая, как известно, наряду с разделением субъекта и объекта, выводит за пределы субъекта познания индивидуальные качества человека, характеризующие его как личность и «мешающие» постижению объективной истины. И если такой подход применим в естествознании, то он оказывается не приемлемым в гуманитарных науках и при исследованиях сложных искусственных систем, создаваемых и функционирующих при участии человека [2]. Здесь *познание* необходимо рассматривать как *естественный образ действий личности в целом*, с её чувствами, желаниями, опытом, надеждами и т.п., но не как чисто духовную функ-

цию, взятую изолированно. В этой связи говорят также о психологической «стреле времени», связанной с осознанием времени человеком и содержанием его переживаний [1].

В работе [3] предлагается использовать другой термин – не психологическое, а *феноменологическое время*. Дело в том, что основной принцип феноменологии, как одного из направлений философии XX века, сформулированный М. Хайдеггером в его «Бытии и времени» [4] звучит так: «к самим вещам». А это означает, что необходимо обратиться к первичному, изначальному опыту сознания, в котором вещи предстают не как предметы имеющихся теорий, точек зрения, установок, не как нечто, на что мы смотрим глазами других, а как нечто, что само раскрывается нам в первичном опыте [5]. Феноменологический опыт, таким образом, должен быть огражден от любых привносимых извне допущений [1].

В фундаментальной онтологии М. Хайдеггера, являющейся по своей сути феноменологической онтологией, рассматривается именно феноменологическое время, хотя сам автор это термин не использует, поскольку такое время предшествует времени физическому, «объективному». В данной статье, опираясь на труды М. Хайдеггера, обосновывается необходимость использования понятия феноменологического времени в процессах принятия решений.

1 Время и бытие

Сформулировав вопрос о возможности поставить рядом понятия «время» и «бытие», М. Хайдеггер пишет [6], что «бытие от раннего начала западноевропейской мысли до сего дня значит то же, что присутствие. Из присутствия, присутствования звучит настоящее. Последнее, согласно расхожему представлению, образует с прошлым и будущим характеристику времени. Бытие как присутствие определяется временем... Мы именуем время, когда говорим: у всякой вещи своё время. Этим подразумевается: всё, что когда-либо есть, всякое сущее приходит и уходит в должное ей время и пребывает некоторое время на протяжении отмеренного ей времени». Таким образом, временность является фундаментальной, базисной характеристикой бытия.

Бытие, по Хайдеггеру, – это бытие человека, которое он определяет термином *Dasein*, в русском переводе означающим «присутствие человека здесь и сейчас» и понимаемое как подвижное (а не стабильное) состояние. В своём исходном значении *Dasein* – это перевод на немецкий язык латинского термина «существование», «*existentia*» [7]. *Existentia* используется для обозначения существования вещи, а не существования вообще. М. Хайдеггер под экзистенцией понимает именно существование человека, а свою «фундаментальную онтологию» в [4] он разрабатывает как феноменологию человеческого присутствия [1].

Человек, обладающий памятью, в своём бытии, в своём настоящем, «содержит» и своё прошлое, и своё будущее. Он сейчас переживает собственное прошлое и сейчас воображает своё собственное будущее, строя предположения и планы. В этом смысле человек – *времяобразующее сущее*. Если бытие – это прежде всего человеческое бытие, то окружающий мир – это предметный мир человека. Пространство представляет собой одно из измерений этого предметного мира, но *воспоминания и предвосхищения (ожидания) открывают нам ещё и временное измерение нашего предметного мира и, соответственно, временную структуру нашего собственного сознания* [3]. Это и есть феноменологическое время, отличное от понятия времени как «всеобщей формы бытия материи». Всё, относящееся к феноменологическому времени, таким образом, есть наше, человеческое истолкование.

Выражение «на это у меня нет времени» означает, что будущее образует непрерывную цепь из «связанных по времени» действий, переживаний и т.д. Оно означает также, что для меня является невозможным в этом течении времени «передвинуть» один из этих членов назад или вперёд таким образом, чтобы осуществлялась дальнейшая деятельность и т.д. Выражение «ещё рано» означает, что время ещё не наступило (возможности еще нет), «уже поздно» – что возможности больше нет. Только «теперь» есть «время и возможность»; «теперь» находится

между «ещё рано» и «уже поздно» и должно «восприниматься», замечаться, пониматься для того, чтобы вообще что-нибудь можно было делать [8].

М. Хайдеггер называет *датируемостью* структуру соотнесенности всех «*теперь*» («сегодня»), «*тогда*» («раньше») и «*потом*» («позже»). «При этом надо пока ещё вполне отвлечься от того, осуществляется ли датировка фактично с оглядкой на календарную «дату». И без таких «дат» «теперь», «потом» и «тогда» более или менее определённо датированы. Если определённости датировки нет, то это не значит, что структура датируемости отсутствует или случайна» [4].

Если универсальными характеристиками (модусами) времени признать прошлое, будущее и настоящее, «прежде», «потом» и «теперь», а содержательным наполнением этих пустых форм – соответственно воспоминания, планы и актуальные переживания, то *связь времени (переживаемого феноменологического времени) с человеческим существованием предстаёт как очевидная*. Не менее очевидно и то, что в человеческом сознании происходит некое совмещение прошлого и будущего в настоящем, поскольку отнесённые к «своему времени» планы и воспоминания вместе с тем переживаются теперь [3].

2 Забота и феноменологическое время

Основной принцип человеческого существования, по М. Хайдеггеру, - быть в мире вещей, которые являются предметом *заботы*, то есть по отношению к окружающему миру существование выступает как озабоченность. Даже если человек всем доволен и ничего уже не хочет, он всё-таки озабочен сохранением в будущем этого состояния удовлетворённости. Иными словами, понятие «*окружающий мир*», в отличие от декартовской «геометрической» протяженности, приобретает в фундаментальной онтологии М. Хайдеггера иной смысл: *в него входит всё то, чего касаются повседневные заботы человека* [3].

«Как озаботившееся просчитывание, планирование, предусмотрение и предостережение оно всегда уже говорит, внятно в звучании или нет: «*потом*» - это надо исполнить, «*прежде того*» - другое довести до конца, «*теперь*» - наверстать, что «*тогда*» не удалось и ускользнуло. В «*потом*» озабочение выговаривается ожидая, в «*тогда*» удерживая и в «*теперь*» актуализируя... Существо озабочения временем лежит *не* в применении числовых определений при датировке» [4]. Например, мы озабочены ожиданием сделать что-то «*тогда, когда рассветёт*»; тем самым озаботившее «*тогда*» датируется исходя из восхода солнца. «Тогда, когда оно взойдёт, будет *время для...* «озабоченность» делает употребление из «подручности» дарящего свет и тепло солнца. Солнцем датируется истолкованное в озабочении время. Из этой датировки возникает «естественнейшая» мера времени, день. И поскольку временность присутствия, которое должно взять себе своё время, конечно, его дни также уже и сочтены. Это «*пока длится день*» даёт озаботившемуся ожиданию возможность заботясь вперёд определять разные «потом», чем надлежит озаботиться, т.е. подразделять день. Подразделение проводится опять же с оглядкой на то, чем датируется время: на движущееся солнце» [4].

Иными словами, *связь Я с его предметным миром раскрывается в многочисленных образах озабоченности*, а само человеческое существо, Dasein, понятие онтологически, есть забота. Именно *забота формирует предметы, составляющие предметный мир человека. Время оказывается привязанным к заботе*, из чего следует следующий важный вывод. Когда речь идёт об объективном (физическом) времени, среди его характеристик («прошлое», «настоящее» и «будущее») преимущество отдаётся «настоящему», поскольку поток времени предстаёт как бесконечная последовательность сменяющих друг друга моментов «теперь», из которых складывается столь же бесконечная череда минут, часов, лет и веков. Если же «тайна» времени коренится в «заботе», то приоритет получает как раз «*будущее*» как *пространство возможностей бытия*. Таким образом, если для «отчуждённого» времени отправной точкой выступает «теперь», настоящее, то для феноменологического времени началом координат (в свя-

зи с заботой) является будущее [3], т.е. его «стрела времени» направлена из будущего в настоящее.

Например, озабоченность болезнью близкого человека подразумевает наше желание увидеть его здоровым в ближайшем (или обозримом) будущем. Исходя из этого «начала координат», мы сегодня («теперь») находим время пригласить к больному доктора и купить лекарства, опираясь на свой опыт, предшествующий подобным случаям. В этой связи термин «временность» в философии М. Хайдеггера трактуется как будущее, поскольку оно содержит прошедшее, а также образует настоящее, ибо существование не пребывает в «теперь», а имеется всегда уже «раньше». Эта *временность и является сущностью заботы* [8].

Моё «теперь» может быть согласовано с «теперь» других Я. Однако экзистенциальным приоритетом обладает всё же моё «собственное» Я – и, значит, моё время и моё «теперь», связанное с моей заботой [3].

3 Время в онтологических моделях ситуаций

К своему предметному миру, и к множеству предметных миров других людей, других Я, человек приходит через язык. Предметный мир – это область смыслов языковых выражений, и каждый предмет не что иное, как значение слова. Каждый *предмет существует* (в качестве предмета) *лишь постольку, поскольку он связан с «вопросом»: он не что иное, как ответ на вопрос*. В этом контексте Dasein можно рассматривать как понятие фундаментальной онтологии М. Хайдеггера, которое обозначает первичное отношение человека с «иным» (которое тем самым превращается в предмет), начинается с интереса и выражается в вопросе.

Человек становится человеком тогда, когда он освоил язык и когда он с помощью языка формирует свой предметный мир. Слова же, следуя концепции Э. Гуссерля, стоявшего у истоков современной феноменологии, не столько обозначают объекты из мира независимой от человека реальности, сколько рождаются вместе с предметами человеческого интереса. «Таким образом, конституирование предметного мира человека стало выглядеть земным аналогом Божественного акта творения, и онтология предметного мира заняла место прежней онтологии сущностей в концепциях классических метафизиков» [3].

Здесь уместно обратиться к работе [4], в которой М. Хайдеггер формулирует проблему различения терминов «бытие» и «сущее» и даёт её разрешение. Бытие, по М. Хайдеггеру, – это «само собой разумеющееся понятие», тогда как понятия любого сущего являются результатом процесса изучения этого сущего, будь то чувственные объекты или математические отношения. Такое отличие бытия от сущего позволяет перейти от «объективного мира» в онтологии классической философии к миру «субъективному», «предметному», который состоит не из «вещей, свойств и отношений», а из ответов на наши вопросы. А это означает, что от «космоса» мы переходим к «логосу», от «мира природы» к «миру культуры», от состава и строения вещества к смыслу слов в предложении. Человек, являясь «интересующимся субъектом», становится сопричастным предметному содержанию языка, осваивает его, «а учение о бытии в таком случае предстаёт как «онтология субъективности» [3]: человек располагается в центре своего предметного мира.

Если возникает задача описания этого мира, то человек должен прежде всего установить *связь между собой и предметным миром*, в который входит всё то, чего касаются его повседневные заботы. Именно *озабоченность выполняет эту связующую функцию*. Озабоченность – быть озабоченным, по М. Хайдеггеру, – обычное поведение в отношении вещей, бытие которых состоит в том, что они могут служить чем-либо, благодаря им можно что-то сделать, выполнить, разрешить. Называя орудиями все те предметы, с которыми мы имеем дело в нашей повседневной озабоченности, М. Хайдеггер вводит понятия «подручности» и «наличности». Для подручности характерна её обыденность, незаметность, вследствие чего её сущность обнаруживается тогда, когда нет под рукой различных орудий. Лишь наличествующими, но не

подручными, являются вещи, с которыми нам нечего делать, которые нас не касаются, которыми мы не занимаемся [8], но которые, в принципе, могут быть нам известными. И если предметный мир человека образуют подручные и наличные вещи, то «предметную область» можно определить как совокупность только подручных предметов, необходимых человеку в некоторой *ситуации*. Тогда *персональной онтологией* можно назвать описание именно предметной области человека в форме понятий и отношений между ними [9].

Но в любой жизненной ситуации, оказывается, как правило, не один, а множество людей, которые должны *принять согласованное решение* о том, каким образом найти выход из этой ситуации. Тогда они должны сформировать разделяемую всему *групповую онтологию*, «навязанную» сложившейся ситуацией и состоящую из фрагментов персональных онтологий, которая будет выполнять роль «концептуальной платформы» для достижения взаимопонимания участников ситуации [9].

В групповой онтологии выражается общее понимание сложившейся ситуации, но отсутствует информация о конкретных параметрах, атрибутах этой ситуации. Эти сведения включаются в *объектные* (денотативные) *модели*, которые разрабатываются на основе групповых онтологий [10].

«Интегрированное» описание ситуации в форме персональных, групповых онтологий и объектных моделей в [9] предложено называть *онтологической моделью ситуации* (ОМС), подчёркивая при этом, что важную роль применительно к построению ОМС играет *фактор времени*, поскольку любая ситуация имеет тенденцию к развитию. Но о каком понимании «времени» идет речь? Из логики предыдущего изложения следует, что *в ОМС должно использоваться феноменологическое время*.

Дело в том, что описание ситуации, на первый взгляд, адресуется к «настоящему», к тому, что происходит сейчас, «теперь». И с этим можно было бы согласиться, если предположить, что описание совершает дистанцированный от ситуации внешний наблюдатель, не имеющий к ней никакого отношения, то есть в том случае, когда реализуется важнейший принцип естествознания – «оппозиция субъекта и объекта». В нашем же случае субъект со всеми его заботами «встроен» в ситуацию и переживает её; то есть налицо «суперпозиция субъекта и объекта».

Тогда «прошлое» уже нельзя трактовать как то, чего уже нет, а следует понимать как то, что постоянно присутствует в «настоящем». Но точно также в «настоящем» человек осознает и своё собственное «будущее»; его планы влияют на оценку текущей ситуации. Таким образом, и «прошлое», и «настоящее», и «будущее», являющиеся абстракциями от естественного (феноменологического) времени [8], должны использоваться при построении ОМС в процессах принятия решений.

Каждый человек, осознавая субъективно («экзистенциально») феноменологическое время, в своих рассуждениях датирует его с помощью понятий «тогда», «теперь», «потом» и т.п. Трудно себе представить, чтобы лица, участвующие в подготовке и принятии решений, постоянно оперировали исключительно астрономическими датами – днями, часами, минутами и секундами. Конечно же, это не означает, что объективное время, измеряемое отрезками пути небесных тел, вовсе не используется [8]. Просто человек естественным образом осознает прежде всего феноменологическое время, предшествующее применению часовых механизмов.

А это означает, что описание ситуации в ОМС должно включать в себя формулировки «озабоченности» людей, оказавшихся в этой ситуации, и «привязанное» к ним феноменологическое время. Причём необходимость принятия коллегиальных решений [11] выдвигает на передний план задачу *согласования «озабоченностей»* людей, то есть выявление их «общих забот», и *феноменологических* (в отличие от календарных) *дат*, что должно привести к построению единой структуры соотнесённости всех «теперь», «тогда» и «потом».

Заключение

Понятие феноменологического времени, связанное с осознанием времени человеком и содержанием его переживаний, должно использоваться в процессах принятия решений, поскольку оно предшествует применению «объективного» физического времени. Обладая памятью, человек в своём настоящем «содержит» своё прошлое и «воображает» своё собственное будущее, т.е. человек – времяобразующее сущее. Поскольку, по М. Хайдеггеру, бытие – это прежде всего человеческое бытие, то окружающий мир – это предметный мир человека.

Пространство представляет собой одно из измерений этого предметного мира, но воспоминания и ожидания открывают нам ещё и временное измерение нашего предметного мира и, соответственно, временную структуру нашего собственного сознания. Это и есть феноменологическое время, отличное от понятия времени как «всеобщей формы бытия материи». Всё, относящееся к феноменологическому времени, таким образом, есть наше, человеческое истолкование.

Находясь в некоторой жизненной ситуации, переживая её и будучи озабоченным её развитием, человек (или группа людей) должен принимать решения о том, каким образом изменить её в требуемом направлении, для чего необходимо иметь описание ситуации. С этой целью предлагается использовать онтологическую модель ситуации, в которой важную роль играет феноменологическое время. Онтологическая модель ситуации должна включать в себя формулировки «озабоченностей» людей, оказавшихся в этой ситуации, и «привязанное» к ним феноменологическое время. В случае принятия коллегиальных решений на передний план выдвигается задача выявления «общих забот» и согласования феноменологических дат.

Благодарность

Автор выражает благодарность д.т.н. С.В. Смирнову за ценные замечания и предложения, высказанные при обсуждении рукописи данной статьи.

Список литературы

- [1] Словарь философских терминов. – М.: ИНФРА-М, 2004.
- [2] Виттих В.А. Парадигма ограниченной рациональности принятия решений – 1 // Вестник Самарского государственного технического университета (серия «Технические науки»), №3 (25), 2009, с. 22–31.
- [3] Зотов А.Ф. Современная западная философия. – М.: Проспект, 2010.
- [4] Хайдеггер М. Бытие и время. – Харьков: «Фолио», 2003.
- [5] Философия (учебник) – М.: Гардарики, 2004.
- [6] Хайдеггер М. Время и бытие (статьи и выступления). – Санкт - Петербург, «Наука», 2007.
- [7] Камке В.А. Современная философия. – М.: Издательство «Омега-Л», 2010.
- [8] Философский энциклопедический словарь. – М.: ИНФРА-М, 1997.
- [9] Виттих В.А. Парадигма ограниченной рациональности принятия решений – 2 // Вестник Самарского государственного технического университета (серия «Технические науки»), №1 (26), 2010, с. 23-31.
- [10] Смирнов С.В. Онтологический анализ предметных областей моделирования // Известия Самарского научного центра РАН, том 3, №1, 2001, с. 62-70.
- [11] Виттих В.А. Онтологические модели ситуаций в процессах принятия коллегиальных решений // Проблемы управления и моделирования в сложных системах: Труды XI Международной конференции. – Самара, Самарский научный центр РАН, 2009, с.405-410.

СИСТЕМЫ ОПЕРАТИВНОЙ ОБРАБОТКИ И АНАЛИЗА ОБРАЩЕНИЙ ГРАЖДАН КАК ЭФФЕКТИВНОЕ СРЕДСТВО ПОДДЕРЖКИ РЕШЕНИЙ В СИСТЕМЕ ГОСУДАРСТВЕННОГО И МУНИЦИПАЛЬНОГО УПРАВЛЕНИЯ

А.Н. Визгунов, М.Л. Zubov, А.В. Шишин

Государственный университет - Высшая школа экономики (Нижний Новгород)
603155, Нижний Новгород, ул. Б.Печерская, 25/12_Россия
arseniy.vizgunov@gmail.com, mzubov@hse.ru, anatoly.shishin@bk.ru
тел: +7 (831) 416-95-49

***Ключевые слова:** семантическая обработка сообщений, органы государственного управления, органы муниципального управления, системы поддержки принятия решений*

Abstract

The main goal of the article is to discover the ways to build the platform which supports collecting, storing, retrieval, on line and off line analysis and processing of citizens' appeals to municipal government. The developed platform could be used by responsible persons and decision makers from municipal government.

Введение

В последнее время одной из самых обсуждаемых в обществе тем является налаживание взаимодействия между властью и гражданами. Это обусловлено тем, что именно отсутствие такого взаимодействия лежит в основе тех конфликтных ситуаций, которых очень много в нашей стране.

Конечно, в целом ряде конфликтов есть конкретные виновники, которые принимают конкретные решения, с умыслом нанося вред интересам общества. Такие ситуации нам неинтересны, они рассматриваются в рамках совсем других исследований – уголовных. Также мы не рассматриваем ситуации, когда управленческое решение поддерживается подавляющим большинством граждан, например, принятие решения о строительстве нового моста через реку или решение о ремонте дорог или, наоборот, не поддерживается абсолютным большинством граждан, например, повышение ставки транспортного налога.

В рамках наших исследований нам интересны ситуации, когда чиновники, обладая доброй волей (которая есть у большинства чиновников в России), пытаются выработать и претворить в жизнь комплекс мер, направленных на решение проблем в тех областях деятельности, за которые они отвечают.

Первым пунктом плана выработки такого комплекса должен быть анализ мнения граждан по данному вопросу. К сожалению, способов выполнить этот анализ в настоящий момент совсем немного. Использование таких способов изучения общественного мнения как референдум или социологический опрос (который можно рассматривать как мини-референдум), очень дорого и возможно только в ограниченных случаях. Получается, что единственным способом является анализ средств массовой информации, собственный опыт и консультации с представителями граждан – депутатами всех уровней. Все эти способы получения информации не являются точными и надежными. Получается, что решения принимаются в некотором предположении о том, что думают граждане по данному вопросу. В этом случае, конечно, есть большая вероятность совершения ошибки, принятия ошибочных решений.

Если посмотреть на ту же самую ситуацию со стороны гражданина, то она еще хуже. Граждане в настоящий момент не имеют возможности не только участвовать в процессе при-

нятия решения, но и не могут донести до чиновников свою оценку исполнения уже принятых решений. У гражданина есть следующие способы высказать свое мнение: записаться на прием к депутату или непосредственно к чиновнику, написать письменное обращение в некоторой форме, выйти на митинг. Из приведенного списка понятно, что все перечисленные действия гражданин может совершить только в том случае, если есть решения, которые оказывают на него сильное негативное воздействие. Трудно себе представить, что человек будет записываться на прием к депутату для того, чтобы высказать свое удовлетворение по поводу тщательно отремонтированного подъезда. Даже если это и произошло, то совсем непонятно, что будет делать депутат в этом случае. Вряд ли он будет искать чиновника, который организовывал ремонт подъезда для того, чтобы передать ему благодарность своего избирателя. А ведь это очень важная информация о том, что определенный чиновник добросовестно выполнил свою работу.

Единственным выходом из описанной ситуации является построение эффективных информационных систем оперативной обработки, хранения и анализа обращений граждан. Обращения являются действенным средством обратной связи граждан и органов власти, позволяют включить в процесс анализа и поиска решений заинтересованных граждан, проводить проверку на адекватность действий властных структур в ответ на обращения. Однако для того чтобы подобные информационные системы были эффективны и успешно решали поставленные задачи, требуется решить значительное число научных и технических проблем.

Построение эффективной системы обработки обращений граждан

С нашей точки зрения для построения эффективной системы обработки обращений граждан следует использовать комбинацию следующих существующих средств современных информационных технологий:

- возможность общения людей посредством компьютерной сети, без необходимости личной встречи;
- системы электронного документооборота;
- алгоритмы семантической обработки текстов и теория онтологий;
- построение развитого пользовательского интерфейса «тонкого» клиента и сервис-ориентированная архитектура сервера;
- хранилища данных и средства анализа данных, средства анализа текстовой информации

Разберем по порядку, какой вклад вносит в предлагаемое решение каждая из перечисленных выше технологий.

Возможность общения между людьми без непосредственного визуального контакта является основной для решения нашей задачи. Если для того чтобы высказать свое мнение гражданин должен прийти в определенный час в определенное место, то граждане по-прежнему не будут высказывать свое мнение часто и по-прежнему, большая часть высказанных мнений будет негативной. Чтобы собрать большое количество откликов за короткое время нужно создать систему, которая позволяет гражданину высказать свое мнение в тот момент и в том месте, где это удобно гражданину. Иными словами, гражданин должен иметь возможность высказать свое мнение путем отправки своего мнения в виде электронного сообщения. В лучшем случае это сообщение может быть отправлено с любого компьютера, подключенного к Интернет, в нормальном случае – с некоторых специализированных терминалов, к которым открыт круглосуточный свободный доступ [5].

Использование специализированных терминалов имеет свои плюсы и минусы. Плюс этого решения заключается в том, что автоматически решается задача ответственности гражданина за информацию, содержащуюся в его обращении. Минус состоит в том, что до терминала все-таки надо идти или ехать, терминал может не работать или быть занятым другим гражданином, если он занят, то придется ждать, если сломан, то идти к другому терминалу, что еще

больше увеличит затрачиваемое время. Нетрудно предсказать, что чем больше будет время, которое нужно затратить на посылку сообщения, тем меньше сообщений будет отправлено гражданами.

С нашей точки зрения более правильно использовать для посылки сообщения сеть Интернет, хотя это и приводит к дополнительным сложностям, или использовать оба варианта, как дополняющие друг друга.

До этого момента наш подход почти не отличается от того, который широко используется в настоящее время в так называемых «интернет-приемных». Интернет-приемные есть на многих сайтах государственных организаций. В качестве примера можно привести сайты администрации Нижнего Новгорода [6] и проект «Общественная приемная Минэкономразвития России» [4]. Интернет-приемная, чаще всего, представляет собой форму, заполняемую в Интернет-браузере. В такой форме пользователю предлагается ввести данные по шаблону или из списка в предопределенный набор ряд полей, а также в свободной форме доопределить суть своего обращения. Обращение потом поступает в систему электронного документооборота и проходит стандартный путь, определенный в регламенте. В результате выносится некоторое решение и через некоторое время гражданин получает его в качестве ответа по электронной почте. Интернет-приемная чаще всего интегрирована с развитыми системами диспетчеризации сообщений, которые выполняют ряд стандартных функций по регистрации и обработке электронных обращений граждан и заявок на прием к руководителям.

Основной проблемой Интернет-приемной, с нашей точки зрения, является то, что такой способ фиксации обращений граждан не является удобным. Поскольку обращение пишется в свободной форме, быстро определить, какая именно проблема описана в письме очень тяжело. Приходится затрачивать большое количество времени и, возможно, даже обращаться к гражданину за разъяснениями. Поскольку сделать это не всегда возможно, то в конечном итоге для получения решения, возможно, потребуется больше времени, чем если бы гражданин пришел лично. Получение большого количества обращений через Интернет-приемную без надлежащей обработки приведет к тому, что чиновники будут перегружены запросами.

Чтобы преодолеть эту проблему, мы предлагаем применить систему интеллектуальной обработки обращений. При налаженной работе такой системы обращения от граждан, поданные в электронном виде, обрабатываются гораздо более оперативно, чем обращения, поданные в виде бумажного письма или при личной встрече. Кроме того, все обращения и результаты работы по этим обращениям хранятся в аналитической базе и служат основой для совершенствования алгоритмов разбора обращений и для моделирования последствий принимаемых решений.

Предлагаемая нами система поддержки принятия решений должна способствовать развитию диалога граждан и органов государственной властью. Поскольку полноценный диалог подразумевает эффективное двустороннее взаимодействие, то каждой категории пользователей, с одной стороны, должен быть предоставлен удобный интерфейс и организовано индивидуальное рабочее пространство (виртуальный личный кабинет), с другой – разработан набор сервисов, дополняющий базовые функции подачи и регистрации электронного обращения.

На рисунке 1 изображена подсистема приема и хранения заявок, реализованная на базе Интернет-портала. К данной подсистеме имеют доступ как граждане, так и чиновники, каждая из групп пользователей, конечно, должна обладать определенными правами доступа к информации. Подсистема приема и хранения заявок передает данные в аналитическую базу – в хранилище данных. Хранилище данных используется совместно с оперативной базой сервисами подсистемы анализа заявок.

Разработка сервисов по извлечению значимой информации непосредственно из текстов электронных обращений требует работы с сырыми данными на уровне слов и предложений и привлечения методов морфологического, синтаксического и семантического анализа, эвристи-

ческих алгоритмов анализа неструктурированных текстов и извлечения субъективной информации из них (анализ мнений, оценка настроений).

Анализ входящего потока электронных обращений граждан является частной задачей проблемы автоматической обработки текстов на естественном языке. Одна из основных целей такого анализа есть представление неструктурированной текстовой информации в структурированном виде, что дает возможность в рамках системы принятия решений проводить оперативный анализ данных.

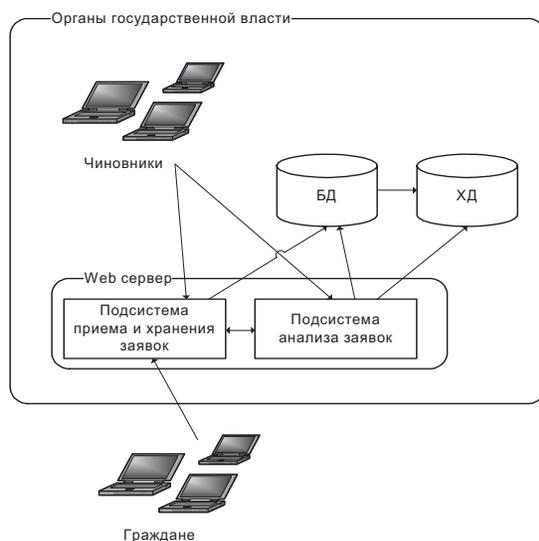


Рисунок 1 – Архитектура предлагаемой программной системы

Каждое обращение является текстовым описанием некоторой ситуации, и задача подсистемы семантического анализа заключается в извлечении в автоматическом режиме максимального объема полезной информации и фиксации ее в виде дополнительных атрибутов электронного обращения, т.е. снабжать наиболее полезными для человека комментариями.

Таковыми атрибутами могут быть:

- категория обращения в терминах таксономии предметной области;
- оценка субъективной информации обращения (негативный, позитивный или нейтральный фон), включая оценку степени удовлетворенности гражданином официальным ответом на обращение;
- факты различного типа, связанные с заданными персонами, организациями или районами местности.

Наличие таких атрибутов дает аналитику возможность проводить комплексный анализ в рамках заданной предметной области и составлять сложные статистические отчеты в автоматическом режиме, например, характеризующие степень удовлетворенности граждан работой конкретной службы в заданной районе города, распределение обращений по категориям и т.д.

Семантический, а затем и прагматический анализ текста основывается на сопоставлении смысловых конструкций текста с некоторыми априорными знаниями об окружающем мире и на анализе самой структуры текста. Таким образом, выбор формы представления знаний оказывает определяющее значение. Рассматриваемая в данной работе система основывается на базе знаний, представляемой онтологиями.

На основе анализа репрезентативной выборки обращений граждан в органы муниципальной власти группой экспертов была разработана онтология проблемных областей (фрагмент приведен на рисунке 2).

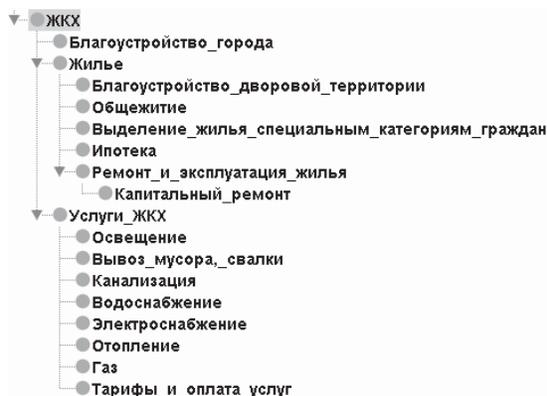


Рисунок 2 – Фрагмент онтологии проблемных областей. Категория ЖКХ

Автоматизация бизнес-процесса обработки заявок может быть проведена за счет добавления в систему сервисов, осуществляющих:

- 1) автоматическую классификацию обращений в терминах разработанной онтологии проблемных областей,
- 2) выбор ответственного исполнителя в терминах онтологии организационной структуры.

Для решения задачи автоматической классификации обращений нами были выбраны основанные на обучении методы: опорных векторов, наивный байесовский подход и др. [1].

В процессе опытной эксплуатации системы должны быть выявлены наиболее эффективные алгоритмы классификации, демонстрирующие лучшие показатели при работе с короткими текстовыми сообщениями.

Сфера ответственности и область компетенции подразделений из онтологии организационной структуры может быть представлена на онтологии проблемных областей множеством узлов. Предполагаем, что перечислением категорий обращений можно задать сферу компетенции практически каждой структурной единицы организации, непосредственно вовлеченной в процесс обработки обращений граждан. Такое множество ключевых категорий (узлов таксономии) будем называть профилем организационной единицы (отдела, сотрудника и т.д.), выраженным в терминах онтологии проблемных областей. Таким образом, система на основе заложенных в нее знаний и извлеченной из текстов информации способна автоматически перераспределять поступающие обращения в соответствии с тематикой запроса и профилями исполнителей.

Объединяя таким образом две и более онтологий, мы создаем ограниченное представление предметной области, в терминах которой осуществляют обработку данных все сервисы подсистемы семантического анализа. Другими словами используемые онтологии определяют множество типов объектов, отношений между ними, а также событий, в которых участвуют объекты.

Извлечение субъективной информации предполагает привлечение эвристических методов оценки мнений и настроений [2], предназначенных для извлечения из текста оценочных суждений, субъективного отношения к какому-либо объекту или явлению.

В процессе эксплуатации системы база знаний, основанная на онтологиях, может содержать разнообразную полезную дополнительную информацию. Так с каждым понятием можно связать множество выражающих его слов, употребленных в разных контекстах. Для формирования такой расширенной базы знаний можно использовать накопленный человечеством опыт, выраженный в текстах на естественном языке, и использовать для этой цели общедоступную сеть Интернет. Определенное выше и ассоциированное с конкретным понятием множество,

может включать как синонимичные, так и любые другие слова, косвенно сообщающие, что речь в тексте идет об интересующем предмете.

Поскольку большинство активно используемых сегодня методов классификации и анализа мнений основаны на пословном представлении документов, то одним из факторов, влияющих на качество результата, является репрезентативность семантической окрестности понятий онтологии предметной области. Семантическая окрестность представляется множеством слов, имеющих веса, рассчитанные на основе частоты их встречаемости в разных контекстах и характеризующих степень их характерности выбранному понятию онтологии. С формирования таких окрестностей должен начинаться процесс обучения, когда ключевые слова сначала выбираются на основе анализа отклика поисковой системы Интернет в ответ на запрос с названием узла онтологии, а затем пополняются из обучающей выборки электронных обращений.

Интегрированная в онтологии возможность с помощью правил логического вывода позволяет осуществлять проверку гипотез и выявлять факты, незафиксированные в явном виде в тексте, но следующие из контекста, является ключевой особенностью выбранной формы представления знаний. Так, например, идентифицированные в процессе анализа обращений географические наименования (представлены в географической онтологии определенной местности), имена персон, а также ключевые ситуации, где упоминаются указанные сущности, могут быть использованы для построения сложных аналитических отчетов.

Пользователям другой категории – гражданам, в процессе редактирования обращения могут быть предложены такие сервисы как поиск родственных заявок, описывающих одну и ту же проблемную ситуацию, использование шаблонов и создание коллективных сообщений. На основе метрики близости на онтологии [3] система может выдавать рекомендации в виде существующих публичных обращений, описывающих похожие проблемы.

Заключение

Задача анализа поступающих обращений граждан является настолько сложной, что многие граждане и чиновники считают ее вовсе нерешаемой. С нашей точки зрения, в настоящее время эту задачу можно решить используя существующие информационные технологии.

Предложенный в статье подход позволяет создать механизм эффективного взаимодействия между гражданами и органами власти, получить адекватные, точные данные по состоянию дел путем включения граждан в процесс подготовки и оценки решений.

Работа была выполнена при поддержке Центра Фундаментальных Исследований ГУ-ВШЭ (проект ТЗ-29.0).

Список литературы

- [1] Ian H. Witten, Eibe Frank. Data Mining: Practical Machine Learning Tools and Techniques (Second Edition) - Morgan Kufmann, 2005
- [2] Liu Bing. Web Data Mining. – Springer, 2007
- [3] Michael Ricklefs and Eva Blomqvist Ontology-Based Relevance Assessment: An Evaluation of Different Semantic Similarity Measures. OTM Conferences (2) 2008: 1235-1252
- [4] Автоматизированная информационная система «Общественная приемная Минэкономразвития России» // Сайт компании Microsoft 4.06.2008. URL: http://www.microsoft.com/Rus/CaseStudies_archive/CaseStudy.aspx?id=568 (Дата обращения 24.03.2010).
- [5] А. Окишев. Владимир Виттих: «Электронное правительство надо создавать в рамках административной реформы» // Сайт РИА «Самара» 5.09.2007. URL: http://www.riamara.ru/rus/analytics_and_comments/politics/article13636.shtml (Дата обращения 24.03.2010).
- [6] Интернет-приемная // Официальный сайт городской администрации города Нижнего Новгорода 24.03.2010. URL: <http://www.admgor.nnov.ru/script/questions.php?item=29> (Дата обращения 24.03.2010).

ИСПОЛЬЗОВАНИЕ ИМИТАЦИОННЫХ МОДЕЛЕЙ ДЛЯ ПРОВЕРКИ ВЫПОЛНИМОСТИ АДМИНИСТРАТИВНЫХ РЕГЛАМЕНТОВ

Н.В. Асеева, Э.А. Бабкин, А.Н. Визгунов, В.П. Князькин, О.Р. Набиуллин, М.С. Шиткова

Государственный университет - Высшая школа экономики (Нижний Новгород)
603155, Нижний Новгород, ул. Б.Печерская, 25/12, Россия
eababkin@hse.ru; arseniy.vizgunov@gmail.com
тел: +7 (831) 416-95-49

***Ключевые слова:** административные регламенты, органы государственного управления, органы муниципального управления, имитационные модели, системы поддержки принятия решений*

Abstract

The subject of the research work is to use the simulation models to evaluate the performance of the enterprise regulations. New modeling language and simulation framework are developed. The developed platform could be used by responsible persons and decision makers from the municipal government.

Введение

В рамках проводимых в настоящее время работ по определению перечня и порядка предоставления государственных услуг одной из важнейших задач является разработка методов оценки реалистичности административных регламентов.

Нами предлагается комплексный подход, заключающийся в построении модели регламента и существующей организационной структуры организации, оказывающей услуги, и использования этой модели при проведении имитационных экспериментов. Выбранный подход позволяет бизнес-аналитику или непосредственно работникам администрации создавать достаточно точные и концептуально-непротиворечивые модели регламента, которые затем автоматизированным образом преобразуются в программный код многоагентной имитационной модели. В ходе имитационных экспериментов с этой моделью появляется возможность оперативной проверки непротиворечивости, выполнимости при заданных ресурсных ограничениях, устойчивости при изменении параметров, возможность сформулировать рекомендации по изменению организационной структуры.

Для достижения поставленной цели были решены следующие задачи:

- проанализированы тексты нескольких регламентов, выделена общая структура регламента;
- проведено критическое рассмотрение существующих средств моделирования процессов административного управления;
- создан свой специфический язык описания бизнес-процессов;
- построена математическая модель процесса имитационного моделирования административных регламентов;
- проведен ряд пробных имитационных экспериментов.

В настоящее время ведутся работы по совершенствованию имитационных сервера и клиента и проведению серии имитационных экспериментов для реальных регламентов Нижнего Новгорода.

Анализ текста ряда регламентов Нижнего Новгорода

Административный регламент представляет собой описание действия или совокупности действий исполнительного органа государственной власти, его структурных подразделений и должностных лиц, направленных на осуществление их полномочий, установленных законом, иным нормативным правовым актом. Другими словами, любой административный регламент описывает процедуру предоставления государственной услуги гражданам и регулирует порядок её выполнения.

Проанализировав большое количество доступных административных регламентов, можно выделить структуру регламента, его основные компоненты. Каждый административный регламент состоит из следующих частей:

- 1) Общие положения
 - наименование государственной услуги;
 - нормативно-правовое регулирование предоставления услуги. В данном разделе перечисляются все Федеральные законы, постановления Правительства, положения, правила и законы других органов власти, в соответствии с которыми осуществляется оказание государственной услуги;
 - орган, предоставляющий государственную услугу;
 - заявители, имеющие право на получение государственной услуги.
- 2) Требования к порядку предоставления государственной услуги
 - порядок информирования о государственной услуге;
 - перечень документов, необходимых для получения государственной услуги;
 - условия и сроки предоставления государственной услуги;
 - перечень оснований для отказа в предоставлении государственной услуги;
 - другие положения, характеризующие требования к предоставлению госуслуги (например, размеры сборов за предоставление государственной услуги).
- 3) Административные процедуры
 - описание последовательности действий при предоставлении государственной услуги;
 - порядок и форма контроля за предоставлением государственной услуги;
 - порядок обжалования действий (бездействия), осуществляемых в ходе исполнения Регламента, и принятых решений.
- 4) Приложения.

Как видно из структуры регламента, в нём есть все необходимые пункты и разделы для моделирования административного процесса, а, следовательно, может быть применен широкий выбор методологий моделирования.

Рассмотрение существующих средств моделирования процессов административного управления

На сегодняшний день известно большое количество методологий и нотаций моделирования «общей направленности», таких как ARIS [1, 2], DEMO [3], IDEF [4], BPMN [5] которые являются достаточно гибкими и позволяют выполнять моделирование в различных областях/ Кроме того известен ряд предметно-ориентированных методологий, созданных специально для государственного административного сектора, таких как методология PICTURE [5], разработанная в университете города Мюнстер (Германия), или российский продукт ГОС-Мастер, компании "Бизнес Инжиниринг Групп", созданный при поддержке Государственного университета – Высшей школы экономики.

В соответствии с проведенным исследованием по целому ряду критериев на сегодняшний день среди наиболее адекватных подходов к моделированию административных процессов можно выделить DEMO [1] и BPMN [5]. К сожалению, даже эти методы при всей своей выра-

зительной силе не могут полностью решить задачу автоматизированного получения имитационной модели по формальному описанию. Например, диаграммы DEMO приходится нетривиальным образом дополнять более простыми, одноуровневыми, но подробными диаграммами для точного понимания процесса имитационного моделирования. Использовать BPMN без внесения изменений также не удастся по причине следующих недостатков этого подхода:

- Невозможно отобразить временной аспект. В нашей задаче крайне важным с точки зрения имитационного моделирования является наглядное отображение времени выполнения шагов процесса и время передачи задания от одного шага/исполнителя другому.
- Невозможно добавить дополнительную информацию об организационной структуре. В BPMN могут быть представлены только роли. В нашей задаче необходимо также представить конкретных людей в организации, ответственных за выполнение тех или иных задач. Необходимо указать связь с ролями, их контактные данные и обязанности.
- Невозможно добавить дополнительную информацию о процессе. Каждое действие соответствует определенному пункту регламента. Это было бы полезно отобразить в модели. С существующими подходами эту информацию можно лишь добавить в название действия, что значительно усложняет модель.
- Невозможно добавить дополнительную информацию о документах (элементе данных). Можно указать лишь один элемент данных, который не говорит о гранулярности – является ли это документом или набором документов. Если использовать элемент данных только для представления одного элемента в некоторых случаях возникнет необходимость связать достаточно много таких документов с одним шагом процесса. Это сделает модель непонятной. Также у каждого документа должен существовать определенный набор атрибутов, которые в данный момент невозможно привязать ни к одному элементу данных.
- Существует определенная проблема с уменьшением сложности представления модели. В основном существующие языки моделирования предлагают использовать иерархичность представления, т.е. добавлять подпроцессы для сложных задач (состоящих из достаточного количества пунктов) или «черные ящики» для частей, которые невозможно смоделировать. При использовании подпроцессов мы в итоге получим несколько диаграмм, которые не очень удобно рассматривать. Они будут не связаны между собой, если не возвращаться постоянно к более верхнему уровню.

Приведенный выше анализ показывает необходимость модификации существующих методик с целью создания нового языка моделирования, пригодного для достижения поставленных целей.

Описание предлагаемого авторами языка описания бизнес-процессов

В рамках нового языка предлагается моделирование регламента разделить на две части – модель регламента верхнего уровня и процессная модель регламента. Модель регламента верхнего уровня должна описывать всю услугу в целом, предоставлять общую картину того, чему посвящена услуга, кто ее выполняет и что необходимо для получения этой услуги. Другими словами, необходимо представить на одной диаграмме следующие аспекты:

- процесс информирования об услуге;
- заявитель (кто может быть заявителем и каковы условия начала выполнения регламента);
- процесс выполнения услуги (некоторое достаточно общее описание из 4-5 основных шагов);
- организационная структура, ответственная за выполнения, участвующая в выполнении (наименование департамента и основных отделов, участвующих в исполнении услуги, а так же роли, выполняемые участниками);
- контроль выполнения услуги (нормативные документы и организационная структура, ответственная за контроль основных этапов предоставления услуги);

- ожидаемый результат выполнения услуги (в случае успешного завершения);
- ошибочный результат (когда выполнение процесса завершилось неуспешно);
- процесс обжалования результата выполнения услуги (в случае ошибочного результата и наличия оснований для обжалования).

На данном этапе развития проекта основным является детальное моделирование процесса выполнения административных процедур в рамках предоставления государственной или муниципальной услуги.

На основе анализа структуры регламента были выделены следующие основные концепции, которые должны быть представлены в модели административной процедуры:

- процесс;
- роли и организационная структура;
- документы;
- временной аспект.

Первые три элемента встречаются практически во всех методологиях моделирования бизнеса. Так как начинать создания нового языка моделирования с чистого листа практически бессмысленно, то наиболее удачные решения для этих элементов было решено позаимствовать из уже существующих нотаций (BPMN, EPC, IDEF3, UML и т.п.).

Временной аспект крайне важен в нашей задаче, так как это один из наиболее значительных параметров, используемых в имитационной модели. Временной аспект делится на две части:

- время, необходимое для выполнения действия (Activity Time);
- время проходящее между концом одного действия и началом следующего действия (Transition Time).

Очевидно, что оба временных аспекта должны некоторым образом быть связаны с задачами, отображаемыми в модели процесса.

В нашем проекте для реализации построения собственного языка моделирования регламентов был выбрано средство моделирования Generic Modeling Environment (GME) [7]. Это настраиваемое средство моделирования позволяет создавать собственные предметно-ориентированные языки, используя нотацию UML. Созданный предметно-ориентированный язык может быть зарегистрирован в качестве парадигмы и в последствии использован для создания моделей предметной области.

При построении собственного языка моделирования активно используется концепция аспектов, представленная в программном продукте GME. В частности реализованы следующие аспекты:

- аспект процесса + роли в виде swimлайнов;
- аспект организационной структуры;
- аспект процесса + документов;
- аспект процесса + времени.

Результатом применения созданного языка моделирования является построение модели регламента, пригодное для подачи на вход разработанной системе имитационного моделирования.

Система имитационного моделирования

Систему для имитационного моделирования административных регламентов удобно представить в виде многоагентной модели. Такой подход позволяет проводить анализ регламента с высокой степенью детализации, т.е. анализировать события и задержки на уровне отдельного должностного лица. Инструментарий, используемый для создания многоагентных имитационных систем характеризуется следующими чертами:

- процесс моделирования основывается на математических методах, при этом позволяя осуществлять автоматическую проверку соответствия математической модели много-агентной реализации;
- ядро многоагентной системы обеспечивает высокую производительность имитации за счет построения распределенных экспериментов и использования оптимизированного кода, выполняющегося напрямую процессором;
- существуют средства исследования корректности, причем как математической модели, так и реализации.

В рамках работ над проектом авторами разработан программный каркас, который удовлетворяет перечисленным требованиям. Верификация имитационных моделей осуществляется с помощью математических метамоделей. Мета модель описывается с помощью формализма машин действий.

Для моделирования административных регламентов предлагается использовать математическую модель, описанную ниже. Можно заметить, что математическая модель регламента представляет собой специализированную систему массового обслуживания. В созданной системе предполагается, что бизнес-процесс (регламент) состоит из некоторого количества блоков-активностей, выполняемых людьми.

Минимальной структурной единицей процесса является сообщение. Характеристикой сообщения является сложность Md . В данной модели считается, что результатом обработки сообщения может являться только это же самое сообщение, т.е. невозможны ситуации, когда после некоторой активности появляются два производных сообщения.

Время T_p , затрачиваемое на преодоление сообщением фрагмента процесса (на выполнение некоторой активности человеком p_i) зависит от профессиональных навыков Rc_i , и сложности обрабатываемой информации Md . Время передачи информации является ступенчатой функцией T_i и зависит только от взаиморасположения p_1 и p_2 , обозначаемого Rsh_2^1 . При этом возможны четыре варианта, в порядке возрастания времени передачи:

- а) обе роли выполняются одним человеком, время передачи минимально;
- б) роли выполняются разными людьми, находящимися в одном отделе;
- в) роли выполняются разными людьми из разных отделов;
- г) роли выполняются разными людьми из разных подразделений.

Сложность информации кодируется параметром Md .

В случае разветвления процесса указываются вероятность $Pr_{ij} = f(A_i, A_j), \forall i \sum_j Pr_{ij} = 1$ того, что после действия A_i будет выбрано A_j .

Исходной информацией для моделирования помимо бизнес-процесса, описанного на разработанном авторами языке, является матрица способностей Ap , описывающая способность человека p_1 выполнять активность A_j . В данной матрице строки задают список способностей человека выполнять роли, а столбцы являются описанием ролей.

Человек определяется своей компетенцией и уровнем затрачиваемых на него ресурсов $\rho_i = \langle Rc_i, S_i \rangle$, причем под ресурсом понимается как заработная плата, так и прочие расходы, связанные с выполнением человеком своих обязанностей. В первом приближении можно считать, что ресурсы являются линейной функцией от компетенции, т.е. $S_i = \alpha * Rc_i + \beta$.

Оптимизация параметров модели осуществляется с помощью манипулирования взаиморасположением людей Rsh_2^1 , выполняющих определенные роли с учетом ограничений, накладываемых матрицей Ap , а также с учетом ограниченности нагрузки на одного человека в течение дня. Для упрощения модели предполагается, что круг служебных обязанностей человека определяется раз и навсегда, т.е. человек не может выполнять по четным дням роли A_1, A_2 , а по нечетным A_3 и A_4 .

Возможны несколько подходов к выбору критерия оптимизации модели. С точки зрения властей интерес представляет минимизация расходов при соблюдении ограничений:

$$\sum S_i \rightarrow \min,$$

$$\sum T \leq T^*,$$

где T^* - максимальное время обработки обращения, прописанное в законе.

С точки зрения граждан требуется минимизировать среднее время прохождения обращения при заданном бюджете

$$\sum T \rightarrow \min,$$

$$\sum S_i \leq B,$$

где B означает бюджет организации.

Результатом имитационного эксперимента, построенного для описанной выше модели, являются таблица–расписание обработанных сообщений, а также список нарушений правил, если они имели место.

Заключение

Предложенный авторами подход к решению поставленной задачи позволяет просчитать заранее нагрузку на чиновников организаций, занятых в процессе предоставления государственных услуг. Важность возможности сделать количественную оценку степени выполнимости регламентов невозможно переоценить. Авторы надеются, что предлагаемый подход будет активно использоваться при разработке, апробации и внедрении административных регламентов в органах государственной власти.

Работа выполнена при поддержке научного фонда НФ ГУ-ВШЭ (проект «Синтез и практическое внедрение методологии информационно-имитационного моделирования административной деятельности муниципальных органов управления»).

Список литературы

- [1] Davis R., Brabandere E. ARIS Design Platform. Springer; London, 2007. 364p
- [2] Моделирование бизнеса. Методология ARIS. / Каменнова М., Громов А., Ферапонтов М., Шматлюк А.// М.: Весть-МетаТехнология, 2001. 327 с.
- [3] Dietz J. Enterprise Ontology: Theory and Methodology. N. Y.: Springer, 2006. 243p.
- [4] Репин В. ARIS Toolset/ВРwin: выбор за аналитиком // КомпьютерПресс. – 2002. - № 1.
- [5] BPMN official web-site [Электронный ресурс]. – Режим доступа: <http://www.bpmn.org/> (Дата обращения: 25.12.09)
- [6] Becker, J. Pfeiffer, D. and Räckers, M. A Process Modeling Language for the Public Sector - the PICTURE Approach.// Wybrane Problemy Elektronicznej Gospodarki, Łódź. – 2007. – p. 271-281.
- [7] Ledeczi A., Maroti M., Bakay A., Karsai G, Garrett J., Thomason C., Nordstrom G., Sprinkle J., and Volgyesi P. The Generic Modeling Environment.//Workshop on Intelligent Signal Processing. 2001.

О ПРОБЛЕМАХ ПРИНЯТИЯ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ В АВТОМАТИЗИРОВАННОЙ СИСТЕМЕ ЭКОЛОГИЧЕСКОГО МОНИТОРИНГА И УПРАВЛЕНИЯ Г. НОВОМОСКОВСКА ТУЛЬСКОЙ ОБЛАСТИ

В.Ю. Волков, В.В. Батышкина, Самаха Башир Аббас

Новомосковский институт (филиал) ГОУ ВПО РХТУ им. Д.И. Менделеева
301665, Новомосковск, Тульская область, ул. Дружбы, 8, Россия
duga@mail.ru, vbatyshkina@mail.ru
тел: +7 (48762) 6-06-76, 6-12-50

***Ключевые слова:** атмосферный воздух, концентрация загрязняющих веществ, автоматизированная система контроля, когнитивная модель, модели технического анализа, принятие решений*

Abstract

Possible implementation of dynamic forecasting and control functions in real time is considered when it comes to air pollution by making optimal administrative decisions on various levels of system of control in order to lower concentration level of air pollutants based on active system of automated control "Atmosfera" in Novomoskovsk, Tula region.

Введение

Существующая автоматизированная система контроля АСК «Атмосфера» г. Новомосковска Тульской области включает в себя: измерительную систему и информационную систему. Система не позволяет воздействовать на экологическую ситуацию, которая находится под влиянием объектов промышленности и транспорта, не позволяет дать прогноз изменения степени загрязнения атмосферного воздуха г. Новомосковска. Также в ней отсутствуют функции управления, с помощью которых можно было бы предпринять действия для управления экологической ситуацией, в частности, снижения концентрации загрязняющих веществ в атмосферном воздухе.

Для реализаций функций управления в АСК «Атмосфера» г. Новомосковска предлагается добавить блок принятия решений, который способен обеспечить:

- получение адекватной информации о возможном источнике загрязнения и доведение этой информации до всех заинтересованных сторон;
- получение краткосрочного прогноза развития экологической ситуацией в случае превышения предельно допустимой концентрации и формирование альтернативных сценариев развития ситуации (получение предполагаемого значения роста концентрации и времени роста концентрации);

выработка управляющих воздействий и реализация мероприятий, направленных на снижение концентрации загрязняющих веществ (ЗВ) и предотвращение возникновения чрезвычайной ситуации.

Таким образом, блок принятия решений должен оперативно реагировать на текущее состояние экологической ситуации в режиме реального времени. На современном этапе развития автоматизированных систем экологического мониторинга (АСЭМ) это возможно только при использовании автоматизации, компьютеризации, передовых информационных технологий на всех этапах управления экологической ситуацией: от наблюдения и контроля до принятия управленческих решений.

1 Блок принятия решений

АСК «Атмосфера» г. Новомосковска Тульской области осуществляет контроль по пяти измеряемым веществам (диоксид серы, диоксид азота, оксид азота, оксид углерода, аммиак) с интервалом 20 минут с трех стационарных постов согласно действующим нормам и стандартам [1]. В случае превышения ПДК необходимо не только не допустить непредсказуемого роста концентрации, но и желательно выявить сам источник загрязнения, а также сообщить о возможном повышении концентрации загрязняющего вещества, т.е. спрогнозировать ситуацию.

Структурная схема блока принятия решения представлена на рисунке 1, в результате работы которого формируются управляющие воздействия для снижения концентрации загрязняющих веществ (ЗВ).

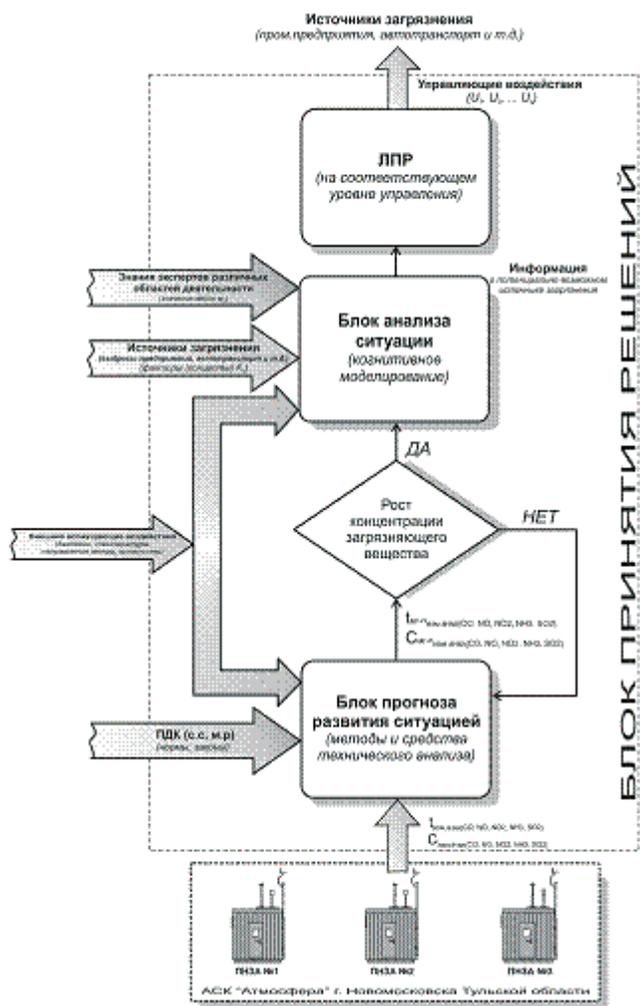


Рисунок 1 – Структурная схема блока принятий решений

Блок принятия решения состоит из следующих элементов:

- блок прогноза развития ситуации, на уровне которого формируется прогноз степени загрязнения атмосферного воздуха в случае превышения ПДК, используя методы и средства технического анализа [2];
- блок анализа, в котором на основе методологии когнитивного моделирования [3] и с учетом полученного прогноза определяется потенциально возможный источник загрязнения атмосферы;
- блок ЛПР, где осуществляется формирование возможных управляющих воздействий на соответствующем уровне управления принятия решений с целью воздействия на потенциально возможный источник загрязнения.

1.1 Блок прогноза развития ситуации

На рисунке 2 представлены графики изменения концентрации СО АСК «Атмосфера» г. Новомосковска за рабочую смену с интервалом 20 мин постом №1.

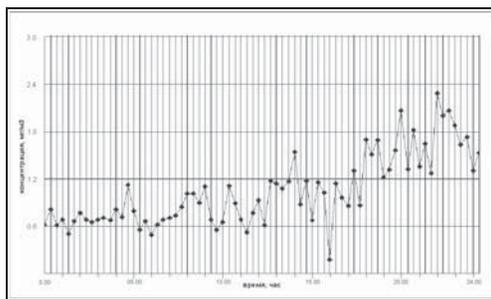


Рисунок 2 – Изменения концентрации СО АСК «Атмосфера» г. Новомосковска за рабочую смену с интервалом 20 мин постом №1

По внешнему виду графика в режиме реального времени невозможно оперативно проанализировать ситуацию и сделать прогноз об уменьшении или об увеличении концентрации ЗВ в конкретный промежуток времени. На основании проведенного анализа был сделан вывод, о том, что существуют закономерности в изменении концентрации ЗВ во времени, но получить соответствующую математическую модель не представляется возможным. Следовательно, для осуществления кратковременного прогноза значения концентрации ЗВ необходимо использовать статистические, графо-аналитические и другие методы. Поэтому, предлагается анализировать полученные данные с помощью моделей технического анализа [4]. Применение этих моделей, использующихся в основном для анализа прогнозирования цен на рынке ценных бумаг, в другой области науки и техники, в частности в экологическом мониторинге, является научной новизной.

Блок прогноза развития ситуацией (рисунок 3) должен обеспечивать динамическое прогнозирование развития текущей экологической ситуацией. Работа включает в себя следующие этапы:

Построение функциональной зависимости. По данным, полученным с АСК «Атмосфера» строятся графики зависимости концентрации измеряемого вещества от времени: $C_{CO-ПНЗА\ 1...3}(t_{20мин})$, $C_{NO-ПНЗА\ 1...3}(t_{20мин})$, $C_{NO2-ПНЗА\ 1...3}(t_{20мин})$, $C_{NH3-ПНЗА\ 1...3}(t_{20мин})$, $C_{SO2-ПНЗА\ 1...3}(t_{20мин})$.

Определение фигуры разворота. В различные моменты времени на графике может образовываться фигура разворота, которая указывает на возможный рост концентрации ЗВ. Среди фигур разворота технического анализа была выбрана модель фигуры «перевернутая голова и плечи» (ПП), как легко реализуемую на практике и не требующую больших вычислительных мощностей. Для образования фигуры разворота необходимо наличие трех локальных минимумов – левый и правый минимум, из которых, образуют «плечи», а средний – «голову», при

этом все три волны своими нижними значениями упираются в «линию шеи» (последний пик перед формированием фигуры), и точки прорыва. После точки прорыва графическая зависимость резко изменяет свое направление и происходит рост значения концентрации ЗВ;

Определение потенциала фигуры разворота. На основании образовавшейся фигуры разворота вычисляется прогноз ожидаемого роста концентрации ЗВ в режиме реального времени. Так как газоанализаторы стационарных постов наблюдения производят замеры каждые 20 мин, то в соответствии с классическим методом ППП нельзя гарантировать, что точка прорыва совпадет по времени с измеренным значением. Для увеличения точности прогноза предложен модифицированный метод «перевернутая голова и плечи» (МППП), который по результатам исследования увеличивает точность прогноза на 1,2-2% по сравнению с классическим методом. Определяются прогнозные значения: время, через которое произойдет возможный рост и соответственно само значение концентрации ЗВ: $t_{\text{МГ-П}}^{\text{ИЗМ.В-ва}}(\text{CO, NO, NO}_2, \text{HN}_3, \text{SO}_2)$, $C_{\text{МГ-П}}^{\text{ИЗМ.В-ва}}(\text{CO, NO, NO}_2, \text{HN}_3, \text{SO}_2)$.

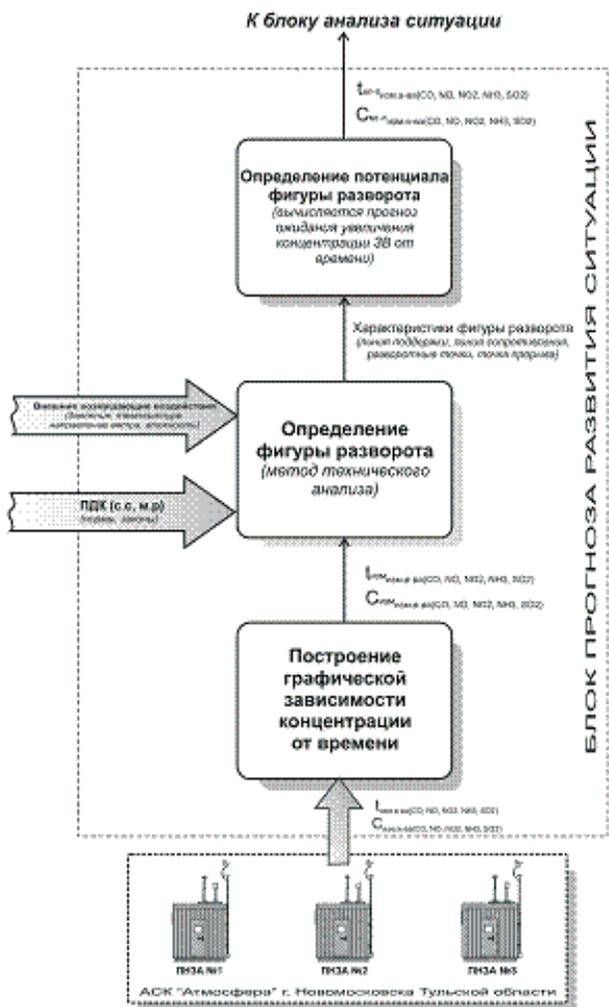


Рисунок 3 – Структурная схема блока прогноза развития ситуации

По графику (рисунок 4) определяются прогнозные значения степени загрязнения атмосферного воздуха г. Новомосковска с помощью метода технического анализа ПГП и МПГП для концентрации СО АСК «Атмосфера», измеренной постом №1.

Для определения прогноза на графиках определяются следующие точки: $X_{изм}$, $t_{изм}$, $X_{2-п}$, $t_{2-п}$, $X_{м2-п}$, $t_{м2-п}$.

где $X_{изм}$, $t_{изм}$ – соответственно измеренное значение концентрации АСК «Атмосфера», кг/м³ и текущее время измерения, час;

$X_{2-п}$, $t_{2-п}$ – соответственно рассчитанное значение концентрации по методу технического анализа «перевернутая голова и плечи», кг/м³ и полученное время прогноза, час;

$X_{м2-п}$, $t_{м2-п}$ – соответственно рассчитанное значение концентрации модифицированного метода «перевернутая голова и плечи», кг/м³ и полученное время прогноза, час.

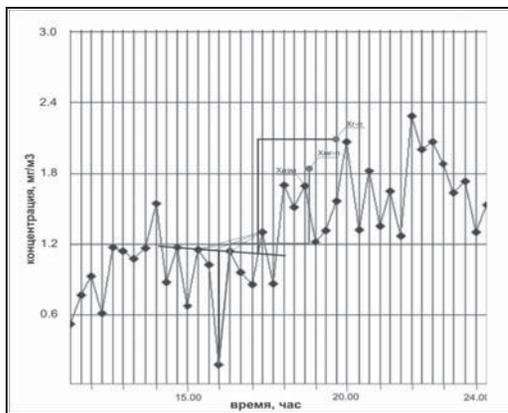


Рисунок 4 – Прогноз с помощью метода технического анализа «перевернутая голова и плечи» и модифицированного метода «перевернутая голова и плечи» для концентрации СО АСК «Атмосфера», пост №1

Для определения точки прогноза $X_{изм}$ приняты следующие допущения:

- Так как для получения точки прогноза по классическому методу ПГП и по МПГП используется один из минимумов пика фигуры, который отмечает конец среднего минимума («голова») и начало правого минимума («правое плечо»), то для определения точки прогноза $X_{изм}$ от второго минимума откладываем по оси времени расстояние, эквивалентное времени формирования фигуры $\pm 5\%$.
- В полученном временном интервале определяется оптимальная точка прогноза $X_{изм}$.
- Абсолютные погрешности по концентрации и по времени измерению по методам технического анализа вычисляются по формулам (1-3):

$$(1) \Delta C_{метод(2-п/м2-п)} = \left| X_{изм} - X_{метод(2-п/м2-п)} \right|,$$

$$(2) \Delta t_{метод(2-п/м2-п)} = \left| t_{изм} - t_{метод(2-п/м2-п)} \right|,$$

$$(3) \Delta_{метод(2-п/м2-п)} = \sqrt{\Delta C_{метод(2-п/м2-п)}^2 + \Delta t_{метод(2-п/м2-п)}^2}$$

1.2 Блок анализа ситуации

Полученные прогнозные значения $l_{м2-п,изм.6-ва}(CO,NO,NO2,HN3,SO2)$, $C_{м2-п,изм.6-ва}(CO,NO,NO2,HN3,SO2)$ сравниваются с ПДК соответствующего ЗВ, и если $C_{мет,изм.6-ва}(CO2,NO,NO2,HN3,SO2)$ превышает $C_{ПДК(изм.6-ва)}$, то необходимо выявить потенциально возможный источник загрязнения и предпринять меры для снижения роста концентрации ЗВ.

Блок анализа ситуации, представленный на рисунке 5 состоит из следующих элементов:

Формирование когнитивной карты и определение силы влияния факторов. Когнитивная карта является моделью представления знаний экспертов о законах развития и свойствах экологической ситуации и представляет собой взвешенный ориентированный граф $G = \langle X, A \rangle$, в котором X – множество вершин, взаимно однозначно соответствующее множеству базисных факторов, A – множество дуг, отражающих непосредственные влияния факторов друг на друга. На основании экспертного опроса специалистов в различных областях деятельности (администрация города, администрация предприятий, природоохранный комитет, Роспотребнадзор, жители города), накопленной информации за прошедший период, с учетом внешних возмущающих воздействий и согласования мнений группы экспертов формируется перечень факторов когнитивной карты степени загрязнения атмосферного воздуха г.Новомосковска и определяется сила влияния одного фактора на другой.

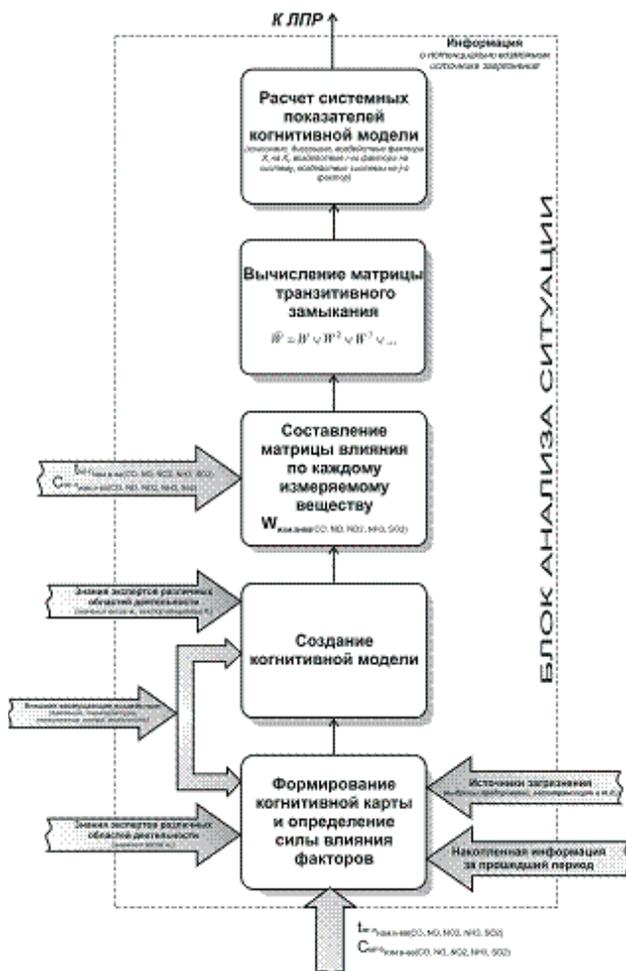


Рисунок 5 – Структурная схема блока анализа ситуации

Создание когнитивной модели. Для анализа динамики задаются шкалы факторов, позволяющие выразить их значение в числовом или лингвистическом виде: x_i – это линейно упорядоченное множество (шкала) лингвистических значений $Z_i = \{z_{i1}, z_{i2}, \dots, z_{iq(i)}\}$, где z_{i1} и $z_{iq(i)}$ –

минимальный и максимальный элементы множества; мощности шкал $q(i)$ для разных факторов различны. Для более точного анализа ситуации определяется отображение: $Z_i[0, 1]$ дискретной лингвистической шкалы $Z_i = \{z_{i1}, z_{i2}, \dots, z_{iq(i)}\}$ на отрезок $[0, 1]$. Значения весов w_{ij} определяются по данным экспертного опроса и на основе статистических методов анализа экспертной информации и показывают значения отношения причинности между каждой парой факторов в когнитивной карте по пяти измеряемым веществам АСК «Атмосфера» (рисунок 6).

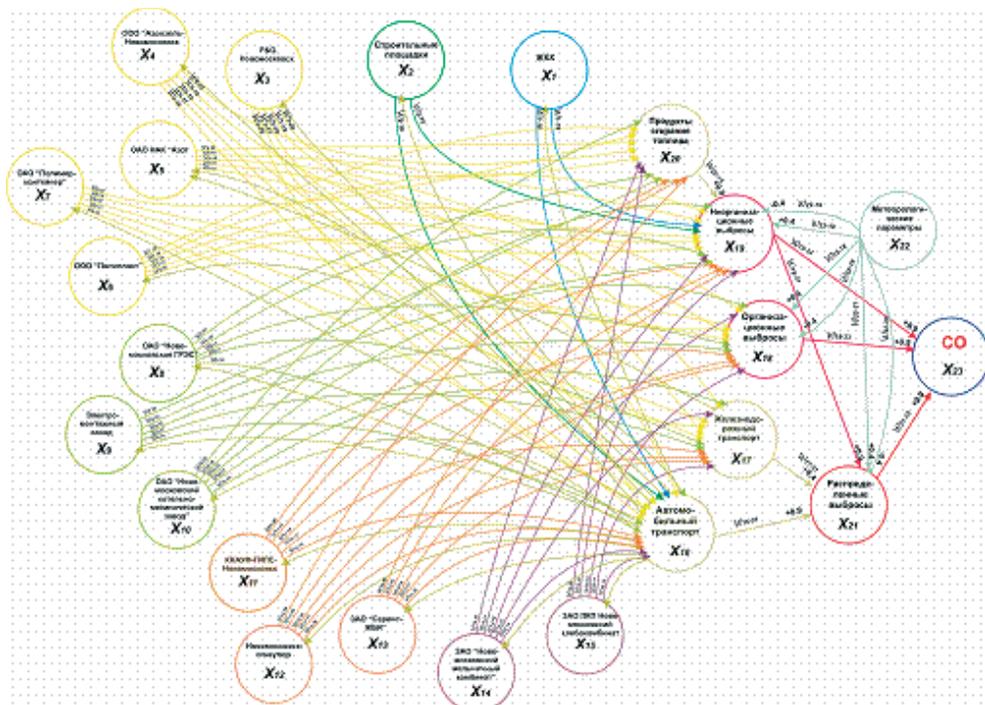


Рисунок 6 – Когнитивная модель выявления источника загрязнения на примере концентрации ЗВ CO

Составление матрицы влияния по каждому измеряемому веществу. Отношение $W_{изм.в-ва(CO,NO,NO_2,NH_3,SO_2)}$ задается в виде матрицы размерности $n \times n$ (где n – число факторов в системе), которая рассматривается как матрица смежности данного графа и называется когнитивной матрицей. Элемент w_{ij} матрицы $W_{изм.в-ва(CO,NO,NO_2,NH_3,SO_2)}$, стоящий на пересечении i -й строки и j -го столбца, характеризует влияние фактора x_i на фактор x_j .

Вычисление матрицы транзитивного замыкания. Для определения опосредованного влияния всех факторов карты друг на друга используется операция транзитивного замыкания квадратной нечеткой матрицы весов:

$$(4) \quad \widehat{W} = W \vee W^2 \vee W^3 \vee \dots$$

где степени нечетких матриц вычисляются на основе операции макстриангулярной композиции:

$$(5) \quad W^k = W^{k-1} \circ W$$

Расчет системных показателей когнитивной модели. По полученной транзитивно-замкнутой когнитивной матрице взаимовлияний для конкретного ЗВ определяются системные и интегральные показатели влияния факторов на систему [5].

В результате работы блока анализа ситуации выявляется потенциально возможный источник загрязнения. Данная информация передается ЛПР, которое принимает решение на соответствующем уровне управления для снижения концентрации ЗВ в режиме реального времени. Оптимальное решение предлагается выбирать на основе критерия Эджворта – Парето [6], который представляет в данном случае задачу многокритериального выбора $\langle X, f, \succ_x \rangle$, где:

- X – множество возможных вариантов (решений), из которого следует осуществлять выбор оптимальных управляющих воздействия (УВ) для снижения концентрации ЗВ;
- $f = (f_1, \dots, f_m), m \geq 2$ – векторный критерий, определенный на множестве X . Критериями оптимальности являются экологическая составляющая, обеспечивающая минимум негативного воздействия f_1 и экономическая составляющая f_2 , которая обеспечивает увеличение прибыли и как следствие максимум воздействия на атмосферу;
- \succ_x – асимметричное бинарное отношение строгого предпочтения ЛПР, определенное на множестве X ; запись $x_1 \succ_x x_2$; $x_1, x_2 \in X$ означает, что вариант x_1 для ЛПР предпочтительнее варианта x_2 .

Таким образом, при принятии решения ЛПР выбирает оптимальное решение из множества допустимых УВ с учетом двух векторных критериев, при чем обязательно должны учитываться оба критерия одновременно.

Заключение

Результатом работы вышеописанных блоков является автоматизированное определение возможного источника загрязнения атмосферного воздуха, реальные значения концентрации ЗВ, регистрируемые АСК «Атмосфера», прогнозируемые значения превышения ПДК ЗВ и время, за которое это превышение может быть достигнуто. Такая схема с одной стороны позволяет предприятию заблаговременно принять действия по снижению концентрации ЗВ и не допустить возникновения чрезвычайной ситуации, а с другой стороны – контролирурующие органы превентивно будут обеспечивать требуемый уровень безопасности проживания людей на контролируемой территории.

Список литературы

- [1] ГН 2.1.6.1338-03. Гигиенические нормативы: Предельно-допустимые концентрации загрязняющих веществ в атмосферном воздухе населенных мест - Утверждено 30.05.03.
- [2] Батышкина В.В., Волков В.Ю., Чистозвонова Л.А. Прогнозирование развития экологической ситуации с помощью моделей технического анализа// Информатика, Экология, Экономика. Вестник международной академии системных исследований (МАСИ). М.: 2010. – С.20–25.
- [3] Волков В.Ю., Батышкина В.В. Проблемы применения когнитивного подхода к созданию интеллектуальной системы экологического мониторинга и управления// Известия высших учебных заведений: Химия и химическая технология. – 2009. – т.52, №6. – С. 109-113.
- [4] Батышкина В.В., Волков В.Ю., Чистозвонова Л.А. Выявление тенденции изменения концентрации загрязняющих веществ в атмосфере с помощью моделей технического анализа// Сб.научн.трудов «Успехи в химии и химической технологии». — 2009. — Том XXIII, №1 (94). — С. 86-90.
- [5] Волков В.Ю. Интеллектуальная система оценки влияния промышленных предприятий на концентрацию загрязняющих веществ в атмосферном воздухе// Датчики и системы. – 2009. - №10. – С. 10-15.
- [6] Кини Р.Л., Райфа Х. Принятие решений при многих критериях: предпочтения и замещения. – М.: Радио и связь, 1981.

ИМИТАЦИОННЫЙ ПОДХОД К МОДЕЛИРОВАНИЮ ПОЖАРООПАСНЫХ СИТУАЦИЙ В ВУЗАХ

И.В. Дрынова, С.В. Дегтярев

Курский государственный технический университет
305004, Курск, ул. Челюскинцев, 19, Россия
ist@kstu.kursk.ru
тел: +7 (4712) 58-71-10, факс: +7 (4712) 58-71-11

***Ключевые слова:** пожар, пожарная опасность, пожароопасная ситуация, риск, имитационное моделирование, информационная система*

Abstract

The paper presents the results of issue of imitational approach to the simulation of the fire situations in the higher education buildings. The general requirements for an automated system for imitational simulation of the fire situations in the higher education buildings were formulated.

Введение

Одной из актуальной проблем современного общества являются техногенные пожары, которые наносят материальный ущерб и приводят к ранениям, а в ряде случаев и гибели людей. В связи с этим ведется разработка новых информационных систем и технологий предупреждения и реагирования в случае возникновения пожароопасных ситуаций. В основе функционирования таких систем лежат модели возникновения пожароопасных ситуаций и развития пожаров.

1 Обзор подходов к оценке пожарной опасности ВУЗов

На сегодняшний день в России реализуются два основных подхода при оценке пожарной опасности, которые могут быть применены к ВУЗам – детерминированный и вероятностный. При детерминированном подходе к оценке опасности объектов вероятность образования горящей среды и появления источника зажигания условно принимается равной единице. При этом в качестве расчетного выбирается наиболее неблагоприятный вариант аварии или период нормальной работы технологического оборудования (аппарата), при котором в пожаре (взрыве) участвует наибольшее количество веществ и материалов, наиболее опасных в отношении последствий пожара (взрыва), т. е. так называемый принцип «максимальной проектной аварии». К достоинствам детерминированного подхода можно отнести относительную простоту его использования, высокую степень завершенности и однозначности решения задач оценки пожароопасности объектов. Недостатками этого подхода являются:

- жесткость в определении расчетного варианта аварии, даже если вероятность максимальной проектной аварии мала. При этом применение защитных мероприятий, в том числе и капитального характера, может быть избыточным и экономически нецелесообразным;
- использование этого подхода не стимулирует предприятие заботиться о повышении безопасности путем применения более надежного оборудования.

Результаты оценки, проведенной с применением детерминированного подхода, дают определенный запас надежности, однако величина этого запаса может приводить к нерациональным материальным затратам на обеспечение пожаробезопасности объекта.

Вероятностный подход основан на более рациональном сопоставлении величин опасных факторов пожара (взрыва), уровня обеспечения безопасности людей, ожидаемого материаль-

ного ущерба и затрат на обеспечение пожаровзрывобезопасности. Результаты оценки, проведенной с применением вероятностного подхода, характеризуются величинами риска.

При исследовании риска обычно выделяют обычно три аспекта проблемы [1-2]:

- 1) анализ риска;
- 2) оценка риска;
- 3) управление риском.

Под анализом риска понимается выявление нежелательных событий, влекущих за собой реализацию опасности (например, пожара), анализ механизмов возникновения подобных событий, выявление и характеристика возможных негативных последствий реализации опасности.

Оценка риска предусматривает процедуру количественного определения его величины.

Управление риском – это совокупность мероприятий, направленных на предупреждение и устранение причин аварий (пожаров) или снижение их последствий, т.е. практическая деятельность, направленная на снижение риска. Сюда же включаются нормативно-правовая деятельность и государственное регулирование пожарной безопасности.

Все известные методы, используемые для оценки и анализа риска возникновения пожароопасных ситуаций, могут быть сгруппированы в три общих группы [2].

Первая группа включает методы, основанные на индексации опасности возникновения аварии (пожара) и ожидаемого ущерба от нее. В большинстве случаев эти методы базируются на обобщении опыта специалистов, занимающихся проблемами обеспечения пожаровзрывобезопасности. Возможности использования индексных методов для решения практических задач весьма ограничены из-за абстрактности показателей (индексы, баллы), не имеющих физической интерпретации.

Вторая группа включает методы, которые предполагают использование статистических данных об авариях, причинах их возникновения и ущербе. При наличии объективных статистических данных указанные методы позволяют получать достоверные результаты.

В настоящее время наибольшее распространение получила третья группа методов – расчетно-аналитические. В научной литературе описаны два различных подхода к оценке риска с помощью расчетно-аналитических методов [3]:

- классический, основанный на рассмотрении деревьев событий, приводящих к реализации того или иного опасного фактора;
- основанный на имитационном моделировании аварий с расчетом их поражающих факторов (разновидность известного метода Монте-Карло). Следует отметить, что методы имитационного моделирования используются также при оценке частот возникновения иницирующих аварию событий.

2 Преимущества применения имитационного подхода к моделированию пожароопасных ситуаций в ВУЗах

Пожар, как объект моделирования характеризуется:

- недостаточной проработанностью некоторых теоретических аспектов его возникновения и распространения;
- большой долей экспертных знаний при описании, структуризации объекта моделирования;
- слабоструктурированностью задач моделирования пожара;
- высоким уровнем неопределенности исходной информации. Различают внутреннюю и внешнюю неопределенность. Внутренняя неопределенность – это совокупность тех факторов, которые не контролируются лицом, принимающим решение полностью, но он может оказывать на них влияние. Внешняя неопределенность определяется характером

взаимодействия с внешней средой – это те факторы, которые находятся под слабым контролем лица принимающего решение;

- пожар представляет собой сложную динамическую систему.

Поэтому в качестве метода моделирования развития пожароопасной ситуации целесообразно выбрать метод имитационного моделирования, поскольку он позволяет адекватно отразить структуру рассматриваемой сложной динамической системы, привести в модель факторы неопределенности. Метод имитационного моделирования обеспечивает итеративный процесс разработки модели, характеризующийся постепенным углублением знаний о системе с участием экспертов и специалистов предметной области.

Имитационное моделирование – один из видов компьютерного моделирования, использующий методологию системного анализа, центральной процедурой которого является построение обобщенной модели, отражающей все факторы реальной системы, в качестве же методологии исследования выступает вычислительный эксперимент.

Имитационное моделирование пожароопасных ситуаций имеет ряд преимуществ по сравнению с другими подходами. В частности, оно дает возможность учитывать большое количество переменных, предсказывать развитие нелинейных процессов, возникновение синергетических эффектов. Имитационное моделирование позволяет не только получить прогноз развития пожара, но и определить, какие управляющие воздействия приведут к наиболее благоприятному развитию событий.

Качественные выводы, сделанные по результатам имитационного моделирования пожароопасной ситуации, позволяют обнаружить такие свойства сложной системы, как ее структуру, динамику развития, устойчивость, целостность и др. Количественные выводы в основном носят характер прогноза рисков пожароопасных ситуаций, характеризующих систему. Одно из основных направлений использования компьютерного моделирования пожара – поиск оптимальных вариантов внешнего воздействия на объект с целью получения минимальных показателей рисков.

Имитационная модель развития пожароопасной ситуации строится строго целенаправленно, поэтому для нее характерно адекватное отображение исследуемого объекта, логико-математическая модель системы представляет собой программно реализованный алгоритм развития пожара. При имитационном моделировании структура моделируемой пожароопасной ситуации адекватно отображается в модели, а процесс ее развития имитируется на построенной модели. Под имитацией понимают проведение на компьютерах различных серий экспериментов с моделями, которые представлены в качестве некоторого набора (комплекса) компьютерных программ. Сравнение характеристик (конструкций, управлений) моделируемого объекта осуществляется путем вариантных просчетов. Особую роль имеет возможность многократного воспроизведения моделируемых процессов пожара с последующей их статистической обработкой, позволяющая учитывать случайные внешние воздействия на изучаемый процесс. На основе набираемой в ходе компьютерных экспериментов статистики делаются выводы в пользу того или иного варианта развития пожароопасной ситуации.

Однако во многих практических случаях моделирования развития пожара формировать решения с помощью формальных методов не удастся – эксперт должен быть включен в процесс принятия решения. Он становится активным компонентом информационной системы – детализирует проблему и модель, осуществляет постановку направленного вычислительного эксперимента на модели, генерацию и ранжирование альтернатив, выбор критериев для принятия решений, а также формирует рациональный вариант управления с помощью базы знаний. Принятие решений в условиях риска, например, требует ведения диалоговых процедур формирования статистически достоверных результатов и поэтапного сопоставления их с функцией цены риска. В таких случаях необходимо осуществлять прямое участие эксперта в формировании оптимального множества вариантов решений и в процедурах вариантного синтеза.

Таким образом, имитационное моделирование развития пожароопасных ситуаций значительно расширяет возможности и эффективность работы лиц, принимающих решения (ЛПР), предоставляя им удобный инструмент и средства для достижения поставленных целей. Имитационное моделирование развития пожароопасных ситуаций реализует итерационный характер разработки модели пожара, поэтапный характер детализации моделируемых компонент, что позволяет постепенно увеличивать полноту оценки принимаемых решений по мере выявления новых проблем и получения новой информации.

Имитационная модель развития пожароопасной ситуации не дает оптимального решения подобно классическому решению задач оптимизации, но она является удобным вспомогательным средством для поиска решения проблемы снижения пожарных рисков. Область применения имитационных моделей развития пожароопасных ситуаций практически не ограничена, это может быть сфера образования, производство и т.д. Главным преимуществом имитационного моделирования пожара является то, что эксперт может ответить на вопрос: «Что будет, если ...», т.е. с помощью эксперимента на модели вырабатывать стратегию развития.

В последнее время ведутся работы по разработке систем, способных оказать помощь эксперту при ответе на обратный вопрос «Что надо, чтобы ...». Это можно назвать как «целевое моделирование», при котором на вход системы подаются показатели целевого состояния, а также перечень возможных регуляторов с указанием диапазона и шага их изменения. Система в автоматическом или полуавтоматическом режиме находит сочетание значений этих регуляторов для достижения заданного целевого состояния.

3 Особенности информационных систем имитационного моделирования пожароопасных ситуаций в ВУЗах

Имитационное моделирование пожароопасных ситуаций проводится в интересах планирования мероприятий противопожарного обеспечения, опирается на систему оценок рисков возникновения и развития пожаров и используется в качестве базиса для построения систем информационного обеспечения.

Задача информационной системы имитационного моделирования пожароопасных ситуаций заключается в том, чтобы помочь ЛПР, и указать, при каком значении контролируемого параметра какие мероприятия целесообразно проводить, чтобы не выполнять лишних мероприятий и в то же время уменьшить возможный ущерб.

Процесс принятия решений должен включать в себя:

- сбор и комплексирование информации об объекте и среде управления;
- обработку и хранение этой информации с приемлемыми степенями агрегации;
- автоматическую или человеко-машинную оценку состояний объекта и среды управления с прогнозированием возможной смены состояний;
- определение целей и критериев эффективности (безопасность людей, имущества и др.);
- автоматический и/или человеко-машинный поиск решений по управлению, инициируемый оценками состояний объекта и среды управления или неблагоприятными прогнозами тенденций развития контролируемых процессов;
- автоматическую и/или человеко-машинную оптимизацию найденных и рекомендуемых человеку, коллективу людей решений;
- человеко-машинное принятие решений с возможностями вызова для анализа как данных, лежащих в основе поиска рекомендуемых решений, так и используемой логики и математических методов, на которых базировался поиск предлагаемых решений, согласование и утверждение мероприятий; реализация решений – приведение объекта в целевое состояние;
- оценка результатов (корректировка целей, способов и путей достижения целей).

Применительно к учреждениям высшего профессионального образования задачу прогнозирования и предупреждения возникновения пожаров можно представить в виде задачи ситуационного управления динамической системой в условиях неопределенности. Ее решение построено на основе имитационного моделирования процессов возникновения и развития пожаров. Модели развития пожароопасных ситуаций включаются в контур управления пожарной безопасностью для имитации реальных процессов, определения возможных ситуаций и вероятности их возникновения на основе сценарного подхода.

В рамках данного подхода могут быть сформированы следующие требования к информационным системам имитационного моделирования пожароопасных ситуаций в вузах:

- построение имитационных моделей с одновременным использованием математических закономерностей и знаний экспертов предметной области;
- обеспечение возможности обработки данных для расчетов на моделях и прогнозирования в реальном масштабе времени;
- поддержка визуализации разрабатываемых моделей;
- открытость и возможность дополнения и уточнения созданных имитационных моделей;
- обеспечение возможности создания прикладных приложений на основе сформированной имитационной модели пожароопасной ситуации;
- обеспечение возможности использования созданных прикладных приложений совместно с существующими системами пожарной безопасности вузов.

С учетом необходимости использования в ходе имитационного моделирования как формализованных, так и слабоформализованных данных и знаний предметной области в качестве основных задач проектирования информационной системы противопожарного обеспечения вуза должны быть выделены:

- 1) разработка функциональной модели информационной системы противопожарного обеспечения;
- 2) обоснование подхода к количественной оценке рисков возникновения и развития пожаров;
- 3) выбор метода представления и формализация экспертных знаний предметной области;
- 4) разработка способа формирования и представления альтернативных вариантов возникновения и развития пожаров.

Заключение

Современные тенденции в области имитационного моделирования связаны с развитием проблемно-ориентированных систем, созданием встроенных средств для интеграции моделей в единый модельный комплекс; технологический уровень современных систем моделирования характеризуется большим выбором базовых концепций формализации и структуризации моделируемых систем, развитыми графическими интерфейсами и анимационным выводом результатов. Имитационные системы имеют средства для передачи информации из баз данных и других систем, или имеют доступ к процедурным языкам, что позволяет легко выполнять вычисления, связанные с планированием факторных экспериментов, автоматизированной оптимизацией и др.

Список литературы

- [1] Азанов С.Н. Ещё раз о риске [Текст] / С.Н.Азанов, С.Н.Вангородский, Ю.Ю.Корейчук // ВИНТИ. Проблемы безопасности при чрезвычайных ситуациях, 1999. Вып. 7. С. 32-51.
- [2] Алымов В.Т. Техногенный риск. Анализ и оценка [Текст] / В.Т.Алымов, Н.П.Тарасова. М.: Академкнига, 2004. 118 с.
- [3] Белов П.Г. Системный анализ и моделирование опасных процессов в техносфере [Текст] / М.: Академия, 2003. 512 с.

ВЫЯВЛЕНИЕ ОТЛИЧИТЕЛЬНЫХ ПРИЗНАКОВ МОДЕЛЕЙ ОТРАСЛЕЙ ЖИЛИЩНО-КОММУНАЛЬНОГО ХОЗЯЙСТВА И ПРОГНОЗИРОВАНИЕ ИХ РАЗВИТИЯ

Д.В. Абарин

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
cscmp@iccs.ru
тел: +7 (846) 332-39-27, факс: +7 (846) 333-27-70

Ключевые слова: реформирование, концепция развития, мировые модели ЖКХ

Abstract

The article presents the detection of the distinctive signs of models of the housing and communal services sector and the prediction of their development.

Введение

Сейчас для большинства процесс реформирования жилищно-коммунального хозяйства в России представляется как хаотичное развитие. К данному большинству относятся инвесторы, рассматривающие данную отрасль как доходную, но в тоже время высокорискованную сферу экономических интересов, которая из-за этого имеет только краткосрочные и среднесрочные прогнозируемые перспективы. Действительно, сложно инвестировать в государстве, деятельность которого в нормативно-правовых актах и официальных договорах приравнивается к форс-мажору. Деловые интересы развития бизнеса в ЖКХ требуют разработки эффективных инструментов прогнозирования развития методов управления объектами коммунальной инфраструктуры, и, как следствие, построения партнерских отношений нового типа между публичной властью (собственником коммунальных систем) и бизнесом на долгосрочной основе.

В настоящее время ввиду несовершенства законодательства частный бизнес в коммунальной отрасли подвержен максимальному риску. По этой причине наилучшей способностью выжить могут обладать только крупные компании, способные нивелировать убыточные проекты за счет рентабельных проектов, а так же собственных оборотных средств. Наиболее простой и эффективной оценкой развития внешних факторов влияющих на реализацию любого инвестиционного проекта в жилищно-коммунальной отрасли будет являться поиск аналогичных условий, и как следствие, использование готовых данных о ее развитии с реальными, а не прогнозными данными, характеризующие инвестиционный климат. В настоящее время Россия развивается самостоятельным путем, прогнозировать который можно исключительно с философской точки зрения. Широким диапазоном моделей развития располагают 83 субъекта Российской Федерации, исходя их политической конъюнктуры, экономической и технической ситуации в регионе в целом и в отрасли в частности. Вместе с тем, совокупность признаков развития регионов, общая тенденция развития большинства регионов-доноров (что важно как показатель самостоятельности процесса) в совокупности с действующим законодательством Российской Федерации формируют вектор развития жилищно-коммунальной отрасли ЖКХ. Попытки произвести анализ направления данного вектора производились Правительством РФ с 1999 года.

В настоящее время Концепцией федеральной целевой программы «Комплексная программа модернизации и реформирования жилищно-коммунального хозяйства на 2010-2020 годы» Правительством РФ осуществлена очередная попытка формирования перспективной

модели развития жилищно-коммунальной отрасли. Авторы концепции планируют два варианта развития отрасли, а именно:

- разработка программы комплексного развития и единой муниципальной базы информационных ресурсов за счет местных бюджетов;
- создание единой муниципальной базы информационных ресурсов и подготовку конкурсов на заключение концессионных соглашений с последующим предоставлением кредитов на реализацию указанных мероприятий с включением понесенных затрат в стоимость пакета правоустанавливающих документов, выставляемых на конкурс инвестиционных проектов (либо на конкурс на заключение концессионного соглашения), либо софинансирование указанных расходов за счет ассигнований федерального бюджета (эта модель, очевидно, требует разработки типовых механизмов банковского кредитования).

На первом этапе программы согласно концепции (2010-2011 годы) предусматривается разработка нормативной базы для реализации Программы и запуск нескольких пилотных проектов. Таким образом, авторы концепции признают неопределенность перспективы развития отрасли ЖКХ, изначально предлагая два пути развития с последующей их коррекцией по результатам пилотных проектов и получения практических данных. Таким образом, Правительством РФ очередной раз пролонгирован срок завершения реформы ЖКХ и определения устойчивой ее модели.

В то же время в мире в основном используется многовековой опыт в этой сфере, и разработано несколько классических моделей ЖКХ, с установленными правилами привлечения инвестиций, оценкой их риска, условиями государственной защиты инвестора и т.д. Новые модели развития ЖКХ, появляющиеся в мире, представляют собой модели переходного периода, и рано или поздно перейдут в классические вариации. Поэтому наиболее удобным и наименее затратным способом прогнозирования развития отрасли в связи с фатальной неопределенностью является выявление признаков формирования стабильной классической модели, определение сроков ее формирования исходя из количества совпадений отдельных ее признаков с признаками анализируемой модели.

1 Основные модели отраслей жилищно-коммунального хозяйства

Анализ мирового опыта функционирования отраслей ЖКХ позволяет выделить модели Германии, Финляндии, Сингапура и США как наиболее успешно зарекомендовавшие себя в отношении бизнеса. Отличительная особенность этих моделей заключается в форме собственности объектов жилищно-коммунального хозяйства и роли государства в вопросе их управления. Необходимо отметить, что в Германии фактически имеются две модели – Восточной и Западной ее территории, что обусловлено различными условиями их становления. В таблице 1 представлены основные отличительные признаки указанных мировых моделей жилищно-коммунального хозяйства.

Любому возможному развитию жилищно-коммунальной отрасли в России можно найти аналог за рубежом, который завершил свое развитие одной из пяти основными моделями. Любое текущее временное развитие отрасли является промежуточным из числа шести обозначенных моделей. Динамика смещения к одной из рассматриваемых моделей есть признак её формирования. Число совпадений признаков модели говорит о степени рисков неопределенности – форс-мажора.

Таблица 1 – Основные мировые модели жилищно-коммунального хозяйства

<i>Государство</i>	<i>Отличительные признаки</i>	<i>Примечание</i>
Германия (западная)	Собственник объектов коммунальной инфраструктуры - частник. Управление объектами комму-	Близкая форма управления к концепции развития ЖКХ России.

Продолжение таблицы 1

	<p>нальной инфраструктуры – частное (контролируемое государством).</p> <p>Основная роль государства: контроль управляющих и муниципальных компаний.</p>	
Германия (восточная)	<p>Собственник объектов коммунальной инфраструктуры - смешанный.</p> <p>Управление объектами коммунальной инфраструктуры – смешанное.</p> <p>Основная роль государства: контроль управляющих и муниципальных компаний, целевые субсидии на решение вопросов ЖКХ.</p>	<p>Наиболее схожая исходная ситуация с состоянием жилищно-коммунального хозяйства России. Совершенствование ценообразования и модернизация инфраструктуры: первым шагом явился переход к адресному субсидированию. Субсидии предоставлялись малообеспеченным семьям в зависимости от получаемых ими доходов. Освободившиеся бюджетные средства были направлены в качестве инвестиций в модернизацию сетей и оборудования. Средства выделялись путем предоставления грантов и кредитов под низкие проценты. Необходимо отметить значительный рост стоимости услуг и одновременно переход к расчету за предоставленные услуги согласно их фактическому потреблению. Это привело к экономии ресурсов населением и их индивидуальному потреблению.</p>
Сингапур	<p>Собственник объектов коммунальной инфраструктуры - государство.</p> <p>Управление объектами коммунальной инфраструктуры – частное (контролируемое государством).</p> <p>Основная роль государства: контроль накопления средств жителей для решения жилищных проблем.</p>	<p>Основу ЖКХ составляет Центральный фонд сбережений (ЦФС). Каждый сингапурец обязан ежемесячно отчислять туда 20 процентов своего заработка. Столько же каждый месяц переводит на его счет наниматель. Вместе эти взносы составляют значительную сумму, равную 40 процентам фонда заработной платы.</p> <p>Сбережения в ЦФС не облагаются налогом, в инвестиционном банке на них ежегодно начисляются проценты. Достигнув 55-летнего возраста, сингапурец получает всю эту сумму на руки. Но еще до ухода на пенсию он вправе использовать три четверти сбережений в ЦФС, дабы приобрести квартиру, а четверть - чтобы при необходимости оплатить лечение в больнице.</p> <p>Льготами государственной программы жилищного строительства вправе пользоваться только граждане Сингапура, причем люди с доходами не выше определенного уровня, не имеющие недвижимости. Чтобы получить квартиру, достаточно накопить в ЦФС 20 процентов ее стоимости, а остальное внести в рассрочку. С целью ускорить покупку квартиры родители и дети, братья и сестры могут объединять средства, которые они накопили в ЦФС. Если через пять лет после въезда построенная государством квартира полностью оплачена, ее можно продать по рыночной цене без всяких ограничений.</p> <p>Имея в штате всего 12 тысяч человек, Управление жилищно-коммунального хозяйства справляется с огромным объемом работ по строительству и</p>

		эксплуатации потому, что распределяет на открытых торгах подряды на все виды коммунального обслуживания, включая водоснабжение, канализацию, вывоз мусора, стрижку газонов, содержание автостоянок. Оно сохраняет в своих руках лишь эксплуатацию 12 тысяч лифтов, чтобы обеспечить их бесперебойную работу в многоэтажных домах и не допустить несчастных случаев.
США	Собственник объектов коммунальной инфраструктуры - частник. Управление объектами коммунальной инфраструктуры – частное (контролируемое жителями). Основная роль государства: компенсация процентов по кредитам для решения жилищных проблем.	По существующему законодательству частные компании используются для обслуживания коммунального хозяйства и строительства только в исключительных случаях (временно или тогда, когда этот сервис не может быть обеспечен государственной компанией).
Финляндия	Собственник объектов коммунальной инфраструктуры: - частник. Управление объектами коммунальной инфраструктуры – частное (контролируемое жителями). Основная роль государства: компенсация процентов по кредитам для решения жилищных проблем.	В Финляндии накоплен богатый опыт функционирования акционерных обществ созданных для совместного управления жилыми домами. Вся жилая недвижимость управляется собственниками. Председателем правления акционерного общества не может быть собственно житель этого дома. Председатель - это наемный менеджер, работа которого жестко контролируется. В Финляндии всю жилищно-коммунальную систему регулирует бизнес. Государство фактически отстранилось от помощи собственникам жилья. Единственное, в чем проявляется участие государства, - это погашение процентов по кредиту.

2 Отличительные признаки классических моделей отраслей жилищно-коммунального хозяйства

В таблице 1 мировых моделей жилищно-коммунального хозяйства, представлены следующие основные отличительные признаки и их значения:

- 1) собственник объектов коммунальной инфраструктуры (1 – государство; 2 – смешанный; 3 - частник);
- 2) форма управления объектами коммунальной инфраструктуры (1 – государственное; 2 – смешанное; 3 - частное);
- 3) роль государства в вопросе реализации жилищно-коммунальной политики (контроль управляющих и муниципальных компаний: 1 – общий контроль; 2 - помощь управляющим компаниям; 3 - помощь частным лицам).

Для первого признака значение 2 – «смешанная форма собственности» соответствует 50% частной форме собственности и 50% - государственной.

При принятых обозначениях таблица 1 трансформируется в матрицу инцидентности, представленную таблицей 2.

На рисунке 1 представлена схема формирования конкретной модели жилищно-коммунальной отрасли. Здесь M_i – типовая модель, $i \in \{1, 2, 3, 4, 5\}$, M_n -вариант модели, представленной к анализу возможных инвестиционных рисков.

Если считать M_i вектором в пространстве признаков, то минимизация риска достигается при $\|M_n\|/\|M_i\| \rightarrow 1$.

В целом величина риска для модели жилищно-коммунального хозяйства может быть оценена суммой $\sum_{i=1}^5 \|M_n\|/\|M_i\|$.

Таблица 2 – Матрица «модели-признаки» для жилищно-коммунальной отрасли

Государство	Основные отличительные признаки моделей ЖКХ		
	Собственник объектов коммунальной инфраструктуры	Управление объектами коммунальной инфраструктуры	Основная роль государства
Германия (западная)	3	3	1
Германия (восточная)	2	2	3
Сингапур	1	2	3
США	2	1	1
Финляндия	3	3	3

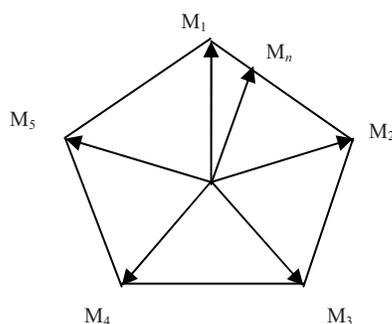


Рисунок 1 – Графическая иллюстрация к выбору модели ЖКХ

Выводы

Развитие любой модели отрасли жилищно-коммунального хозяйства в конечном итоге приводит к так называемым «идеальным» моделям. Конструируя облик модели как по отдельным значениям, так и по их совокупности, можно выявить признаки, формирующие вектор развития отрасли. Рассмотренный подход к прогнозированию не исчерпывает свои возможности отдельной отраслью. Данный метод можно успешно применить и на уровне муниципальных образований, и на уровне субъектов РФ. Наибольшие перспективы метода просматриваются при анализе совокупности субъектов РФ для оценки формирующейся модели жилищно-коммунального хозяйства по России в целом.

Список литературы

- [1] Распоряжение Правительства РФ от 02.02.2010 №102-р «Об утверждении Концепции федеральной целевой программы "Комплексная программа модернизации и реформирования жилищно-коммунального хозяйства на 2010 - 2020 годы».
- [2] Письмо Минфина РФ от 09.01.2001 №10-05-22 «Об обзоре результатов анализа эффективности мероприятий, проводимых субъектами Российской Федерации в соответствии с Концепцией реформы жилищно-коммунального хозяйства».
- [3] Распоряжение Правительства РФ от 17.11.2008 №1662-р (ред. от 08.08.2009) «О Концепции долгосрочного социально-экономического развития Российской Федерации на период до 2020 года».

КОНЦЕПЦИЯ ПРОЕКТА СОЗДАНИЯ СИСТЕМЫ КОЛЛЕКТИВНОГО УПРАВЛЕНИЯ ОБРАЗОВАТЕЛЬНЫМ УЧРЕЖДЕНИЕМ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ СИТУАЦИОННОГО ЦЕНТРА

Д.П. Маевский¹, В.А. Филимонов²

¹Омский государственный институт сервиса
644090, Омск, ул. Певцова, 13, Россия
jobdim@yandex.ru

тел: +7 (381) 223-82-26, факс: +7 (381) 224-94-45

²Омский филиал Института математики СО РАН
644090, Омск, ул. Певцова, 13, Россия
filimono@ofim.oscsbras.ru
тел: +7 (381) 223-67-39

Ключевые слова: субъекты управления, кризис образования, проекты, прогнозы, прорывные технологии, сетевая организация

Abstract

The concept of collective management by the educational institution, adequate to subjects of management is considered. The collective is understood as all people and the organizations which functioning is connected with educational institution, including the personnel, pupils, their relatives, employers, administrations of territories. The project of a network control system is offered.

Введение

Настоящий текст отражает определённую авторскую позицию и ориентирован на достаточно узкую целевую аудиторию. Авторская позиция, в общем, состоит в оценке глобальных изменений в системе образования России как негативных. Целевая аудитория (к ней относятся и сами авторы) – люди, которые имеют авторские проекты, относящиеся к системе образования и соразмерные ресурсам авторов. Тем самым из рассмотрения заведомо исключаются проекты обустройства образования России в целом. Сложившаяся ситуация существенно ограничивает степени свободы при реализации таких проектов. Однако существуют технологии, позволяющие управлять процессами из различных точек при наличии критического ресурса – понимания принципов и механизмов действия систем, в которых существуют такие проекты [1].

1 Концепция системы

Система коллективного управления образовательным учреждением понимается в данном тексте как регулярное взаимодействие персонала учреждения и представителей социума. Взаимодействие осуществляется по проектному регламенту в ситуационном центре, поддерживается специализированной сервисной командой и документируется. Результаты теоретических исследований проходят экспериментальную проверку на прототипе – учебно-исследовательском ситуационном центре.

Научная проблема состоит в отсутствии теоретических основ и практических методик создания механизмов коллективного управления образовательными учреждениями с использованием потенциала современных информационных технологий. Под коллективным управлением понимается систематическое взаимодействие персонала образовательного учреждения

и представителей социума в процессе определения и реализации стратегии и тактики образовательного учреждения.

Актуальность проекта обусловлена необходимостью интеграции образовательного учреждения (администрации и профессорско-преподавательского состава) и социума (органов местного самоуправления, работодателей, общественных организаций, родственников студентов) для обеспечения эффективного управления с учётом разнородных интересов всех прямых и косвенных участников образовательного процесса. Потребность в такой интеграции многократно возрастает в период резких социальных перемен.

Цель проекта состоит в формировании знания о способах создания и закономерностях функционирования механизма коллективного управления образовательным учреждением. Достижение цели требует постановки и решения задач, перечисленных ниже.

- Разработка теоретических основ коллективного управления образовательным учреждением.
- Разработка рекомендаций по созданию учебно-исследовательского ситуационного центра образовательного учреждения на основе доступных ресурсов.
- Разработка регламента коллективного управления образовательным учреждением с использованием технологий ситуационного центра.
- Создание действующего прототипа системы коллективного управления и его экспериментальное использование.
- Проведение экспертизы полученных результатов.
- Публикация истории исследования и аналитических материалов по его результатам.

Основной результат состоит в разработке теоретических основ создания коллективных систем управления образовательными учреждениями. Результатами являются также учебные курсы для подготовки сервисных команд ситуационных центров, и опыт практической реализации прототипа для руководителей проекта, молодых учёных и студентов.

Практическое значение состоит в повышении общего уровня понимания и, соответственно, доверия участников процесса к системе взаимодействия. Дефицит доверия является одним из главных препятствий к формированию и реализации эффективной социальной политики. Использование комплексного текстового и графического представления информации повышает уровень понимания ситуации. Мониторинг участников повышает уровень ответственности. Участие экспертов в реальном времени с использованием Интернет повышает уровень экспертизы.

Новизна поставленной задачи состоит в формировании нового способа взаимодействия участников управления образовательным учреждением: целенаправленного, проектно-ориентированного, поддержанного возможностями информационных технологий и сервисной командой ситуационного центра. Иначе, новизна состоит в попытке дополнить обычные способы взаимодействия – совещания, конференции, круглые столы и т.п. – проектной работой, включающей систему управления проектом, анализ истории и мониторинг участников.

Концепция предлагаемого подхода заключается в организации процесса управления как коллективного взаимодействия персонала образовательного учреждения и среды существования этого учреждения – социума – по технологии ситуационного центра (СЦ). Основные аспекты технологии СЦ:

- Уникальность анализируемой ситуации и решаемой задачи.
- Формирование, наряду с индивидуальными образами, коллективного образа ситуации, возможно, в виде совместной декларации и т.п.
- Наличие центра как точки сборки, иначе, места в пространстве, которое является определяющим для действий организации.
- Зависящее от решаемой задачи разделение ресурсов на критические и прочие.

2 Концепция реализации

Концепция реализации предполагает, что будут использованы следующие основные методы и подходы [2-3]:

- Систематическое использование промежуточных технологий и прототипов (наиболее простой вариант системы, содержащий наиболее сложный компонент) в соответствии с правилом Парето: 20% правильно выбранных ресурсов обеспечивают 80% результатов.
- Выполнение основных функций ситуационных центров с использованием доступного для вуза набора технических и программных средств.
- Использование для организации работы в ситуационном центре специализированного регламента (методологическая игра) и сервисной команды ситуационного центра.
- Использование проектных групп из участников, работающих по регламенту ситуационного центра с поддержкой сервисной команды.
- Использование для представления информации арсенала как естественнонаучных, так и гуманитарных дисциплин.
- Активное и систематическое использование ресурсов Интернет и средств телекоммуникации (организация сетевого взаимодействия участников разработки, а также участников, вовлечённых в процесс управления).
- Организацию документирования, анализа истории взаимодействия, мониторинга участников, их активности, компетентности и ответственности.
- Создание технологии (системы трансляции подхода), обеспечивающей передачу результатов работы заинтересованным организациям, в том числе подготовка по отдельным компонентам технологии видеоматериалов и размещение их в Интернет.

Кроме того, будет использовано несколько авторских и заимствованных компонентов технологии ситуационного центра. В числе этих компонентов:

- Средства когнитивной визуализации [4, 5].
- Полисенсорный (синестетический) интерфейс.
- Система распознавания речи для оперативного документирования.
- Методики рефлексивного анализа В.А.Лефевра [6].
- Комплекс интерактивных досок: участников, экспертов, рефлексивного анализа.
- Рефлексивный театр ситуационного центра, обеспечивающий параллельную реализацию художественного воплощения и рефлексивного комментирования различных феноменов и приёмов социального действия [7].

Предусмотрено создание пилотного варианта (прототипа) системы на базе учебно-исследовательского ситуационного центра Омского государственного института сервиса (ОГИС), который входит в состав Научно-образовательного центра ОГИС. В качестве членов сервисной команды привлечены аспиранты, темы научных исследований которых связаны с технологиями ситуационного центра. Экспериментальные исследования предполагают проведение тренинговых мероприятий и реальных совещаний по вопросам управления научно-образовательными процессами ОГИС. Участниками этих совещаний, помимо административно-управленческого и профессорско-преподавательского персонала, будут представители городской администрации, общественных организаций, родственники студентов, представители СМИ, а также других социальных групп. Итоги работы будут предметом общественной экспертизы, проводимой в формате конференций с участием всех заинтересованных лиц и организаций.

С 2001 г. в ОГИС осуществляется подготовка студентов специальности «Прикладная информатика (в сфере сервиса)», включающая учебные курсы подготовки группы специалистов для ситуационных центров. С 2009 г. открыта аспирантура по специальности 05.13.01 «Сис-

темный анализ, управление и обработка информации (по отраслям)». Также с 2009 г. образован учебно-исследовательский ситуационный центр ОГИС. Проведены три Всероссийские конференции «Рефлексивный театр ситуационного центра» в 2007, 2008 и 2009 гг. На базе этих конференций сформировался многодисциплинарный сетевой коллектив разработчиков из Абакана, Владивостока, Москвы, Новосибирска и Омска. В первой половине 2010 г. издательством ОГИС будет выпущена монография «Информационные технологии и ситуационные центры», подготовленная этим коллективом.

Заключение

Существуют организации и люди, чья деятельность существенно продвигает применение технологий ситуационного центра. Отметим здесь Клуб Инновационного Развития (<http://www.reflexion.ru/club/>), возглавляемый В.Е.Лепским, а также Центр разработки технологий развития административных и бизнес-структур (<http://razvitiye-plan.ru/>) Г.Г.Малинецкого. Гораздо чаще, однако, трудно определить цели и ресурсы авторов различных проектов.

Наш прогноз состоит в том, что в каждом образовательном учреждении будет оставаться всё меньше пространства для «пузырей» - оболочек реальной проектной деятельности. Однако сетевая реализация проектов вполне реальна. Примером такого проекта является взаимодействие Научно-образовательного центра ОГИС и Владивостокского государственного университета экономики и сервиса (ВГУЭС) по созданию компонентов описанной здесь системы коллективного управления. Помимо личных контактов и телеконференций, готовятся совместные публикации, в том числе и упомянутая выше монография.

Настоящий текст может рассматриваться как приглашение к сотрудничеству по формированию и реализации инновационных проектов в сфере образования.

Список литературы

- [1] Филимонов В.А. Чудесные пузыри образования// Доклады и тезисы Седьмой международной конференции «Перспективы систем информатики» 15-19 июня 2009 г.// Секция «Информатика образования»// Новосибирск: ООО «Сибирское Научное Издательство», 2009, С. 132-135.
- [2] Поляк Ю.Г., Филимонов В.А. Статистическое машинное моделирование средств связи.- М.: Радио и связь, 1988.- 176 с.
- [3] Филимонов В.А. Спинной мозг ситуационного центра// «Ситуационные центры и перспективные информационно-аналитические средства поддержки принятия решений». Сб. материалов Международной научно-практ. конф. 7 - 9 апреля 2008 г. // М.: Российская Академия государственной службы при Президенте РФ, 2009, с. 361-366.
- [4] Tufte E. Beautiful Evidence, 2006, Cheshire, CT: Graphics Press, 213 p.
- [5] Kovalerchuk B. Visual and Spatial Analysis. Advances in Data Mining, Reasoning, and Problem Solving.- Schwing, James (Eds.).- 2004, XX, 576 p. 16 illus. in color.
- [6] Лефевр В.А. Лекции по теории рефлексивных игр.- М.: «Когито-центр», 2009.- 218 с.
- [7] Рефлексивный театр ситуационного центра (РТСЦ-2007)// Всеросс. конф. с междуна. участием : сборник статей / под общей ред. проф. В.А.Филимонова.- Омск: Омский гос. институт сервиса, 2007.- 140 с.: ил. + 16 с. вкл. ил.



New information technologies in Control and Management
Новые информационные технологии в управлении

INTEGRATION OF TELECOMMUNICATION MANAGEMENT SYSTEMS USING ONTOLOGY-BASED METHODS

E.A. Babkin

State University – Higher School of Economics (Nizhny Novgorod)
603155, Nizhny Novgorod, B. Pechorskaya St., 25/12, Russia
eababkin@hse.ru
phone: +7 831-416-9536, fax: +7 831-4169651

Key words: *telecommunication, information models, semantic integration, ontology, information flow theory, relational logics*

Abstract

В этой статье предлагается новый метод интеграции для семантической интероперабельности между разнородными информационными моделями SID и CIM, применяемыми в современных системах управления телекоммуникационным оборудованием. Предлагаемый метод основан на применении формальных онтологий и комбинации теории информационных потоков Барвиза с реляционной логикой системы Alloy Analyzer.

Introduction

In last decades the Telecommunication industry became a bright example of the interoperability challenge during integration of heterogeneous software and hardware distributed systems. In order to support of all stages of life cycle of this product it is required to use in various combinations the multilevel and geographically distributed technical infrastructure uniting diverse network resources, to support constant interaction of different divisions, systems and equipment, to provide fast reaction to changes and their prediction. In such situation the "through-out" business-processes with complete or partial automation, and also completely new business-processes of strategic planning become necessary. Effective organization and management of such processes are hardly achieved by means of the isolated software decisions presently used in various divisions of the telecommunication companies. Local software systems are not ready for such situations, when the majority of the processes of operation, planning and the managements of services cross borders of divisions.

Only integrated systems can support coordinated process-oriented decision of the problems of service planning and management at all levels of the organizational structure of the operator. It was offered to use the specialized class of such systems, which have received the name "Operations Support Systems" (OSS), and also a related class of the software for support of the business-oriented decisions (BSS) [1]. OSS/BSS concept provides for development of standard multilevel information modeling paradigm within the framework of all company. Such standard modeling paradigm should facilitate creation of a comprehensive model of the network resources being managed. Officially several different international organizations are responsible for standardization of information architecture in the telecommunication domain. Such situation leads to existence of different partially incoherent information modeling paradigms in OSS/BSS domain.

Currently TeleManagement Forum (TMF) and Distributed Management Task Force group (DMTF), two leading international consortiums, develop and adopt internationally accepted standards in domains of telecommunications and enterprise IT-infrastructure management. Because these standards are embodied in terms of widely used modeling languages like UML and have reference software implementations many OSS vendors and operators have already recognized significance of the proposals for future sustainable development of telecommunication industry.

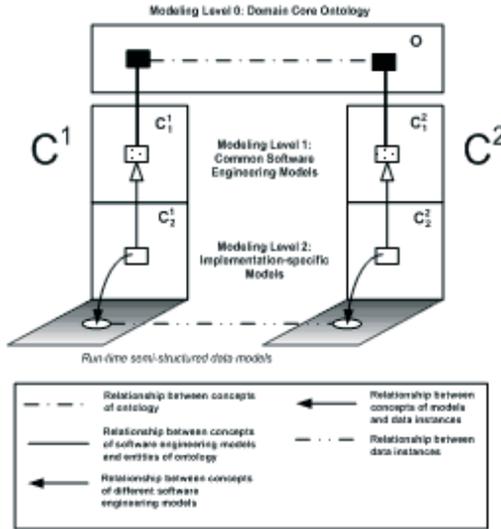
The TMF mainly serves the information services and telecommunication industry. First important TMF's contribution to standardization of telecommunication activities consists of so called enhanced Telecom Operations Map (eTOM). Providing the common modeling language for all activities in eTOM TMF developed comprehensive information model which is called Shared Information Datamodel (SID). In terms of UML class diagrams SID describes major physical and logical entities which are involved in the process of network service provisioning, for example, Customer, Product, Service, Network, Resource. Also general business concepts like location, contract and party are included into the SID model. Such business-oriented structure facilitates comprehensive description of the telecommunication activities which can be easily understood from the business point of view, and, at the same time, has close relationships with the models of software architecture. In parallel DMTF works on standards for configuration and control of general-purpose enterprise IT infrastructure. In order to describe all enterprise resources in a uniform or 'normalized' fashion DMTF developed own information model which is called Common Information Model (CIM) [2]. In the latest CIM specification each schema is expressed in the form of UML class models. In narrow and clearly identified applications domains CIM and SID show good modeling capabilities. For example, the CIM approach allows fault isolation, alert correlation, root-cause and service impact analysis tasks to be performed across multiple domains of equipment, network and service management efficiently. However, now planning and management of modern telecommunication services require so tight integration of information reflecting different points of view, that using single modeling approach (either CIM or SID) becomes impractical.

In absence of a single modeling standard for telecommunications networks software designers and network managers need solid and comprehensive scientific principles of sharing and integration of information models in order to achieve interoperability among heterogeneous OSS components. To develop such principles advanced techniques for knowledge representation should be inevitably applied because it may not be always possible to map low level concepts defined within individual models to each other since they have different semantic meanings.

In that case the role of domain ontologies in the course of integration dramatically increases. It is ontology which contains shared domain concepts and provides principal foundations for connecting technology- or aspect-dependent constituents of particular modeling elements of different modeling paradigms like SID and CIM. However ontologies alone are insufficient for development of complete engineering methodology. In our opinion, which reflects the end-user's point of view, the consistent methods of transformation of ontologies should also be proposed. The origins of ontology transformation are formulated at the highest level of abstraction where concepts of application domain are placed, but software engineering practice must find correspondent answers at any intermediate level, and finally the ultimate response to the interoperability challenge is given at the level of processing of actual data.

1 Proposed method of hierarchical transformation of ontologies

Following the guidelines of the discussion above we developed a theory of hierarchical ontology transformation suitable for support of dynamic integration in telecommunication domain. The process of ontology transformation is initiated at the level of conceptual modeling, then it continues at the architectural software levels (correspondent information models), and finally at the data level it is practically implemented in dynamics during runtime execution of mappings between semi-structured data models. Without loss of generality describing the theory of ontology transformation we use three abstraction levels of modeling hierarchy and two components C^1 , C^2 of the OSS/BSS (Picture 1).



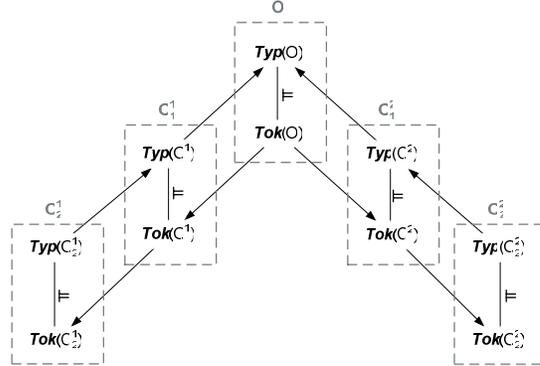
Picture 1 – The illustrative structure of the distributed system

We selected the same formalism of relational logic and the language of Alloy Analyzer system [6] for representation of structure and relationships in ontological and software engineering models. In this case the ontology structure is directly mapped onto the set of atoms of the relational logic and relations identifiers are mapped onto the relational operators of the relational logic. Additional axioms in terms of first-order logic define domain-specific constraints. Further refinement of the theory consists of restricting the available modeling technique and end-user modeling languages for software engineering models. Pragmatic considerations force to use widely accepted object-oriented modeling, and specifically UML and OCL. Because Cranfield in [3] shows applicability of UML and OCL for representation of ontology as well, our theoretical framework obtains the elegant uniform language for the models of all three levels of the modeling hierarchy. In our case the same language of UML class diagrams is used for representation of the domain core ontology and the models of lower levels. OCL is used as the language of axiomatization of domain- or implementation-specific constraints both for domain core ontology and software engineering models. In order to avoid semantic ambiguity of pure UML and OCL structures they are translated to the form of relational logic.

Picture 1 shows that different components of the distributed system are modeled by different software engineering models labeled C_j^i , $i, j = 1..2$ correspondingly. At the same time the developers of components share the same domain core ontology O . Semantic interoperability, which should be achieved in the result of ontology transformation, gives the ability of interpretation of data structures and sharing of the meaning of the data between remote parts of the integrated system. In such a case we can say that when semantic interoperability is achieved then information models are integrated and seamless information flow between remote parts of the integrated systems is maintained. Integration can be viewed as creation of new constraints, which restrict diversity of information models at the software engineering levels in accordance with the logic of the whole system, and mappings between partial views of remote components and the whole system.

Such formulation of integration leads to application of the major principles of information flow theory [4] for expression of ontology transformation. According to the information flow theory in a general case the distributed system \mathcal{A} consists of an indexed family $\text{cla}(\mathcal{A}) = \{A_i\}_{i \in I}$ of classifications together with a set $\text{inf}(\mathcal{A})$ of infomorphisms all having both domain and codomain in $\text{cla}(\mathcal{A})$ [4]. In our specific case of the layered information modeling of two components each separate level of the

software engineering models is represented as a classification C_j^i . Similarly the level of ontology is represented as a classification O . In our theory we propose to use the language of relational logic L as a foundation for definition of types and tokens of classifications C_j^i and O . In this case types are correct sentences of L which define a certain logic micro-theory of UML class diagrams and tokens are L -structures which in fact become the model of the corresponding micro-theory. Correspondingly to the proposed classifications five local logics are defined to express valid constraints on the classification: $\lambda = \langle C_j^i, \vdash_{\lambda_j}, N_{\lambda_j} \rangle$ $i, j = 1..2$; $\lambda = \langle O, \vdash_{\lambda}, N_{\lambda} \rangle$. The theory of each classification has as constraints just the sentences of relational logic L that are valid in the usual sense, and the consequence relation \vdash_{λ} is logical consequence in the language of relational logic. According to the Picture 2 infomorphisms are defined between the classifications of the models and ontology.

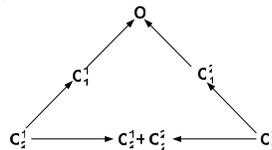


Picture 2 – The structure of infomorphisms between different classifications of information models

The structure of infomorphisms makes up the information channel $\mathcal{C} = \{f^i : C_1^i \iff O\}$ $i = 1..2$, where the classification of ontology O plays a role of the core of the information channel. On tokens f^i assigns to each token of the ontology the correspondent L -structure of the software engineering models. On types, f^i assigns certain concepts of ontology defined by L -sentences to the subset of the software engineering models. The core of the information channel represents the whole system view, and tokens define connections between information models of the separate components. Given the described structure of infomorphisms between classifications, the sum of classifications of implementation-specific models ($C_2^1 + C_2^2$) can be defined (picture 3). It is the classification defined as follows [4]:

- 1) The set $Tok(C_2^1 + C_2^2)$ is the Cartesian product of $Tok(C_2^1)$ and $Tok(C_2^2)$.
- 2) The set $Typ(C_2^1 + C_2^2)$ is the disjoint union of $Typ(C_2^1)$ and $Typ(C_2^2)$.

Using the Universal Mapping Property for Sums [4] the following diagram is made.



Picture 3 – The sum of classifications in the information channel

As the sum ($C_2^1 + C_2^2$) has common properties of every infomorphisms of the information flow theory, it has own local logic. According to [4] the local logic of the sum ($C_2^1 + C_2^2$) is called the distributed logic $Dlog_{\mathcal{C}}(\lambda)$ of the information channel \mathcal{C} generated by λ . $Dlog_{\mathcal{C}}(\lambda)$ represents the

reasoning about relations among the information models of the components of the distributed system justified by the logic of the core. In other words the created structure of infomorphisms allows moving local logics around from one classification to another. Analysis of formal properties of $\text{Dlog}_{\mathcal{C}}(\mathcal{A})$, which was done by Barwise in the framework of the information flow theory, shows that it is not generally sound. Because moving of local logics does not preserve soundness and completeness, we have a theoretical justification for necessity of manual procedures in the course of ontology transformation. Concepts of Infomorphism and Information Channel give answers to “what” and “why” questions – they explain what part of information is moved and why elements of one information model move information about other information models in the context of the whole system. But in order to produce real-time data elements of semi-structured data models we need to perform manual procedures of mapping, answering the important pragmatic question “how”.

In terms of our theory that is of manually defined mapping μ between the tokens of the sum classification $(\mathbf{C}_2^1 + \mathbf{C}_2^2)$:

$$\mu : \text{Tok}(\mathbf{C}_2^1) \xrightarrow{\text{map}} \text{Tok}(\mathbf{C}_2^2).$$

As our theory shows that mapping of the tokens of the sum classification is a minimally necessary manual procedure for the process of ontology transformation. In principle different computational paradigms may be applied for definition and execution of that procedure.

2 Illustration of the theory application

To get insight to the presented theory of ontology transformation and its practical implications a simple but important example of identity transformation is briefly considered. In that case the shared domain core ontology has a single entity which connects two fragments of different software engineering models. In our example the shared domain ontology defines the generic concept of rack mounted physical equipment. There are two common software engineering UML models which describe engineering concept of rack mounted equipment in terms of different modeling paradigms (SID and CIM correspondingly). These models become the models \mathbf{C}_1^1 and \mathbf{C}_1^2 in accordance with the theory of ontology transformation (Picture 1). Applying principles of the theory of ontology transformation to the given models we need to define precisely semantics of infomorphisms f^i between classifications which comprise the information channel $\mathcal{C} = \{f^i : \mathbf{C}_1^i \rightleftarrows \mathbf{O}\} \ i = 1..2$, and correspond to the software engineering models and the shared domain ontology. Such formal semantics allows consistent expressing of UML structure of ontology and software engineering models in terms of the language of relational logic.

In the case of ontology transformation the most important part of infomorphism f^i defines mapping between types of classifications: $\text{Typ}(\mathbf{C}_1^i) \xrightarrow{\text{map}} \text{Typ}(\mathbf{O})$. Semantics of this part of the infomorphism between classifications can be naturally represented as the UML association “one-many” between the ontology concept and semantically related UML classes of the common engineering models. This approach defines foundation principles for reusable and expandable expression of ontology concepts and correspondent classes of UML models in terms of relational logic and Alloy language. In accordance with this approach reusable definition of shared domain core ontology includes a generic logical micro-theory. This micro-theory states, that in the result of mapping between the types of the classifications any ontology concept links two different software engineering models via classes relations and associations relations. The infomorphism mapping should possess the property of closeness: for all connected classes their associations should be included in the infomorphism.

The next step is reformulation of UML class diagrams in terms of relational logics. With the help of freely available UML2Alloy tool [5] this task is almost automated. Given Alloy-based definition of UML class diagrams the infomorphism mappings between L -sentences of ontology and L -sentences of CIM- and SID-based engineering models are defined as a certain type of specialization of generic micro-theory. By the similar manner reformulation of specific software engineering models should be provided in terms of Alloy language and the following mappings should be established between L -sentences of common models and L -sentences of specific models.

As far as the structure of specific software engineering models was constrained to concrete subclasses of abstract classes in common models, the mappings are defined by a straightforward way: all signatures of the concrete software engineering models extend correspondent signatures of the common software engineering models. Additionally domain-specific OCL constraints are translated to the appropriate facts in Alloy language.

Due to specific capabilities of Alloy analyzer finding of these L -structures may be done automatically in the result of the model analysis. To do this task some auxiliary predicate Model should be defined and the command run is issued. In the result Alloy Analyzer will return several instances of the L -structures which satisfy the posed constraints. These L -structures form the token part of the sum classification ($C_2^1 + C_2^2$) and relate connected instances of CIM and SID models for the given application specific constrains. In the given example there is a restriction on the number of slots in the rack and following L -structures are found.

Because in practice definition of intermediate mappings between classes of the concrete software engineering models is not always possible, detailed specification of correspondences between tokens of the sum classification can not be done automatically in the framework of the theory of ontology transformation. Such specification should be performed manually in accordance with certain engineering methodology.

Conclusion

In this article we introduced theory of hierarchical transformation of ontologies suitable for semantic integration of different information models in Telecommunication domain. We briefly demonstrated the steps necessary for practical application of that theory when relational logic and Alloy analyzer tool is used for integration of two widely used modeling paradigms in the form of UML class models (SID and CIM). The proposed approach facilitates automated integration routines however give some freedom to the integrators for manual definition of mappings between different concepts of the information models.

Literature

- [1] Terplan K. *OSS Essentials: support system solutions for service providers*. Willey Computer Publishing. ISBN 0-471-39240-5. 2001.
- [2] Bumpus W., et all. *Common Information Model: Implementing the Object Model for Enterprise Management*. Willey Computer Publishing. ISBN 0-471-35342-6. 2000.
- [3] Cranefield S. *UML and the Semantic Web*. In Proc. of the Intrl. Semantic Web Working Symposium, SWWS'01. July 2001.
- [4] Barwise J., Seligman J. *Information Flow*. Cambridge University Press, 1997.
- [5] UML2Alloy Project. Online: <http://www.cs.bham.ac.uk/~bxb/UML2Alloy/index.php> . Last visited (February 2010).
- [6] Jackson D. *Software Abstractions: Logic, Language and Analysis*. The MIT press, Cambridge, Massachusetts. 2006.

КОМПЬЮТЕРНЫЙ КОНТРОЛЬ ЗНАНИЙ – МИФ ИЛИ РЕАЛЬНОСТЬ

А.Н. Гинзбург¹, В.М. Дёмкин²

¹Институт прикладной физики РАН
603005, Нижний Новгород, ул. Ульянова, 46, Россия
aginzburg@bk.ru
тел. +7 (831) 436-80-67

²Нижегородский филиал Государственного университета – Высшая школа экономики
603155, Нижний Новгород, ул. Б.Печерская, 25/12, Россия
valerdem@mail.ru
тел: +7 (831) 416-95-49

Ключевые слова: компьютерные формы контроля знаний, открытые и закрытые тесты, мультиагентные системы тестирования, онтологический подход, холоны и акторы

Abstract

This article presents some results of the ontological approach for conceptual design of system of automatic checking of knowledges. “Recognition” and “guessing” serve as foundation of mythologization of both procedure of control of knowledges and results of control. A testable subject and his answers must become the actors of the multi-agent system of automatic checking of knowledges. Whether it will become by a new myth, the future will show.

Введение

Среди ключевых факторов, оказывающих влияние на совершенствование и повышение качества учебного процесса в образовательных учреждениях и определяющих направление развития контроля знаний в целом, особое место по праву принадлежит информационным технологиям. Активное внедрение в учебный процесс различных форм компьютерного тестирования стало неотъемлемой частью современного образования.

Интернет-технологии уверенно вошли в практику работы учебных заведений системы высшего и среднего профессионального образования, подтверждением тому служит увеличение числа участников Интернет-проектов, разработанных Научно-исследовательским институтом мониторинга качества образования [1].

Другим ярким примером стал проводимый Национальным аккредитационным агентством в сфере образования эксперимент по введению Федерального экзамена в сфере высшего профессионального образования (ФЭПО). Содержанием эксперимента является проведение компьютерного Интернет-тестирования в части внешней оценки уровня подготовки студентов на соответствие требованиям государственных образовательных стандартов [2].

Однако следует отметить, что подобного рода мероприятия имеют отношение к так называемым традиционным компьютерным формам тестирования, где проверяются только теоретические знания. Ключевой идеей такого тестирования является “узнавание”, поэтому здесь широко используются закрытые и открытые тесты на основе экранных форм шаблонов. Процедура проверки заключается в посимвольном сравнении шаблона с правильными ответами и заполненного шаблона с ответами тестируемого субъекта. Поскольку наряду с “узнаванием” имеет место и “угадывание”, то такие формы компьютерного тестирования служат основанием для мифологизации как самой процедуры контроля знаний, так и результатов контроля.

Таким же недостатком “страдают” и формы абитуриентского бланочного тестирования, квинтэссенцией которого стали, например, Межрегиональные многопрофильные олимпиады, проводимые под эгидой Государственного университета – Высшая школа экономики [3].

1 Компьютерный контроль знаний – это миф?

Перед тем как ответить на поставленный вопрос, напомним основные достоинства традиционной формы компьютерного тестирования:

- отсутствие затрат на распечатку заданий, бланков ответов, инструкций и другой документации;
- возможность быстрого оценивания результатов;
- возможность выдачи задания разного уровня сложности в зависимости от предыдущих ответов тестируемого субъекта (адаптивное тестирование), и как следствие, более точная оценка его достижений;
- возможность рейтингового контроля ответов тестируемого субъекта;
- возможность дистанционного проведения оценочных мероприятий с использованием Интернет-технологий;
- возможность формирования задания непосредственно в ходе выполнения работы.

Продолжая, укажем на возможный перечень действий со стороны тестируемого субъекта при выполнении тестовых заданий закрытой формы:

- выбрать один правильный ответ;
- выбрать несколько правильных ответов;
- соотнести ответы к вопросам.

Также укажем и на возможный перечень действий со стороны тестируемого субъекта при выполнении тестовых заданий открытой формы:

- ввести правильный ответ;
- ввести ответ в текст задания.

Контроль такого рода “знаний” не способен выявить наличия у тестируемого субъекта ни творческого начала, ни практических навыков. Его только спрашивают, что он может узнать или угадать из указанного контекста, не предоставляя ему возможности оказывать влияние на сам процесс тестирования. Результаты контроля всего лишь фиксируют состояние процессов “узнавания” и “угадывания”. Они не достоверны и иллюзорны, именно поэтому способствуют укреплению мифа о действенности такой формы контроля знаний.

Накопленный опыт сообщества разработчиков программного обеспечения показывает, что предпринимаемые с их стороны усилия по “борьбе” с “узнаванием” и “угадыванием” не оправданы по затратам и бесплодны по своей сути. Миф о “всесилии” компьютерных технологий в вопросе контроля знаний культивируется десятилетиями!

2 Компьютерный контроль знаний – это реальность?

Среди учебных дисциплин особняком стоят те, где присутствует творческое начало. Таковой, например, является дисциплина программирования. Чтобы “побороть” рутину традиционных форм компьютерного тестирования, необходимо от полной проверки программного кода идти к его частичной проверке [4-5]. Лишь в этом случае можно надеяться на успех в достижении контроля практических навыков программирования, а значит и контроля знаний.

Чтобы подобное утверждение было справедливым по отношению и ко всем остальным учебным дисциплинам независимо от присутствия или отсутствия в оных творческого начала, необходима новая парадигма компьютерного тестирования.

Становление новой парадигмы, прежде всего, должно опираться на стремление вернуть человеку активную роль в осуществлении компьютерного контроля его знаний. В противном случае все усилия по совершенствованию традиционных форм компьютерного тестирования окажутся напрасными, так как будут ограничены рамками парадигмы классической рациональности. Как следствие, подобные попытки по-прежнему будут способствовать укреплению мифа о действенности форм такого рода контроля знаний.

Одним из возможных шагов на пути перехода от старых форм к принципиально новым может оказаться попытка применения такого подхода к построению сложных искусственных систем на основе онтологий, где краеугольным камнем новизны будет являться парадигма ограниченной рациональности принятия решений [6-7].

Перед тем как перейти к обсуждению свойств будущей новой системы тестирования знаний, перечислим основополагающие принципы парадигмы ограниченной рациональности:

- 1) Ограниченная рациональность опирается на фундаментальную онтологию М. Хайдеггера, в которой человек не только есть, но соотносится с собой, миром и миром других людей, что позволяет говорить о множестве персональных онтологий.
- 2) Ограниченная рациональность предполагает “встроенность” человека в ситуацию.
- 3) Ограниченная рациональность предполагает нахождение согласованного решения на основе солидарности неоднородных акторов (лиц, не только познающих, но и преобразующих ситуацию).
- 4) Ограниченная рациональность опирается на взаимодействие по принципу “часть-целое” элементов открытых систем.
- 5) Формализованное описание ситуаций осуществляется с помощью онтологических моделей ситуаций, включающих в себя персональные и групповые онтологии, а также объектные модели, разрабатываемые на основе взаимодействия неоднородных акторов.

Итак, парадигма ограниченной рациональности опирается на отношение “часть-целое”. Каждая часть рассматривается как целостность (или целое), являющееся частью другого большего целостного (или целого). Целостность, участвующую в становлении новых целостностей, принято называть холоном. Принцип целостности предполагает взаимодействие целостных компонентов (холонов). Ключевой фигурой парадигмы ограниченной рациональности является человек-актор, берущий на себя полномочия принимать решения и несущий ответственность за последствия принимаемых решений. В этом смысле человек – это главное “действующее лицо” холона.

Следуя основополагающим принципам парадигмы ограниченной рациональности, укажем, что новая система тестирования знаний:

- функционирует при участии тестируемых субъектов, являющихся “одушевленными акторами”, и их ответов на задания теста, являющихся “неодушевленными акторами”;
- является открытой системой, взаимодействующей с окружением и информацией, хранящейся и порождаемой в “недрах” системы;
- характеризуется неопределенностью и изменчивостью возникающих в ней проблемных ситуаций, побуждающих акторов принимать решения;
- организуется в качестве динамически упорядоченных целостностей, в которых образуются и совершенствуются взаимосвязи между частями целого.

Итак, в системе взаимодействуют тестируемые субъекты и их ответы, все они обладают статусом холонов. Холоны “толстеют”, т.е. становятся составными благодаря новым “правильным” ответам, или “худеют” в противном случае. Акторы как “действующие лица” получают полномочия по управлению холонами и несут ответственность за последствия принимаемых решений. Онтологические модели ситуаций исполняют роль платформы, на базе которой происходят процессы согласования “точек зрения” и “взаимных уступок” акторов. Акторы изучают проблемную ситуацию, сложившуюся в системе, и принимают решения о путях выхода из нее, согласуя свои действия с другими акторами. В конечном итоге будет найдено одно из решений: “актор обладает знанием” или “актор не обладает знанием”. Принятие решения будет получено в темпе развития ситуации, т.е. в реальном масштабе времени.

Становление новой парадигмы позволит разрушить миф о действенности современных форм компьютерного контроля знаний. Тестируемый субъект и его ответы как активные творческие элементы открытой системы будут создавать и разрешать ситуации, которые позволят выявить глубину знаний. Станет ли новый подход очередным мифом, покажет будущее.

Заключение

Сформулируем основные результаты работы:

- 1) Оставаясь в рамках традиционного подхода к компьютерным формам тестирования, будем камень за камень “укреплять” фундамент мифологизации как самой процедуры контроля знаний, так и результатов контроля. Как говорят, “какие мельницы, такая и вода”.
- 2) Смена парадигмы компьютерного тестирования вынудит разработчиков программного обеспечения перейти к разработке открытых систем в рамках парадигмы ограниченной рациональности, опирающейся на отношение “часть-целое”.

Список литературы

- [1] <http://www.i-exem.ru>
- [2] <http://www.fepo.ru>
- [3] <http://talent.hse.ru/olimp/mmo/materials>
- [4] Гинзбург А.Н., Дёмкин В.М., Рыбалкин А.Н., Сатунин С.В. Адаптивный подход к управлению процессом разработки систем автоматического тестирования решения задач по программированию. Результаты и перспективы разработки // Проблемы управления и моделирования в сложных системах: Труды IX Международной конференции (Самара, 22-28 июня 2007г.). – Самара: Самарский научный центр РАН, 2007, с. 603–613.
- [5] Антропов А.А., Баранов Е.В., Гинзбург И.А., Дёмкин В.М., Мотолин А.В. Архитектура Web-ориентированной системы автоматического контроля знаний дисциплины программирования // Проблемы управления и моделирования в сложных системах: Труды XI Международной конференции (Самара, 22-24 июня 2009г.). – Самара: Самарский научный центр РАН, 2009, с. 658–663.
- [6] Виттих В.А. Парадигма ограниченной рациональности принятия решений – 1 // Вестник Самарского государственного технического университета. Серия: Технические науки, №3 (25), 2009, с. 22–31.
- [7] Виттих В.А. Парадигма ограниченной рациональности принятия решений – 2 // Вестник Самарского государственного технического университета. Серия: Технические науки, №1 (26), 2010, с. 23–31.

КАТЕГОРИАЛЬНЫЙ АНАЛИЗ В ЗАДАЧАХ ПРИНЯТИЯ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ

О.И. Бедняк¹, М.А. Кораблин¹, А.А. Салмин¹, С.В. Смирнов²

¹Поволжский государственный университет телекоммуникаций и информатики
443010, Самара, ул. Льва Толстого, 23, Россия
ist@psati.ru

тел: +7 (846) 228-00-21, факс: +7 (846) 228-00-05

²Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
smimov@iccs.ru

тел: +7 (846) 332-39-27, факс: +7 (846) 333-27-70

Ключевые слова: категориальный анализ, свойства и категории качества объектов, идентификация объектов, сегментирование, байесовская вероятность

Abstract

The questions of the object's identification of "hidden essence" in the problems of management decision making are considered. A category analysis over of object's properties on the basis of Bayesian approach is produced. The algorithm for identifying objects is presented.

Введение

Существует класс трудно формализуемых задач, связанных с идентификацией скрытой сущности объекта. Под термином «скрытая сущность» подразумеваются особые черты объекта (например, личностные свойства индивида, его склонности к определенным поступкам, его соответствие сложившейся ситуации и т.п.). Эти и аналогичные им стороны объекта, как правило, не лежат на поверхности и в то же время они в определенных ситуациях весьма важны для принятия управленческих решений, связанных с этим объектом.

Задачи идентификации скрытой сущности многочисленны и могут принадлежать разным предметным областям, например:

- выдача кредитов клиенту банка (идентификация порядочности клиента);
- оценка лояльности и оттока клиентов телекоммуникационной компании;
- соответствие должности, занимаемой объектом (идентификация соответствия);
- достоверная оценка качества продукции;
- выявление авторства неизвестного текста;
- идентификация спама;
- определение диагноза болезни по набору признаков (симптомов) и др.

Порядочность, лояльность, достоверность, соответствие, адекватность, – все это виды скрытой сущности, которая должна быть оценена не только качественно, но и количественно. Формирование такой оценки обычно связывают с вероятностной мерой, что определяет наличие реальных рисков ошибочной идентификации объекта. Насколько конкретный объект соответствует его имиджу, реальный ли это имидж или мнимый – ответ на эти вопросы имеет важное, а иногда решающее значение.

В качестве объекта может выступать фирма (предприятие), клиент, техническая система и т.п. Не умаляя общности можно утверждать, что особое значение методы идентификации скрытой сущности приобретают для задач, в которых объектом исследования является человек

как представитель определенного социума (клиент компании, сотрудник предприятия), как носитель болезни, обладатель определенных знаний и т.п.

Скрытую сущность объекта трудно оценить, - традиционные математические модели идентификации в этих задачах редко оказываются полезными, здесь чаще используются такие методы как поиск прецедентов, формирование «кредитных историй», выявление достоверных фактов, сравнение с аналогами и т.п. Использование этих и им подобных методов в совокупности приводит к формированию общего многогранного имиджа объекта, собранного из разнородных сведений, данных, косвенных признаков и неочевидных связей.

История этого направления восходит к концепциям экспертных систем [1], дискриминантного анализа и проверки гипотез, методам Data Mining [2] и т.п.

В отличие от этих традиционных направлений мы рассматриваем категориальный анализ как средство получения результатов путем «слепого» поиска на множестве свойств объекта прямых и косвенных признаков, присущих объекту, и анализа стереотипов.

Концептуальная основа категориального анализа строится на постулате, что набор таких свойств и признаков неявно связан с идентифицируемой сущностью объекта, причем, чем более представительным является этот набор, тем сильнее такая связь.

1 Свойства и категории качества объектов

В основу категориального анализа ложится состояние объекта, характеризуемое перечнем (обычно статическим) всех свойств данного объекта и текущими (обычно динамическими) значениями каждого из этих свойств, которые могут быть как присущи объекту, так и приобретены с течением времени. Именно эти свойства делают объект индивидуальным, особенно если в качестве объекта рассматривается клиент. Позиционирование такого объекта и составляет задачу идентификации скрытой сущности.

Любой объект характеризуется свойствами и категориями качества. Свойства характеризуют общие черты объекта, например, «запах», «цвет», «возраст» и т.п. Категории определяют значения качества свойств (например, «резкий» запах, «красный» цвет, «пожилой» возраст). Чем больше свойств и категорий качества участвуют в описании объекта, тем конкретнее описан объект, тем он определеннее, тем полнее его описание. Свойства и категории качества объектов рассматриваются как основные черты, характеризующие его с разных точек зрения. Например, свойство «отцовства» – объект (отец) имеет детей (является отцом). При этом качество «отцовства» характеризуется количеством детей. Такого рода качество имеет несколько категорий: многодетный отец, малодетный отец, бездетный отец и т.п.

Любое свойство несет в себе определенную информативность, если, например, рассматривается возможность использования конкретного индивида в качестве водителя автобуса, то свойства «Пол» (иметь пол) и «Возраст» имеют интуитивно большую информативность, чем свойство образованности («Образование»).

Свойства объекта могут быть тесно связаны с интересами компании (например, материальное состояние клиента или количество иждивенцев). Свойство обладания материальными ценностями является очень важным для, например, выдачи кредита, в то же время свойство образованности клиента для этой ситуации может играть второстепенную роль. Все резко меняется, если кредит выдается в виде гранта для продолжения научной работы.

Главное и второстепенное в категориальном анализе подвержено резким изменениям, в некоторых случаях фактор изменчивости понятий может полностью изменить оценку сущности, в этом состоит особая важность изучения свойств, когда второстепенные на первый взгляд свойства в определенных ситуациях могут стать важнее тех, которые считаются основными.

Кроме того свойства и категории можно разделить на очевидные и неочевидные. Очевидные характеризуются тем, что объекту можно поставить в соответствие значение категории качества. Например, мы знаем, что возраст объекта имеет значение «пожилой». Неочевидные

свойства таковы, что мы не можем уверенно поставить в соответствие объекту значение категории качества, но субъективно мы можем оценить такое значение, используя субъективные вероятности, характеризующие отношения между категориями в свойстве.

Совокупность свойств и категорий качества, используемых в задачах категориального анализа, в общем случае целесообразно представить в виде таблицы с буквенно-индексными обозначениями категорий (таблица 1), где каждое свойство объекта определено совокупностью категорий качества (в каждой строке число категорий больше единицы).

Таблица 1 – Свойства и категории качеств

Свойства	Категории качества					
	1	2	...	m	...	M
1	K_1^1	K_2^1	...	K_m^1	...	K_M^1
2	K_1^2	K_2^2	...	K_m^2	...	K_M^2
...
N	K_1^N	K_2^N	...	K_m^N	...	K_M^N

Верхний индекс в записи K_j^i идентифицирует свойство, нижний – категорию качества, присущую объекту, подчеркивание определяет выбранную категорию соответствующего свойства. Выбор категории качества интерпретируется как категориальное событие – конкретный факт индивидуальной характеристики объекта (ИХО).

Категории в любой строке таблицы альтернативны, т.е. они образуют полную группу несовместных категориальных событий. Любая строка таблицы свойств и качеств может быть расширена путем введения более «тонких» категорий или сокращена введением более «грубых».

Для каждого свойства должна быть выбрана только одна категория качества. Формируемый при этом набор категорий интерпретируется как совокупность фактов, определяющих ИХО. Для таблицы 1 такой набор может быть представлен следующим образом:

$$(1) \text{ ИХО} = (\underline{K_1^1} \ \& \ \underline{K_2^2} \ \& \ \dots \ \& \ \underline{K_m^N}).$$

Категории качества в этом наборе конъюнктивны (совместные категориальные события).

Набор категорий качества должен характеризовать объект с разных точек зрения, чем богаче и разностороннее этот набор, тем выше достоверность оценки скрытой сущности объекта.

ИХО представляет собой не только характеристику одного отдельно взятого объекта, - в общем случае эта аббревиатура представляет собой стереотип, который рассматривается как множество объектов с одинаковыми ИХО.

Особое значение в этих рассуждениях отводится понятию «скрытая сущность» и её категории качества. Например, сущностное свойство клиента банка - «Претендовать на получение кредитов». В этой ситуации важно знать сущность клиента, – к какой категории его отнести: либо его следует охарактеризовать как исправного плательщика, аккуратно оплачивающего долг, либо оценивать его как ненадежного плательщика. В этом примере имеем две категории сущности:

- 1) «неплательщик» или
- 2) «надежный плательщик».

Отметим, что эти категории альтернативны, т.е. с позиций теории вероятностей они образуют полную группу несовместных событий. Множество объектов – заемщиков, попадающих в соответствующую категорию сущности (1 или 2), будем называть сегментом. В общем случае сегмент S_g – это совокупность объектов, принадлежащих одной из категорий сущности. В этом примере имеем две категории сущности и соответственно два сегмента, в общем случае

количество сегментов может быть и другим, бóльшим 2-х, но поскольку существенные категории всегда альтернативны, любой объект может попасть только в один из рассматриваемых сегментов.

2 Посегментная статистика свойств и категорий

Посегментная статистика свойств и категорий качеств существенно расширяет структуру таблицы 1 в двух направлениях (таблица 2):

- 1) путем введения «карманов» S_j^i для статистических данных по каждой категории качества K_j^i .
- 2) путем формирования статистики свойств и категорий для каждого сегмента S_{g_n} .

Таблица 2 – Посегментная статистика свойств и категорий качеств для сегмента S_{g_n}

Свойства	Категории качества					Объем выборки
	1	2	...	m	M	
1	K_1^1	K_2^1	...	K_m^1	K_M^1	—
	S_1^1	S_2^1	...	S_m^1	S_M^1	Σ_1
2	K_1^2	K_2^2	...	K_m^2	K_M^2	—
	S_1^2	S_2^2	...	S_m^2	S_M^2	Σ_2
...

N	K_1^N	K_2^N	...	K_m^N	K_M^N	—
	S_1^N	S_2^N	...	S_m^N	S_M^N	Σ_N

В каждом из выше упомянутых карманов содержится сумма объектов S_j^i , попавших в статистику соответствующих категорий. Поскольку для каждого i -ого свойства статистика формируется на одном и том же множестве объектов, очевидно условие:

$$(2) \quad \Sigma_1 = \Sigma_2 = \dots = \Sigma_N \cdot$$

Таблица 2 представляет сегмент S_{g_n} , для других сегментов используются аналогичные по форме таблицы, но статистика, собираемая в них, для различных сегментов отличается.

Отношение

$$(3) \quad P_j^i = \frac{S_j^i}{\Sigma_i}$$

интерпретируется как наполняемость соответствующей категории (соответствующего «кармана»). В то же время эти отношения определяют эмпирические вероятности (частоты) категориальных событий.

Интуитивно ясно, что чем больше статистика одного сегмента (например «Надежный плательщик») отличается от статистики другого («Неплательщик»), тем с большей вероятностью можно идентифицировать скрытую сущность объекта. Худший случай, когда статистика сегментов не имеет существенных различий. Для того чтобы провести надежную идентификацию, необходимо априори сформулировать наиболее информативные свойства, несвязанные друг с другом, кроме того, их число должно быть достаточно велико.

Формирование по сегментной статистики может проводиться с использованием обучающей выборки объектов, заранее классифицированных с высокой степенью достоверности. Отсутствие такой выборки в отдельных случаях можно заменить введением дополнительных свойств и категорий качества.

Таблица 2 может иметь и иную структуру, например, в «карманах» могут размещаться непосредственно отношения типа (3) или субъективные вероятности [3], - смысловое содержание таблицы при этом остается прежним.

При использовании отношений (3) для любого i -го свойства в соответствии с альтернативностью категорий и отношением (2) выполняется условие нормировки:

$$\sum_{j=1}^M P_j^i = 1.$$

Вероятности P_j^i следует рассматриваться как условные, то есть относящиеся к категориальным событиям в конкретном сегменте S_{g_n} : $P_j^i = P(K_j^i | S_{g_n})$, причём для всех S_{g_n} и для любого i $\sum_{j=1}^M P(K_j^i | S_{g_n}) = 1$.

3 Алгоритм идентификации объекта

Система сегментов S_{g_n} ($n = 1, 2, \dots, L$) может быть представлена условными эмпирическими вероятностями $P_j^i = P(K_j^i | S_{g_n})$, $i = \overline{1, N}$; $j = \overline{1, M}$, которые в совокупности образуют многомерный информационный образ сегмента. Каждый сегмент имеет собственный образ, отличающийся от остальных.

Посегментная статистика свойств и категорий, которую образуют эти образы, рассматривается как статистическая база данных.

Кроме статистики свойств и категорий в каждом из сегментов, имеется набор априорных вероятностей по сегментам. Эти две компоненты являются основой для выявления скрытой сущности объекта.

В качестве меры «тяготения» объекта к соответствующему сегменту предлагается использовать байесовскую вероятность [4], которая интерпретируется как оценка влияния свойств и категорий объекта (фактов ИХО) на апостериорную вероятность. Таким образом, ИХО рассматривается как основной компонент, влияющий на «превращение» априорной вероятности тяготения в апостериорную.

Несмотря на множество различных технологий по выявлению скрытой сущности объекта (и в первую очередь Data Mining), полагаемся на широко известное мнение, что, «несмотря на простоту байесовских процедур, результаты их работы могут превзойти результаты работы более сложных алгоритмов классификации».

Собственно пересчёт априорных вероятностей в апостериорные на основе фактов, входящих в ИХО, реализуется хорошо известной формулой Байеса [5], которая при введенных выше обозначениях выглядит следующим образом:

$$(4) \quad P(S_{g_n} | ИХО) = \frac{P(ИХО | S_{g_n}) \cdot P(S_{g_n})}{\sum_{n=1}^L P(ИХО | S_{g_n}) \cdot P(S_{g_n})}.$$

В формуле (4) $P(S_{g_n})$ – априорные вероятности попадания клиента в сегмент S_{g_n} , $P(S_{g_n} | ИХО)$ – условные вероятности тяготения объекта к сегменту S_{g_n} при условии, что ИХО

является фактом, $P(IХО|S_{g_n})$ – обратная вероятность (вероятность наличия фактов ИХО в сегменте S_{g_n}).

Необходимо отметить, что формула (4) предполагает, что параметры в наборе ИХО независимы. Таким образом, предположение независимости свидетельств позволяет записать:

$$P(IХО|S_{g_n}) = P(K_1^1|S_{g_n}) \cdot P(K_m^2|S_{g_n}) \cdot P(K_k^3|S_{g_n}) \cdot \dots \cdot P(K_M^N|S_{g_n})$$

Отсюда (4) примет вид:

$$P(S_{g_n}|IХО) = \frac{P(K_1^1|S_{g_n}) \cdot P(K_m^2|S_{g_n}) \cdot P(K_k^3|S_{g_n}) \cdot \dots \cdot P(K_M^N|S_{g_n}) \cdot P(S_{g_n})}{\sum_{n=1}^L P(K_1^1|S_{g_n}) \cdot P(K_m^2|S_{g_n}) \cdot P(K_k^3|S_{g_n}) \cdot \dots \cdot P(K_M^N|S_{g_n}) \cdot P(S_{g_n})}$$

Одним из показателей адекватности выдвигаемых предположений в соответствии с представленной методологией категориального анализа может выступать траектория изменения априорной неопределённости, которая показывает зависимость апостериорной вероятности от последовательно представляемых условий (набора ИХО). Данная траектория показывает, как происходит изменение апостериорных вероятностей в соответствии с алгоритмом пересчёта и последовательном представлении условий при приближении к результату. Причём наиболее достоверный результат будет в том случае, когда траектория распределения апостериорных вероятностей в зависимости от факторов, влияющих на результат, будет наиболее «гладкой». Например, в результате пересчёта получили, что из 5-и предъявляемых ИХО вероятность принадлежности объекта к сегменту при 2-ом условии равна 0.1, а вероятности при предъявлении остальных ИХО изменяются в пределах от 0.4-0.6. Результат 2-го условия резко изменяет траекторию апостериорной вероятности, что сказывается на достоверности полученного результата. Анализ подобного рода траектории изменения априорной неопределённости позволит визуально охарактеризовать изменения априорной вероятности с учётом предъявляемых ИХО и сделать вывод о предъявляемом наборе условий и об адекватности полученных результатов пересчёта в целом.

Заключение

Таким образом, в развиваемом подходе основой категориального анализа является общая и посегментная статистика объектов, дифференцированная по секторам, свойствам и категориям. Наличие и представительность такой статистики определяет достоверность оценки скрытой сущности объекта на основе свойственных для него фактов и косвенных признаков как по отдельности, так и в совокупности.

Формула (4) предполагает независимость свойств, входящих в ИХО. Вместе с тем это предположение нередко выглядит весьма ограничительным. Например, зависимость свойства «Образование» от свойства «Возраст» (чем больше возраст, тем вероятнее наличие высшего образования). На основе предположения о независимости свойств строится так называемый наивный байесовский классификатор [6], который использует прямое произведение вероятностей категориальных событий. Такое предположение характеризуется линейной сложностью процедуры байесовской классификации, отсутствие независимости свойств придает этой процедуре экспоненциальную сложность.

Следует также отметить, что категориальный анализ, основанный на вероятностном подходе, позволяет не только количественно оценить принадлежность объекта к определённому сегменту, но и дополнить имеющуюся базу новыми апостериорными значениями по новым объектам. Кроме того, подобного рода анализ является универсальным, так как внесение каких-либо новых переменных не изменяет принцип анализа, а лишь усложняет определение условных вероятностей в связи с увеличением исследуемых факторов. Вместе с тем такой пе-

перасчёт способен значительно увеличить информативность процесса получения новых знаний на основе основного свойства информации – её аддитивности.

Кроме того, при идентификации объекта важно оценить вклад каждого отдельного взятого свойства в изменение апостериорной вероятности принадлежности объекта сегменту. Такие вклады позволяют ранжировать влияние того или иного свойства на оценку скрытой сущности.

Список литературы

- [1] Балтрашевич В.Э Реализация инструментальной экспертной системы. – СПб.: Политехника, 1997.
- [2] Назначение систем Data Mining [Электронный ресурс]. – Режим доступа: <http://www.iso.ru/journal/articles/276.html>, свободный. – Загл. с экрана.
- [3] Бернштейн С.Н. Собрание сочинений. Т.4. Теория вероятностей. Математическая статистика. – М.: Гостехмиздат, 1964.
- [4] Байесовские процедуры классификации: Вводный обзор [Электронный ресурс]. – Режим доступа: <http://www.spc-consulting.ru/DMS/Machine%20Learning/MachineLearning/Overviews/NaiveBayesClassifierIntroductoryOverview%20.htm>, свободный. – Загл. с экрана.
- [5] Кораблин М.А., Мелик-Шахназаров А.В., Салмин А.А. Оценка лояльности клиентов телекоммуникационной компании на основе байесовского подхода // Информационные технологии. 2006. №4. - С. 63–67.
- [6] Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. — Springer, 2001.

ПОДХОД К РАЗРАБОТКЕ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ДЛЯ ДОБЫВАЮЩЕГО ПРЕДПРИЯТИЯ НЕФТЕГАЗОВОГО КОМПЛЕКСА

Ю.А. Загорулько, Г.Б. Загорулько, С.В. Булгаков

Институт систем информатики имени А.П. Ершова СО РАН

630090, Новосибирск, пр. Лаврентьева, 6, Россия

zagor@iis.nsk.su

тел: +7 (383) 332-83-59, факс: +7 (383) 332-34-94

Ключевые слова: система поддержки принятия решений, онтология задач, онтология предметной области, мониторинг, диагностика, выработка рекомендаций

Abstract

The paper presents an approach to development of the system providing decision support for decision-maker which has aim to reduce the power inputs of oil-and-gas production enterprise and enhance an environmental safety of oil and natural gas production. Architecture and principles of operation of such system are discussed. Examples of decision support for monitoring and diagnostics of production equipment are given.

Введение

В настоящее время для повышения энергоэффективности и экологической безопасности добывающих предприятий нефтегазового комплекса широко используются информационные технологии [1], в частности АСУ ТП. Но, как правило, эти АСУ ТП установлены на локальных объектах и обслуживают только основной производственный процесс (добычу углеводородного сырья), не затрагивая при этом транспортное обслуживание и проведение плановых и экстренных ремонтов. В связи с этим такие системы не могут обеспечить ЛПР (лицо, принимающее решение) необходимой информацией обо всех процессах, влияющих на эффективность и экологическую безопасность работы предприятия. К тому же эти системы, действуя локально, обеспечивают управление и мониторинг только отдельно взятого объекта или процесса, в то время как для принятия эффективных решений требуется иметь интегральную информацию обо всех объектах и процессах технологической инфраструктуры добывающего предприятия.

Для снижения энергетических затрат добывающих предприятий нефтегазового комплекса и повышения экологической безопасности их работы создается система оперативного мониторинга их технологической инфраструктуры (СОМТИ). Важным компонентом этой системы является аналитическая подсистема (АП СОМТИ), которая должна обеспечить поддержку принятия решений для ЛПР, стремящегося снизить энергозатраты на предприятии и повысить экологическую безопасность процесса добычи углеводородов.

Аналитическая подсистема СОМТИ должна решать следующие задачи:

- Анализ состояния объектов технологической инфраструктуры добывающего предприятия нефтегазового комплекса (далее – просто объектов) с целью предотвращения аварийных ситуаций.
- Анализ состояния объектов с целью улучшения показателей их работы.
- Анализ статической и динамической информации об объектах с целью выработки рекомендаций для ЛПР о необходимости проведения планового технического обслуживания и/или экстренного ремонта объектов, а также о списании объектов и замене их новыми.
- Выработку рекомендаций для ЛПР об оптимизации потоков технологического транспорта и процессов технического обслуживания и ремонтов объектов.

В данном докладе рассматривается архитектура и принципы построения такой подсистемы.

1 Представление знаний в АП СОМТИ

Так как технологическая инфраструктура предприятия может быть подвержена как структурным, так и качественным изменениям, вся система в целом и АП СОМТИ, в частности, должна быть настраиваема на предметную область (ПО) и типы задач. В связи с этим в состав АП СОМТИ в явном виде входит модель ПО, представленная онтологией, а ее архитектура допускает подключение дополнительных модулей, обеспечивающих поддержку принятия решений новых задач.

Онтология системы [2] состоит из двух взаимосвязанных онтологий – онтологии ПО и онтологии задач [3], для описания которых используется язык представления знаний интегрированной системы Semp-ТАО [4, 5].

Онтология ПО описывает модель предметной области в виде понятий и отношений между ними. Базовыми понятиями (классами) онтологии ПО являются: *Объект анализа*, *Нормативно справочный объект*, *Состояние*, *График ремонтов*, *Маршрут*, *Результат*.

В классе *Объект анализа* выделяются подклассы *Оборудование* (насосные установки, трансформаторы, трубопроводы, линии электропередач и т. п.) и *Подвижные объекты* (автоцистерны, грузовики, автобусы и другие виды автотранспорта). Эти объекты могут находиться в том или ином состоянии, для описания которого вводится класс *Состояние*. В свою очередь класс *Состояние* имеет следующие подклассы: *Неисправность*, *Поломка*, *Штатное состояние*, *Предаврийное состояние*.

Для контроля соответствия параметров объектов анализа нормативным значениям вводится понятие *Нормативно справочный объект*. Экземпляры этого понятия создаются для каждого типа объектов (например, для насосов одной марки) и связываются с соответствующими объектами отношением «Описывается».

Каждый объект ПО имеет свой график ремонтов, в котором собрана информация о плановых, текущих и срочных ремонтах данного объекта.

Результатами работы АП СОМТИ являются сообщения о состоянии оборудования и подвижных объектов, рекомендации для ЛППР, решения транспортных задач, скорректированные графики ремонтов и т. п.

Заметим, что онтология ПО не только служит для описания модели ПО, но и определяет формат представления информации, извлекаемой из БД ПХИ, в виде объектов (экземпляров понятий) предметной области. Именно с таким представлением будут работать все модули поддержки принятия решений, включенные в АП СОМТИ.

Онтология задач включает описания решаемых системой задач и модулей принятия решений, реализующих решения этих задач, а также понятия, определяющие настройки и стратегии решения задачи. В онтологии вводятся отношения между задачами и связи задач с модулями принятия решений.

Для описания задачи вводится понятие (класс) *Задача*, имеющий атрибуты «Имя задачи» и «Параметры задачи».

Атрибут «Параметры задачи» представляет собой множество пар вида $\langle Ct, Lt \rangle$, где Ct – класс онтологии ПО или онтологии задач, Lt – список имен объектов класса Ct , для которых должна быть решена данная задача (предполагается, что имя объекта является ключевым атрибутом объекта). Если Lt – не задано (пусто), то считается, что задача решается для всех объектов класса Ct .

Для представления модуля принятия решений в онтологии задач вводится класс *Модуль*, имеющий атрибут «Имя модуля».

На задачах определены отношения «Подзадача» и «Порождает». Первое отношение связывает некоторую задачу с другими задачами (ее подзадачами), решение которых требуется для решения данной задачи. Отношение «Порождает» определяет потенциальную возможность порождения одной задачи другой. Вводится также отношение «Реализует», связывающее модуль принятых решений с задачей, решение которой он обеспечивает. Модуль может связываться только с терминальной задачей.

Все модули принятия решений хранятся в репозитории модулей и для каждого из них определены типы входных и выходных данных, а также решатель.

Описание входных данных задает типы (классы) объектов, необходимых для функционирования модуля. Формально описание выглядит следующим образом: $\langle Cin, Rin, Ain \rangle$, где Cin – множество классов онтологии ПО и онтологии задач, Rin – множество отношений онтологии ПО, заданных на классах Cin , Ain – множество ограничений на значения атрибутов объектов классов из Cin . Ограничение Ain служит для фильтрации объектов, необходимых для решения задачи. Если модуль реализует задачу, у которой есть параметры, то параметры задачи добавляются к множеству ограничений Ain .

Описание выходных данных определяет множество классов, объекты которых могут быть порождены в ходе работы модуля. Его содержимое имеет вид $\langle Cout, Rout \rangle$, где $Cout$ – множество классов онтологии, $Rout$ – множество отношений онтологии, заданных на классах $Cout$.

Для каждого модуля задается имя решателя, т.е. программной системы, которая будет исполнять (интерпретировать) данный модуль принятия решений. В настоящее время в АП используется два решателя: интерпретатор продукционных правил системы Semp-ТАО и решатель задач целочисленного линейного программирования (ЦЛП) GLPK [6]. Для каждого решателя разработан адаптер для обмена данными между ним и локальной памятью АП СОМТИ.

Набор модулей может быть расширен с помощью конфигуратора системы, который позволяет регистрировать новые модули и включать их в репозиторий модулей и онтологию задач.

2 Архитектура и принципы функционирования АП СОМТИ

АП СОМТИ при выработке решений и рекомендаций для ЛПР использует информацию из базы данных подсистемы хранения информации СОМТИ (БД ПХИ СОМТИ), в которой представлены данные о структуре и состоянии технологической инфраструктуры предприятия. В связи с этим АП СОМТИ (см. рисунок 1) реализуется в виде двух взаимодействующих модулей – супервизора и адаптера БД ПХИ СОМТИ, а также набора динамически подключаемых решателей, выполняющих требуемые АП вычисления. Дополнительно АП СОМТИ включает онтологию, репозиторий модулей принятия решений, а также ряд вспомогательных инструментальных средств, обеспечивающих конфигурирование и управление подсистемой.

Супервизор АП СОМТИ является основным модулем, отвечающим за координацию всех элементов подсистемы. Выступая в качестве посредника между подсистемой управления СОМТИ (ПУ СОМТИ), ПХИ СОМТИ и решателями задач, супервизор выполняет обработку полученных заданий, предварительную подготовку входных данных для решателей, сбор и отправку в БД ПХИ СОМТИ результатов вычислений.

Адаптер ПХИ включает компоненты, обеспечивающие получение задач от ПУ СОМТИ, загрузку и выгрузку данных в БД ПХИ СОМТИ, а так же взаимодействие с супервизором.

Основной цикл работы АП СОМТИ состоит в последовательном выполнении заданий, поступающих от подсистемы управления СОМТИ.

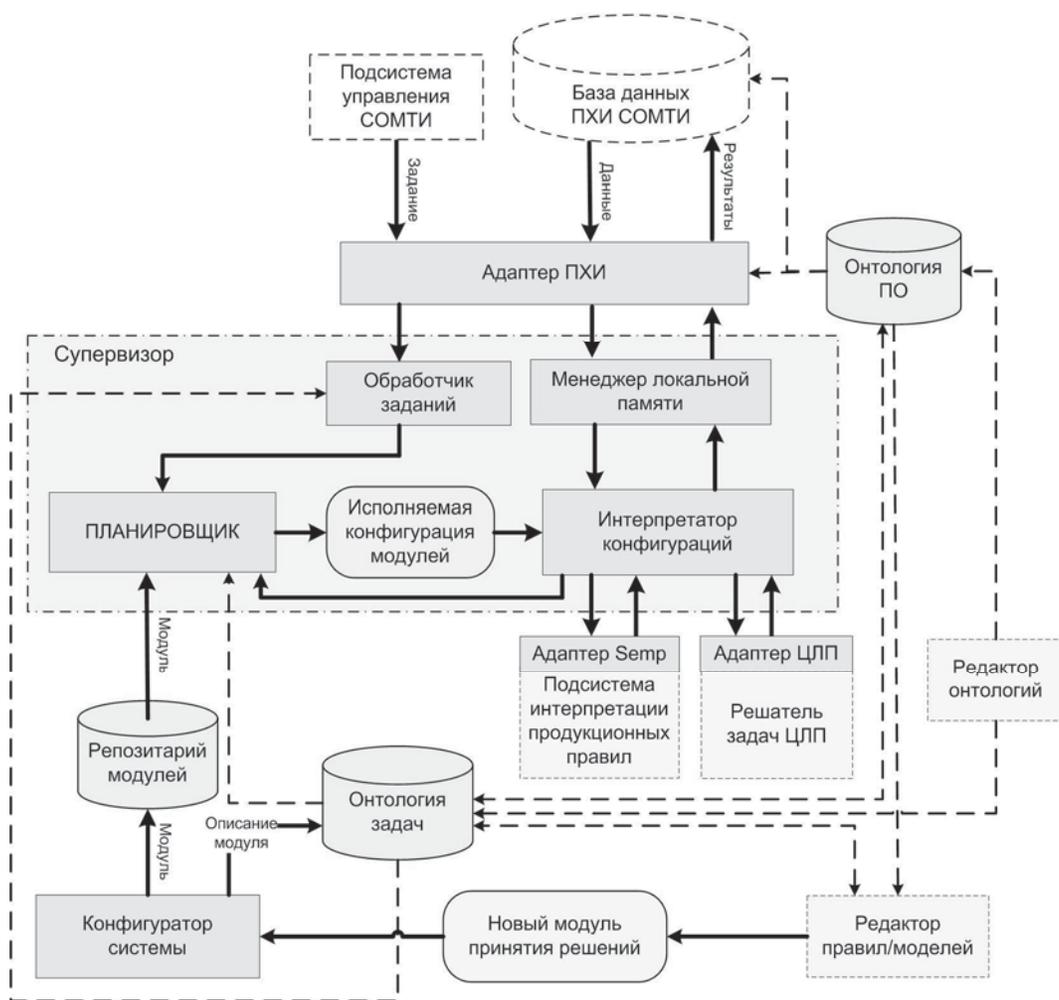


Рисунок 1 – Архитектура АП СОМТИ

Обработчик заданий анализирует задание, извлекает из него очередную задачу, представляет ее в формате, заданном онтологией задач, и передает планировщику.

Планировщик для каждой поступившей на его вход задачи порождает исполняемую конфигурацию модулей, реализующих ее. При этом он использует онтологию задач и репозиторий модулей. Если задача в задании была снабжена параметрами, то они учитываются при сопоставлении данной задачи с задачами, представленными в онтологии задач, и при успешном сопоставлении передаются в модуль, реализующий решение этой задачи.

Цикл работы интерпретатора конфигураций состоит в выборе из конфигурации модулей описания очередного модуля принятий решений, загрузке необходимых для его работы данных из БД ПХИ (через адаптер ПХИ) в локальную память АП СОМТИ, вызове соответствующего решателя и исполнении им выбранного модуля, выгрузке результатов работы модуля в локальную память.

После отработки очередного модуля принятия решений супервизор осуществляет мониторинг локальной памяти и при обнаружении в нем объектов типа «Задача» формирует задание из этих задач и вызывает планировщика. После окончания работы планировщика, который

может дополнить конфигурацию новыми описаниями модулей, интерпретатор конфигураций продолжает работу уже над модифицированной конфигурацией модулей.

3 Поддержка принятия решений в АП СОМТИ

Для решения задач диагностики, мониторинга, выработки управляющих воздействий и рекомендаций используются модули поддержки принятия решений, реализованные в парадигме продукционной модели и для исполнения которых используется интерпретатор системы Semp-ТАО. (Такие модули мы будем называть продукционными.) Продукционные модули включают структурированное множество продукционных правил, описанных в терминах онтологии системы. Для спецификации продукционных правил используется язык системы Semp-ТАО.

В качестве примера рассмотрим работу продукционных модулей, осуществляющих мониторинг некоторого куста скважин и диагностику содержащегося в нем оборудования: автоматизированной групповой замерной установки, трансформаторной подстанции, кустовой насосной станции, насосных установок и транспортных труб. Описания данных объектов представляются в семантической сети. Продукционные правила, работающие над этой сетью, осуществляют анализ таких параметров объектов куста, как дебит скважины, давление на участках трубопровода, энергопотребление и др.

Результатом работы модуля, выполняющего мониторинг куста, может быть сообщение о наличии утечки на определенном участке трубопровода, либо заключение о предполагаемой неисправности того или иного оборудования.

Детальную диагностику оборудования, выделенного первым модулем, и определение его неисправности осуществляет продукционный модуль, предназначенный для диагностики данного типа оборудования. Важную роль в диагностике оборудования играет понятие *Неисправность*. На основе этого понятия для каждого вида (класса) оборудования (например, Насосная установка) вводится класс (Неисправность насосной установки) с теми же атрибутами (параметрами), что и у данного вида оборудования. Значением атрибута «Название» экземпляра этого понятия будет название неисправности (например, «Разрушение рабочего колеса»). Значениями же атрибутов, определяющих технические характеристики конкретного вида оборудования, будут симптомы рассматриваемой неисправности.

В данном контексте, симптомы – это либо предельные значения параметров оборудования, либо лингвистические переменные, описывающие поведение параметров (падение, возрастание, стабильность и т. д.). Для связи оборудования с неисправностями служит понятие Гипотеза. Для каждой конкретной пары оборудования и его неисправности в семантической сети порождается объект класса Гипотеза, первым атрибутом которого является диагностируемый объект, вторым – возможная неисправность, третьим – вероятность того, что данный объект имеет данную неисправность. В процессе проверки гипотезы обнаружение у диагностируемого объекта симптомов неисправности увеличивает значение ее вероятности. Если окажется, что вероятность гипотезы больше некоторого порогового значения, установленного экспертом, то результатом работы продукционного модуля будет сообщение о наличии у диагностируемого объекта данной неисправности.

Поддержка решения задач оптимизации транспортных потоков и ремонтов оборудования осуществляется вычислительными модулями, для исполнения которых служит решатель задач ЦЛП. В используемом в АП решателе GLPK модель решения задачи, входные данные для нее и результаты решения представляются на специальном языке GMPL (GNU MathProg Language). Этот язык позволяет описывать задачу в естественной математической нотации, включающей параметрические уравнения и неравенства, индексированные выражения и т. п., что значительно упрощает составление вычислительных моделей экспертами.

Решение задачи оптимизации транспортных потоков может представлять собой список маршрутов движения автотранспорта, перевозящего нефть, при соблюдении которых будет минимизирован расход горючего при фиксированном объеме собранной нефти.

Заключение

В докладе предложены принципы разработки системы поддержки принятия решений задач повседневной управленческой деятельности для лиц, принимающих решения на нефтегазодобывающем предприятии с целью снижения его энергозатрат и повышения экологической безопасности его деятельности. Ввиду того, что технологическая инфраструктура предприятия может быть подвержена как структурным, так и качественным изменениям, данный подход обеспечивает настройку системы на предметную область и типы задач. Для этого в состав системы в явном виде включена модель предметной области, представленная онтологией, а ее архитектура допускает подключение дополнительных модулей, обеспечивающих поддержку принятия решений при решении новых задач.

В настоящее время реализована пилотная версия системы АП СОМТИ, обеспечивающая поддержку принятия решений таких задач управленческой деятельности, как мониторинг состояния объектов технологической инфраструктуры предприятия с целью своевременного вывода неисправного оборудования из эксплуатации или постановке его на ремонт, оптимизация ремонтных работ и транспортных потоков.

Благодарности

Работа выполняется в рамках Государственного контракта № 02.514.11.4126 от 30.09.2009 г. с Федеральным агентством по науке и инновациям в рамках федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2012 годы» научно-исследовательские работы по теме: «Разработка интеллектуальной системы пространственно-технологического мониторинга на базе глобального спутникового позиционирования с целью повышения энергоэффективности и экологической безопасности существующих методов добычи углеводородов» (шифр заявки «2009-04-1.4-15-13-011»).

Список литературы

- [1] Байков И.Р., Самородов Е.А., Ахмадуллин К.Р. Методы анализа надежности и эффективности систем добычи и транспорта углеводородного сырья. – М.: ООО «Недра-Бизнесцентр», 2003. – 275 с.
- [2] Guarino N. Formal Ontology in Information Systems // Proc. of FOIS'98, Trento, Italy, 6-8 June 1998. – Amsterdam: IOS Press, 1998. – P. 3–15.
- [3] Добров Б.В., Иванов В.В., Лукашевич Н.В., Соловьев В.Д. Онтология и тезаурусы. – Казань: Казанский гос. ун-т, 2006. – 198 с.
- [4] Загорюлько Ю.А., Попов И.Г., Щипунов В.В. Интегрированная технологическая среда для создания систем обработки знаний // Известия РАН. Теория и системы управления. –1995. – № 5. – С. 210–213.
- [5] Загорюлько Ю.А., Попов И.Г. Представление знаний в интегрированной технологической среде Semp-ТАО // Проблемы представления и обработки не полностью определенных знаний / под ред. И.Е. Швецова. – М.-Новосибирск, 1996. – С. 59–74.
- [6] GLPK (GNU Linear Programming Kit). Дата обновления: 2008.10.16. URL: <http://www.gnu.org/software/glpk/glpk.html> (дата обращения: 16.03.2010).

РАСШИРЕНИЕ ПОДСИСТЕМЫ ВИЗУАЛИЗАЦИИ НАПОЛНЕНИЯ ИНФОРМАЦИОННОГО ПОРТАЛА СРЕДСТВАМИ ВИЗУАЛЬНОЙ АНАЛИТИКИ*

З.В. Апанович¹, Т.А. Кислицина²

¹Институт систем информатики им. А.П. Ершова СО РАН
630090, Новосибирск, проспект Лаврентьева, 6, Россия
apanovich@iis.nsk.su

тел: +7 (383) 330-93-44, факс: +7 (383) 332-34-94

²Новосибирский государственный университет,
630090, Новосибирск, ул. Пирогова, 2, Россия

Ключевые слова: портал знаний, онтология, информационное наполнение, анализ информации, методы визуализации информации, силовой алгоритм, радиальный алгоритм, методы кластеризации графов, иерархическое связывание ребер

Abstract

The process of development of an ontology-based knowledge portal and creation of its content is time-consuming and labor-intensive. The lifetime of such portals is sufficiently long and they collect plenty of valuable information. This information can be analyzed from various points of view. A subsystem for visual analysis of content of ontology-based knowledge portals has been developed in the A.P. Ershov Institute of Informatics Systems. This paper describes an extension of this visualization subsystem with visual analytics tools. An example of information extraction about scientific cooperation from the content of a knowledge portal is presented and utilization of visual analysis tools to this scientometric information is demonstrated.

Введение

Общепризнанным инструментом, обеспечивающим понимание больших объемов абстрактной информации, является визуализация информации с применением графовых моделей. Онтология и наполнение информационного портала могут быть представлены в виде графа, вершины которого изображают сущности, такие как классы, объекты и атрибуты онтологии, а ребра изображают отношения между этими сущностями. Просматривая изображение различных подграфов, генерируемых по описанию данной онтологии и информационного наполнения, разработчик системы может обнаружить как ошибки данных, возникшие при ручном вводе информации, так и ошибки проектирования, незаметные при просмотре текстовых форм.

По мере своего развития, информационные порталы накапливают все больше разнородной информации и становятся бесценным источником информации, пригодной для всевозможных форм анализа. При условии регулярного пополнения такого портала, он становится базой данных, анализ которой может иметь самостоятельный научный интерес. Поэтому возникает задача анализа развивающегося информационного наполнения с целью выявления различных закономерностей. Так, например, в случае информационного портала, посвященного некоторому научному направлению, весьма актуальной является задача наукометрического анализа его наполнения, в частности, выявления шаблонов сотрудничества между различными научными сообществами.

* Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 09-07-00400) и проекта РАН 2/12 Формальные языки и методы спецификации, анализа и синтеза информационных систем.

В данной работе демонстрируются новые возможности подсистемы визуализации информационных порталов, разработанной в ИСИ СО РАН [1, 2]. Эти методы визуализации ориентированы на более глубокий анализ его информационного наполнения. В качестве тестовых данных использовалось информационное наполнение специализированного археологического портала знаний [3], построенного на основе системной классификации археологической науки [4], предложенной Ю.П. Холушкиным и Е.Д. Гражданниковым. В этой классификации определено множество понятий археологии, которые организованы в набор упорядоченных в иерархию фрагментов универсальной структуры, фиксирующих явные и неявные связи между понятиями. Археологический портал знаний продолжает активно пополняться новыми знаниями и данными и на данный момент содержит более 4 тысяч информационных объектов, связанных более чем 15 тысячами онтологических отношений.

1 Генерация и визуальный анализ сетей научного сотрудничества на основе отношения соавторства в публикациях

Археологический портал содержит, в частности, много информации о публикациях исследователей, работающих в разных областях археологии. Помимо того, что эта информация представляет несомненный интерес для специалистов, работающих в области археологии, она может быть использована и для всевозможных наукометрических экспериментов. Одним из распространенных методов анализа библиографической информации является анализ сетей сотрудничества на основе отношения соавторства. Следует заметить, что исследование сетей соавторства в настоящее время весьма популярной научной задачей. На смену изолированным исследователям приходят меж-дисциплинарные команды с различными организационными структурами, выполняющие совместные научные проекты. Поскольку совместные исследования стали доминирующим и наиболее многообещающим способом производства высококачественного научного результата, структуры сотрудничества становятся тоже предметом исследования, прежде всего, в целях создания новых структур управления. Факт наличия совместных научных проектов, совместных публикации, и много-дисциплинарных сетей сотрудничества указывает на современный способ производства научного знания [5, 6]. Сети соавторства изучаются с таких точек зрения, как распределение степеней [7], извлечение научных сообществ [8], предсказание будущих партнерских отношений [9, 10], ранжирование социальных сущностей [11] и др.

Для анализа сети соавторства, необходимо сгенерировать граф, в котором каждый автор изображается вершиной, а ребро соответствует факту наличия совместной публикации между двумя авторами. В случае, если публикация имеет n авторов, создается клика, соединяющая всех авторов одной публикации. Поскольку авторы могут иметь несколько совместных публикаций, каждое ребро, соединяющее научных исследователей, имеет вес, равный количеству совместных публикаций. К сожалению, в тестовой онтологии, с которой осуществлялись эксперименты, отношения соавторства в явном виде не существовало. В то же время, отношение соавторства, как и многие другие отношения, могут быть описаны при помощи суперпозиции отношений, уже имеющихся в онтологии. Например, в онтологии имеется отношение «Автор-Публикация», связывающее классы «Исследователь» и класс «Публикация». Ясно, что отношение соавторства может быть описано как композиция отношения «Автор-Публикация» и его инверсии. Поэтому первым шагом нашей работы было встраивание во внутренний язык подсистемы визуализации средств, позволяющих описывать суперпозицию отношений. Это позволило нам построить граф соавторства на основе информации о научных публикациях и их авторах, имеющейся в наполнении портала.

2 Применение кластеризации для визуализации и анализа отношения сотрудничества

Как правило, анализ сетей соавторства начинается с извлечения научных сообществ. Следует сказать, что стандартные силовые алгоритмы мало пригодны для визуализации научных сообществ, поскольку они стремятся разместить все вершины графа на одинаковом «идеальном» расстоянии. Нам нужен алгоритм, который позволял бы визуально выделять группы исследователей, сильно связанных между собой, и слабо связанных с другими группами. Стандартным способом выделения научных сообществ являются различные методы кластеризации, позволяющие выделять группы исследователей, сильно связанных между собой, и слабо связанных с другими группами. Обзор, посвященный методам кластеризации и их применению для визуализации графов, можно найти в работе [12]. Мы использовали в качестве алгоритма кластеризации адаптацию алгоритма кластеризующего размещения LinLog [13]. Эксперименты осуществлялись с моделью вершинного отталкивания и с моделью реберного отталкивания.

Энергия вершинного отталкивания вершин $U_{\text{nodeLinLog}}$ для оптимизации размещения p определяется как:

$$U_{\text{NodeLinLog}}(p) = \sum_{\{u,v\} \in E} \|p(u) - p(v)\| - \sum_{\{u,v\} \in V^{(2)}} \ln \|p(u) - p(v)\|.$$

Первый терм в определении энергии $U_{\text{nodeLinLog}}$ можно интерпретировать как притяжение между смежными вершинами, а второй терм – как отталкивание между различными вершинами. В этой модели позиции вершин соответствуют прежде всего степеням вершин – вершины высокой степени располагаются в центре, а вершины с низкой степенью – на периферии.

В модели реберного отталкивания отталкивание между вершинами заменено на отталкивание между ребрами. Считается, что отталкивание действует не между всеми ребрами, а только между их концевыми вершинами. Поэтому отталкивание между двумя вершинами имеет вес, равный количеству ребер, у которых они являются концевыми вершинами (то есть, степени вершин). Энергия реберного отталкивания вершин $U_{\text{EdgeLinLog}}$ имеет вид:

$$U_{\text{NodeLinLog}}(p) = \sum_{\{u,v\} \in E} \|p(u) - p(v)\| - \sum_{\{u,v\} \in V^{(2)}} \ln \|p(u) - p(v)\|.$$

Такая функция энергии позволяет вершинам небольшой степени быть расположенными в центре, если они сильно связаны с другими вершинами. На рисунке 1 показан пример применения функции LinLog для выявления исследователей, сильно связанных отношениями соавторства. На этом изображении окружности соответствуют отдельным исследователям, размер вершин – количеству научных публикаций (что позволяет легко идентифицировать наиболее плодотворно работающих ученых). В частности, большая окружность розового цвета изображает директора института, большая окружность красного цвета – его заместителя, а самая большая окружность синего цвета – автора информационного портала. Цвета различных вершин соответствуют кластерам, вычисленным на основании принципа модулярности [8].

В целом, такое изображение позволяет составить общее представление о группах людей, связанных отношениями партнерства и группах, такими отношениями не обладающих. Для нас индикатором приемлемого качества кластеризации послужил тот факт, что все сотрудники ИСИ СО РАН автоматически попали в один кластер с сотрудниками сектора археологической теории и информатики ИАЭТ СО РАН, с которыми их связывает несколько совместных проектов. Хотя внимательно изучив изображение, можно заметить, что имеются кластеры, вершины которых при размещении оказались разбросанными достаточно далеко друг от друга (например, вершины, принадлежащие кластеру коричневого цвета). С нашей точки зрения, основным недостатком такой визуализации является ее поверхностный характер. Ведь специализированные информационные порталы содержат, помимо библиографической, огромное

количество дополнительной информации: о научных организациях и разделах науки, в которых работают исследователи, о предметах исследования и различных видах научной деятельности, таких, например, как научные экспедиции и т.д. Поэтому весьма интересной видится задача сопоставления отношения сотрудничества с другими подструктурами, которые можно выделить в информационном наполнении портала. Возможность такого сопоставления реализована при помощи метода иерархических пучков ребер [14], который будет описан в следующем разделе.

Class: Исследователь
 id: 227
 Источник информации (URL):
 пол: мужской
 Отчество: Тимофеевич
 Имя: Василий
 Ученое звание:
 Фамилия: Воронин
 Ученая степень: кандидат
 экономическ. наук
 e-mail: bas@sabi-gw.
 archaeology.nsc.ru
 Координаты:
 p. t.: +7383302733

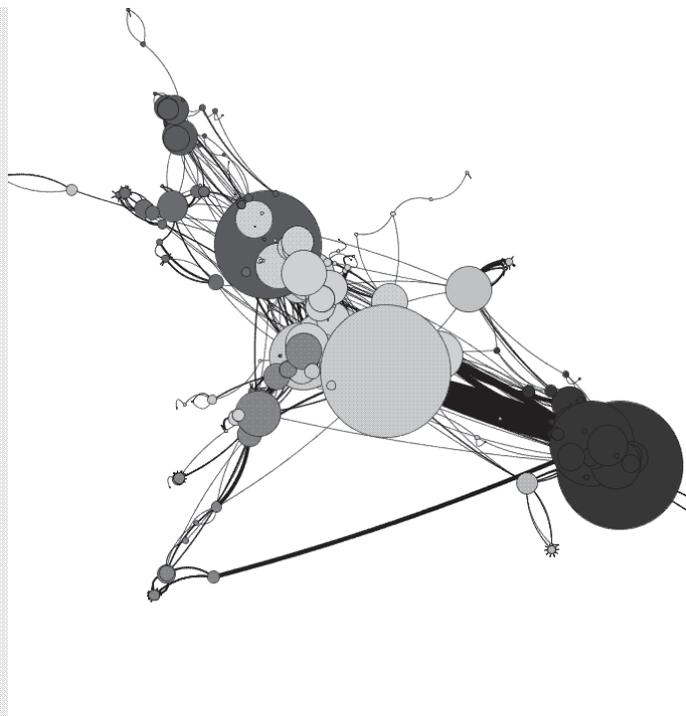


Рисунок 1 – Кластеризация сети соавторства, полученная на основе модели $U_{EdgeLinLog}$

3 Применение иерархического связывания ребер для визуализации и анализа отношения сотрудничества между исследователями

Метод иерархического связывания ребер позволяет извлекать из онтологии произвольное отношение, например, дерево связей разделов науки или же информацию об организационных структурах, в которых работают научные сотрудники, и накладывать на это изображение информацию о сетях соавторства. Для отображения иерархической структуры можно использовать любой алгоритм визуализации деревьев, а для неиерархических связей используется метод иерархического связывания ребер. Этот метод изображает каждое заданное ребро в виде кубического β -сплайна (кривой). Его основные преимущества заключаются в следующем:

- Может быть использован в сочетании с технологиями визуализации деревьев, и легко интегрирован в существующие программные средства.
- Уменьшает визуальную перегруженность изображения графа с большим количеством смежных ребер за счет того, что кривые объединяются в визуальные «жгуты».
- Можно регулировать степень кривизны кубического β -сплайна, получая различную информацию: о связи между объектами низкого уровня (между двумя концами ребра) и о связи вершин с объектами верхнего уровня.

Для сопоставления отношения сотрудничества с другой информацией, доступной в информационном наполнении портала, была реализована следующая стратегия. Отношения, связь с которыми мы хотим исследовать, изображаются в виде дерева. Их изображение строится либо при помощи радиального, либо кругового алгоритма. Затем координаты вершин построенного дерева используются в качестве контрольных точек, через которые проводится каждое ребро отношения сотрудничества. Таким образом, каждое ребро отношения сотрудничества представляет собой многоугольный путь по дереву, описываемый последовательностью точек P_0, P_1, \dots, P_{N-1} . Зная эти контрольные точки, можно построить кривую, изображающую ребро графа сотрудничества по формуле:

$$P' = \beta \cdot P_i + (1 - \beta) \left(P_0 + \frac{i}{N-i} (P_{N-1} - P_0) \right), \text{ где}$$

N – это количество контрольных точек,

i – номер контрольной точки, $i \in \{0, \dots, N-1\}$,

P_i – позиция i -ой контрольной точки

β – сила связывания, $\beta \in [0, 1]$.

Тестовая онтология информационного портала содержит значительное количество отношений, которые задают различные иерархии, пригодные для изображения при помощи дерева. К таким отношениям относятся «Раздел-Науки-Включает», «Организация_Включает», «Метод_Исследования-Включает» и другие. Данная методика позволяет исследовать отношение сотрудничество совместно с любым из этих отношений.

Рассмотрим, например, небольшую по размеру иерархию научных организаций, показанную на рисунке 2. Она изображена при помощи кругового алгоритма. Маленькие черные окружности изображают научные организации, маленькие рыжие окружности – отдельных сотрудников организаций, зеленым цветом изображены ребра, показывающие принадлежность отдельных сотрудников научным организациям, желтым цветом показаны ребра, соответствующие отношению «Организация-Включает».

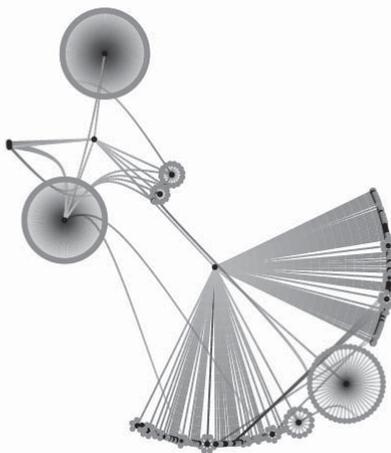


Рисунок 2 – Изображение иерархии «Организация–Включает» при помощи кругового алгоритма

Можно заметить, что существуют сотрудники, работающие одновременно в нескольких организациях. Как правило, вторая организация – это образовательное учреждение (университет).

Теперь рассмотрим эту же самую иерархию, но совместно с отношением сотрудничества, как это показано на рисунке 3. Широкие светлые жгуты показывают все имеющиеся связи на-

учного сотрудничества между изображенными организациями, а более тонкие жгуты темного цвета – связи ИСИ СО РАН с сотрудниками института Археологии и Этнографии СО РАН. Имеется также возможность проследить и проанализировать связи сотрудничества каждого отдельного сотрудника.

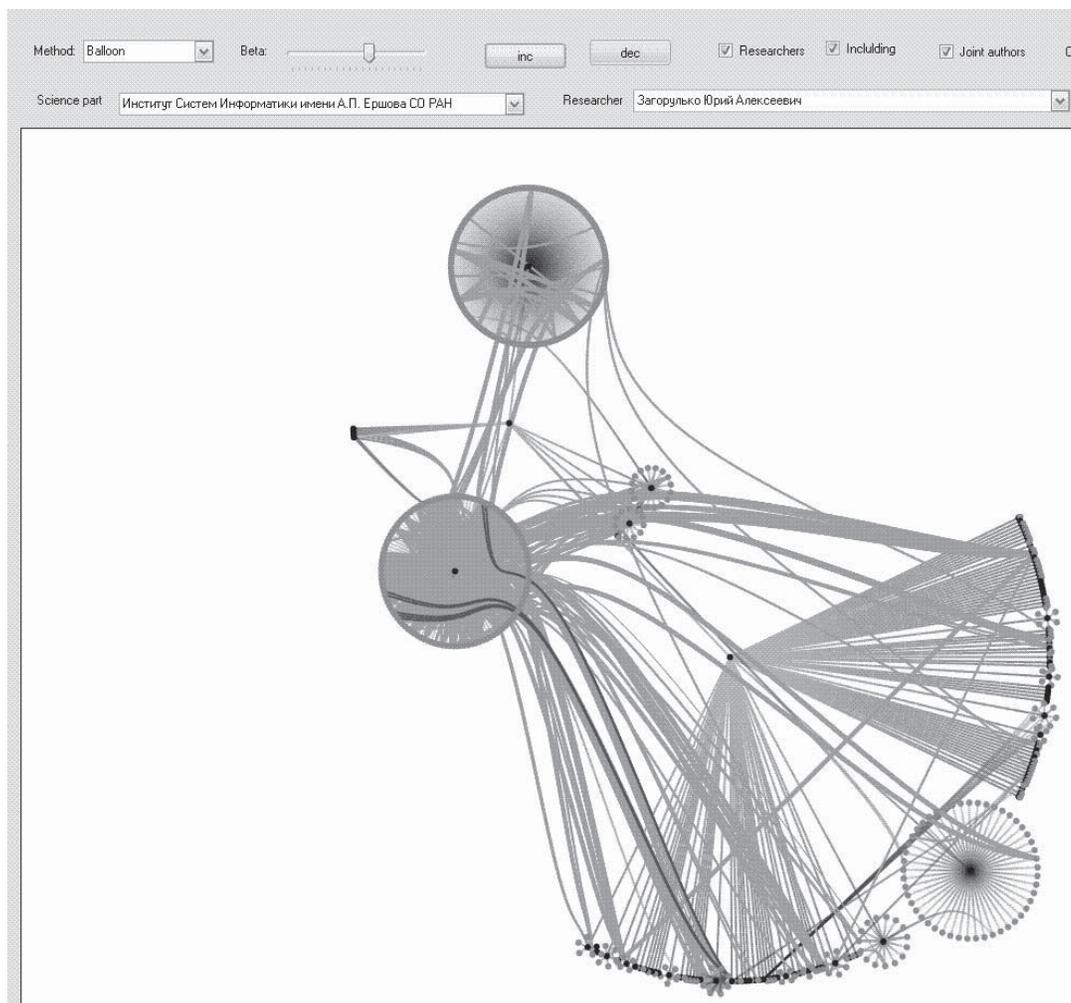


Рисунок 3 – Связи научного сотрудничества в сочетании с изображением научных организаций

Наконец, на рисунке 4 можно проследить связи внутриинститутские связи отдельного исследователя. Достаточно хорошо просматривается корреляция количества отношения сотрудничества отдельных исследователей с занимаемой должностью и ученой степенью.

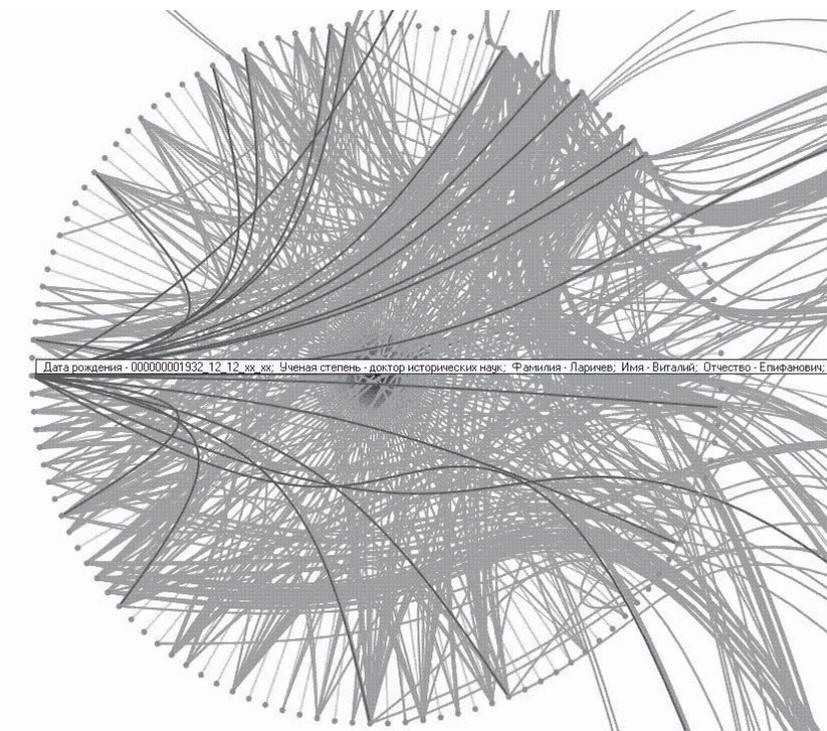


Рисунок 4 – Структура научного сотрудничества между исследователями, работающими в одном институте

Заключение

В данной работе были рассмотрены первые результаты по применению средств визуальной аналитики для анализа информационного наполнения порталов знаний. В ближайшее время планируется развитие этих исследований по нескольким направлениям. Прежде всего, предполагается расширение множества методов кластеризованного размещения новыми методами, учитывающие различные характеристики центральности, в частности, центральность промежуточности. Далее мы собираемся рассмотреть изображения, строящие иерархию на основе одного из методов кластеризации в сочетании с иерархическим связыванием ребер. Наконец, планируется развитие входного языка подсистемы визуализации в направлении максимального приближения к языку OWL.

Благодарности

Авторы выражают благодарность Ю.А. Загорулько и С.В. Булгакову за предоставленные тестовые данные в xml-формате.

Список литературы

- [1] Apanovich Z. V., Vinokurov P. S., Elagin V. A. An approach to visualization of knowledge portal content// Bulletin of NCC . – Issue 29. – 2009. – P.17-32.
- [2] Апанович З.В. Методы навигации при визуализации графов// Вестник НГУ, Том 6, выпуск 3. – 2008. – С. 35-47 .

- [3] Загорюлько Ю.А., Боровикова О.И., Холюшкин Ю.П. Построение предметной онтологии для археологического портала научных знаний//Информационные технологии в гуманитарных исследованиях. Вып 10. – Новосибирск, 2006.
- [4] Холюшкин Ю.П., Гражданников Е.Д. Системная классификация археологической науки (элементарное введение в науковедение) Новосибирск. – Новосибирск. – 2000.
- [5] Wuchty, S. Jones, B. Uzzi, B. The Increasing Dominance of Teams in Production of Knowledge.// *Science Express*. – Vol. 316, N. 5827. – 2007. – P. 1036 – 1039.
- [6] Jones, B. F. Wuchty, S, Uzzi, B. Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science// *Science* 322 (5905). – 21 November 2008. – 1259.
- [7] Barabasi, A.-L. The origin of bursts and heavy tails in human dynamics// *Nature*. – N 435. – 2005. – P. 207-211.
- [8] Newman M. E. J., Girvan M. Finding and evaluating community structure in networks// *Physical Review E*, 69. – 26113. – 2004.
- [9] Getoor L., Diehl C. P. Link mining: a survey// *ACM SIGKDD Explorations Newsletter*, N7. – 2005. – P. 3-12.
- [10] Huang J. et al. Collaboration over time: characterizing and modeling network evolution// *Proceedings of the international conference on Web Search and Web Data Mining*. – 2008. – P. 107-116.
- [11] Zhou D. et al. Co-ranking authors and documents in a heterogeneous network. // *Proc ICDM'07*. – USA 2007. – P. 734-744.
- [12] Herman I., Melancon G., Marshall M. S. Graph visualization and navigation in information visualization: A survey// *IEEE Transactions on Visualization and Computer Graphics*. – Vol. 6, N1. – 2000. – P 24-43.
- [13] Noack A. Energy Models for Graph Clustering//*Journal of Graph Algorithms and Applications*. – vol. 11, N. 2. – 2007. – P. 453-480. 2007. – <http://jgaa.info/>
- [14] Holten D. Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data//*Transactions on Visualization and Computer Graphics*. – vol. 12, N5. – 2006. – P. 741-748.

АДЕКВАТНАЯ РЕАЛИЗАЦИЯ КОГНИТИВНОГО ПОТЕНЦИАЛА ОБОЗРЕВАТЕЛЯ ОНТОЛОГИЧЕСКИХ МОДЕЛЕЙ

С.В. Смирнов, Е.Г. Суворова

Институт проблем управления сложными системами РАН

443020, Самара, ул. Садовая, 61, Россия

smirnov@iccs.ru

тел: +7 (846) 332-39-27, факс: +7 (846) 333-27-70

Ключевые слова: онтологическая модель, семантическая сеть, интерактивная визуализация информации, метод «фокус + контекст», 3D-организация и анимация изображения

Abstract

The purpose of this article is the enhancement of visual interface of the Ontological models browser – a component of “gB-system” modeling toolset, being developed in ICCS of RAS. Models visualization is based on the synthesis of radial, central-radial and complex “Fisheye” views. The article provides an overview of a set of not trivial geometrical and analytical problems, which were to be resolved for building the view diagram, with would utilize in full the cognitive potential of the proposed interface.

Введение

В [1], а затем наряду с описанием прикладных задач управления в статьях [2, 3] был представлен *новый графический интерфейс* одного из интерактивных программных приложений общецелевой системы объектно-ориентированного моделирования gB [4]. Это приложение – «*Мастер моделей*», реализует функции *обозрения и редактирования* объектных денотативных моделей предметных областей (ПрО).

Уникальность Мастера моделей состоит в том, что он, вообще говоря, *покрывает* функциональность всех других специализированных приложений gB-системы [4] (например, может использоваться для конструирования *онтологий* как денотативных моделей ПрО, описываемой *метаонтологией* [5]). Однако реализация этой функциональности *в части редактирования* моделей возможна лишь *универсальными «низкоуровневыми» средствами* gB-системы – фактически интерфейсными методами объектно-ориентированной СУБД. Вследствие этого семантическая правильность, или «целостность», создаваемых объектных моделей практически целиком определится корректностью действий пользователя: в данном случае у него отсутствуют посредники в виде *специализированных компонентов* тех приложений gB-системы, которые созданы для работы в определенных *контекстах моделирования* [4, 5] и поэтому снабжены адекватным инструментарием и процедурами семантического контроля действий пользователя в пределах этих контекстов.

Ремарка о специфике работы Мастера моделей оправдана тем, что поясняет целесообразность выделения функции *обозрения моделей* в gB-системе и ее автономную реализацию специализированным приложением «*Обозреватель моделей*». При этом принципиальным является тождество возможностей *Обозревателя моделей* и *Мастера моделей* в визуализации информации и навигации по информационному пространству моделей¹. Таким образом, фактически в данной статье речь пойдет о пользовательском интерфейсе обоих названных gB-приложений.

¹ Подобная практика разработки облегченных, и в частности, выпуск исключительно «читающих» (reader-) приложений наряду с базовыми программными процессорами (текстовыми, графическими и т.п.) весьма характерна для рынка сложных программных продуктов.

Разработка графического интерфейса для обозрения онтологических моделей относится к обширной, насыщенной разнообразной проблематикой сфере создания методов и средств интерактивной визуализации информации и когнитивной графики. Оптимистической предпосылкой здесь служит то обстоятельство, что онтологии и, вообще, онтологические модели с методической точки зрения – наиболее «систематический» способ структурирования информационного контента с возможностью наглядного представления в виде *семантической сети*, т.е., в конечном счете, в виде *графа* [1, 6]. Применительно же к графам имеется громадный задел методов интерактивной визуализации, а равно и примеров соответствующих наукоемких программных продуктов [7, 8]; появляются и разработки, прямо ориентированные на визуализацию онтологий [9, 10]. Новизны и практической полезности результатов самостоятельных проектов в этой области можно ожидать и в продуктивном сочетании ранее известных подходов. Критерии такой продуктивности разнообразны и специфичны [8], но, несомненно, для онтологических моделей особую важность имеет уровень реализации *когнитивного потенциала* предлагаемых методов и средств интерактивной визуализации информации [6].

В этом смысле представленный в [1-3] графический интерфейс лишь до некоторой степени реализовывал вполне оригинальную идею трехмерной (3D-) организации представления онтологических моделей с элементами анимации: в частности, в определенных ситуациях 3D-образы соединялись с двумерным представлением элементов модели. Возникающий в подобных случаях геометрический диссонанс существенно дискредитировал роль когнитивно-ориентированных элементов изображения, «размывал» внедряемые в изображение аллюзии, затруднял интегральное восприятие доставляемого пользователю информационного образа (сопоставление рисунков 1а и 1б позволяет в какой-то мере ощутить влияние обсуждаемого эффекта). Эти обстоятельства и полученный опыт применения инструментальных средств моделирования [2, 3] выдвинули в линию актуальных проблему адекватной реализации когнитивного потенциала предложенного способа интерактивной визуализации онтологических моделей в *gB*-системе.

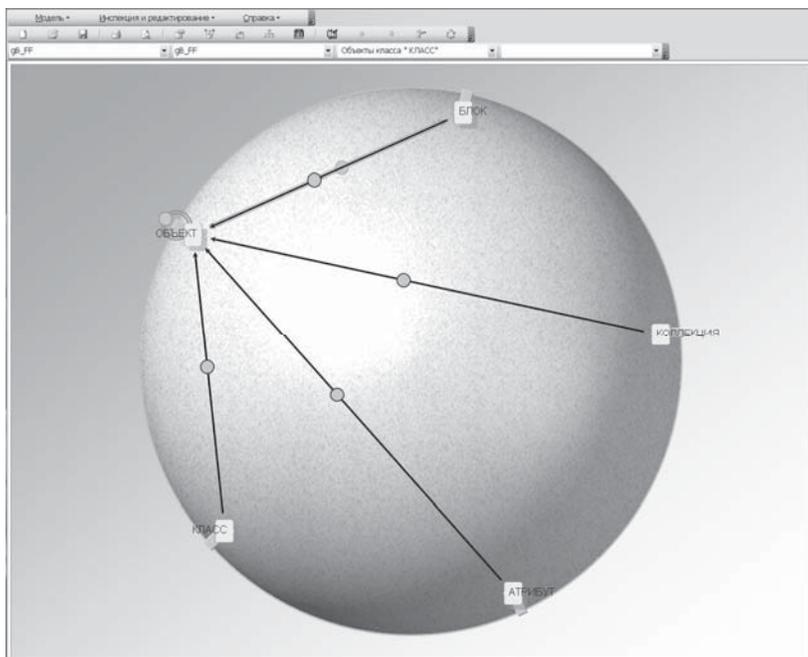
В предлагаемом сообщении формулируются основные принципы обозрения онтологических моделей с помощью разработанного графического интерфейса, а также дается обзор геометрических и аналитических задач, решение которых служит основой визуализации моделей.

1 Принципы организации обозрения моделей

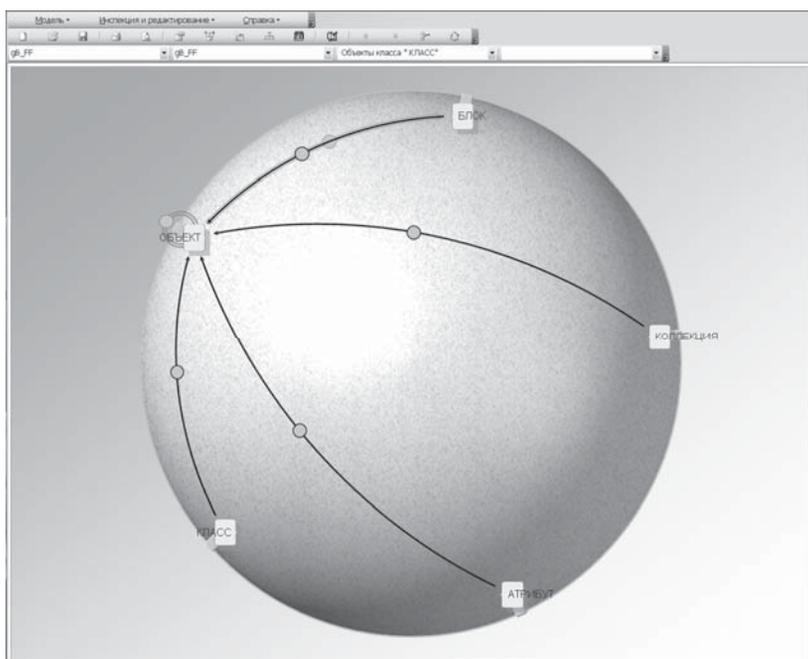
Знакомство с обширной литературой, представляющей методы и инструментальные средства визуализации информации в форме графов, многочисленными Интернет-ресурсами подобной направленности² и следование (ставшими уже традиционными при разработке *gB*-системы) ценностям *минимализма, однородности теоретических и технологических конструкций* привело к выбору следующих взаимодополняющих принципов организации обозрения онтологических моделей:

- опора на *блочную организацию* модели как простейшего и эффективного метода ее иерархической структуризации, которая, вообще говоря, может вводиться *независимо* от семантики модели (принцип «работает» на представление *общих планов* и визуализацию *больших моделей*) [12];
- использование *кругового и центрально-радиального* изображений графа соответственно для обозрения коллекций объектов общего вида и коллекций объектов-смежников некоторого избранного объекта модели (в целом это *предельно упрощает* отображение данных на геометрическую плоскость, а центрально-радиальное изображение практически является интуитивно понятным и удобным *порталом* к предоставлению по требованию самых «мелких» деталей модели – свойств отдельных объектов);

² Например, сайт [11] дает возможность что называется «вживую» ознакомиться с быстро растущим множеством идей в этой области.



а)



б)

Рисунок 1 – Пример представления онтологической модели с разной степенью качества:
 а) соединение в обозрении трех- и двумерных (для дуг семантической сети) образов;
 б) согласование «размерности» изображений различных компонентов обозрения

- использование вершин-псевдообъектов и дуг-псевдосвязей (принцип, служащий «бритвой Оккама» для уменьшения разновидностей образов, которые формируют диаграмму обозрения, и действующий в задачах представления *больших моделей*, а также при поддержке работы с несколькими моделями одновременно);
- следование идее методов «фокус-контекст» (известных также как различные реализации концепции экранного представления «рыбий глаз») [8], предполагающей совмещение на изображении деталей актуального фрагмента модели с условным, составным образом остальной фактически невидимой части модели (это, прежде всего, обеспечивает *взаимодействие* пользователя с большими моделями).

Синтез этих позиций с приматом использования средств прямого манипулирования элементами изображения (в частности, устройства «мышь») и позволил предложить новый конкурентоспособный вариант системы интерактивной визуализации онтологических моделей.

В целом пространство обозрения модели организует 3D-фигура «Инфоглобус», на поверхности которого размещаются вершины-объекты модели и соединяющие их дуги-связи (см. рисунок 1). В отличие от известных способов использования сферической поверхности для размещения вершин визуализируемого графа Инфоглобус, во-первых, *непрозрачен* и, во-вторых, обращен к наблюдателю *всегда одной стороной* (сравните с Луной).

Предлагаемое для интерактивной визуализации развитие метафоры Инфоглобуса состоит в поддержании *иллюзии постоянного присутствия на экране всей обозреваемой модели* (в виде графа, вершины которого, представляют совокупность неоднородных объектов моделируемой ПрО, а дуги - множество различных отношений, связывающих объекты ПрО):

- на *видимой стороне* Инфоглобуса размещается тем или иным способом указанный пользователем *актуальный фрагмент модели* (иначе говоря, фокус обозрения), изображение которого в целом имеет круговую или центрально-радиальную организацию;
- все компоненты модели, не вошедшие в состав актуального фрагмента, располагаются на *обратной, невидимой стороне* Инфоглобуса;
- имеющиеся *связи видимого объекта с невидимыми объектами*³ показываются специальной интегрированной псевдосвязью - *коннектором*, образ которой - неориентированный отрезок прямой, исходит из образа видимого объекта и «скрывается за горизонтом» Инфоглобуса;
- смена актуальных фрагментов модели производится путем *анимированного перемещения объектов* по видимой поверхности Инфоглобуса с «восходом» актуализируемых объектов из-за линии горизонта и «заходом» неактуальных объектов за линию горизонта Инфоглобуса;
- указания пользователя о *смене актуального фрагмента* модели осуществляется как с помощью выбора в предоставляемых меню, так и прямым формированием состава этого фрагмента с использованием пальцев руки и сенсорного экрана, применения манипулятора «мышь» и т.п.

В целом в рамках теоретической концепции «фокус-контекст» именно Инфоглобус и совокупность коннекторов образуют необходимо присутствующий при визуализации *составной образ невидимой части обозреваемой модели*.

2 Геометрические и аналитические задачи, решение которых служит основой визуализации моделей

Задачи, решение которых составляет основу усовершенствованного интерфейса Обозревателя моделей gB-системы, условно можно сгруппировать вокруг двух главных проблем:

³ Для центрально-радиальных диаграмм речь может идти и о связях и между видимыми объектами-смежниками «обозреваемого объекта», располагающегося в центре изображения.

- 1) построение образов видимых связей между объектами модели в виде дуг, проводимых на поверхности Инфоглобуса;
- 2) построение образов видимых объектов – плоских фигур различной формы, размещаемых в различных точках поверхности Инфоглобуса с имитацией различной удаленности этих фигур от наблюдателя.

Анализ этих проблем и решение возникающих задач выполнялись при следующих предположениях и ограничениях:

- Инфоглобус находится за плоскостью экрана, и фактическая глубина конструируемых 3D-изображений равна радиусу Инфоглобуса;
- наблюдение формируемых на экране изображений осуществляется из точки, расположенной перед плоскостью экрана и принадлежит оси, которая перпендикулярна экранной поверхности и проходит через центр Инфоглобуса.

2.1 Образы связей между объектами

При сделанных предположениях задачи построения образов связей при круговом и центрально-радиальном изображениях обозреваемого фрагмента модели радикально отличаются.

Несколько идеализируя картину можно сказать, что в первом случае образ связи двух видимых объектов – это *кривая*, соединяющая на экране две точки кругового горизонта Инфоглобуса, где точки суть названные объекты (см. рисунок 1б). Во втором случае образами видимых связей будут отрезки *прямых*, соединяющие ближайшую к наблюдателю точку Инфоглобуса, которая идеализирует «*обозреваемый объект*», с точками на круговом горизонте Инфоглобуса, которые представляют *объекты-смежники* обозреваемого объекта (рисунок 2).

Хотя и в том, и в другом случаях каждый образ связи есть проекция отрезка линии, кратчайшим путем соединяющей две точки сферы и проходящей по ее поверхности, ясно, что никаких принципиальных затруднений с построением центрально-радиальной диаграммы не имеется. Трудности возникают с изображением связей на круговой диаграмме, когда линейная аппроксимация проекции 3D-отрезка очевидным образом не работает (см. рисунок 1а). Найти более адекватное решение обнаруженной проблемы оказывается важным еще и потому, что оно лежит в основе построения многих других элементов обновленного интерфейса Обозревателя моделей.

Анализ проблемы позволил выдвинуть гипотезу о возможности представления искомого адекватного образа связи *дугой окружности*, параметры которой связаны уравнением, определяющим в первой четверти прямоугольных координат одну из замечательных плоских кривых - *трактрису* [13]:

$$r = R \left(1 + \ln \frac{R + \sqrt{R^2 - h^2}}{h} \right) - \sqrt{R^2 - h^2},$$

где R – неотрицательный параметр трактрисы, интерпретируемый в нашей задаче как радиус Инфоглобуса; h - ордината трактрисы и расстояние от центра плоской проекции Инфоглобуса до хорды, стягивающей концы интересующей нас плоской дуги (далее будем для краткости именовать такой образ связи двух объектов *h-дугой*); r – абсцисса трактрисы и радиус h -дуги.

Для нахождения радиуса h -дуги, соединяющей две различные точки на горизонте сферы, достаточно сведений из элементарной геометрии. Однако реальная задача построения образа связи двух объектов на круговой диаграмме оказывается значительно сложнее, т.к. образы объектов в действительности располагаются не на горизонте Инфоглобуса, а на его поверхности на некотором фиксированном расстоянии от линии горизонта (см. рисунки 1 и 2). Эту задачу удастся свести к предыдущей путем решения ряда подзадач с привлечением методов аналитической геометрии и поиска экстремума функций. В целом доказано, что в результате применения предлагаемой методики всегда может быть построен и притом единственный образ связи двух объектов в виде отрезка h -дуги.

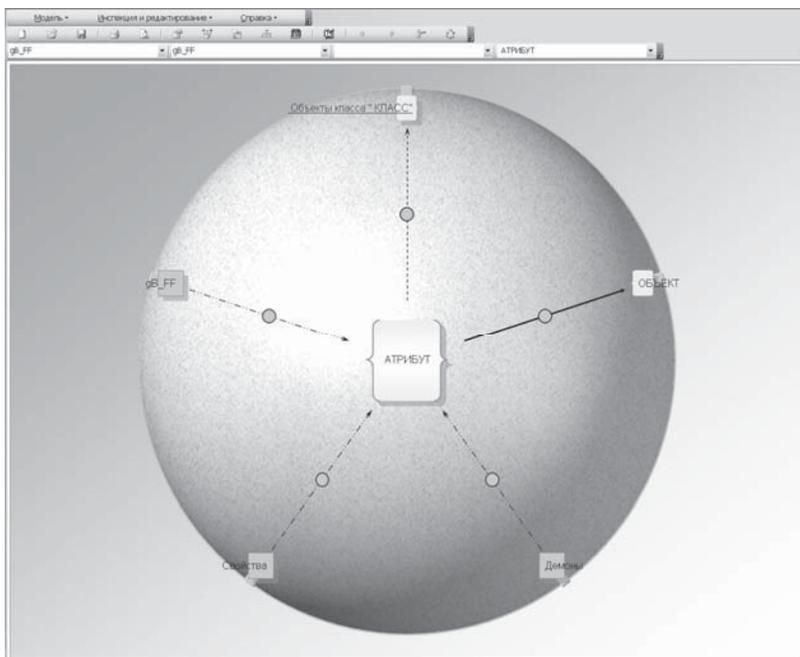


Рисунок 2 – Пример центрально-радиального представления фрагмента онтологической модели (обозрение объекта модели и его смежников)

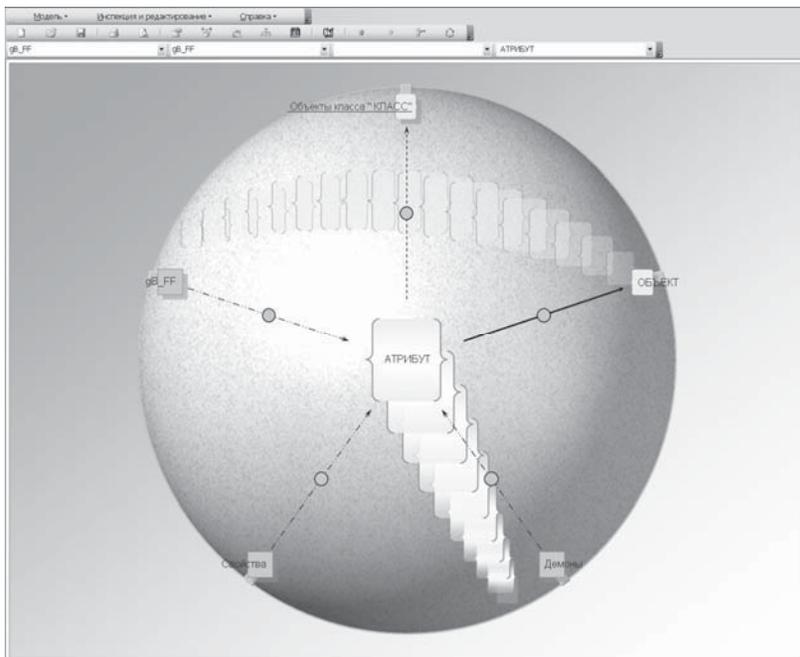


Рисунок 3 – Спрайт-анимация перемещения образов объектов при навигации в объектном пространстве онтологической модели: объект «АТРИБУТ» становится предметом обозрения, занимая ближайшую к наблюдателю точку Инфоглобуса; «ОБЪЕКТ» смещается в отводимое ему место среди объектов-смежников обозреваемого объекта

2.2 Образы объектов

Круг задач, связанных с построением на диаграммах образов объектов, оказывается существенно шире.

Во-первых, для имитации глубины предъявляемой диаграммы предложено аппроксимировать тело Инфоглобуса вертикально ориентированными дисками так, что поверхность Инфоглобуса, обращенная к наблюдателю, представляет собой совокупность колец, параллельных плоскости экрана. Здесь оказывается рациональным использовать равномерную дискретизацию 3D-пространства в смысле приращения телесного угла видимости дисков из центра аппроксимируемой ими сферы.

Второй важной задачей является расчет радиуса образа объекта в зависимости от его удаления от наблюдателя. В целом варианты удаления определяются набором аппроксимирующих Инфоглобус дисков, и показано, что в качестве огибающей образов объектов, по-разному удаленных от наблюдателя, но выстраиваемых по линии диаметра Инфоглобуса, может быть использована h -дуга. Тогда, если априори задать радиус образа объекта на круговой диаграмме (см. рисунок 1), рассматриваемую задачу удастся однозначно решить методами элементарной геометрии и поиска экстремума функций. Кроме того, удаленность от наблюдателя влияет на яркость образа объекта; соответствующие расчеты выполняются на основе известных физических закономерностей [14].

Третья группа задач связана с обеспечением спрайт-анимации при построении диаграмм обозрения. Здесь центральной становится задача вычисления прямоугольных (экранных) координат точки h -дуги при имитации равномерного движения вдоль нее образа объекта из некоторого исходного положения в заданное целевое. Дополнительно в зависимости от удаленности образа объекта от наблюдателя по априори рассчитанным опорным вариантам размеров и яркостей образов методом линейной интерполяции определяются действительные значения этих параметров в процессе анимации. На рисунке 3 зафиксированы промежуточные положения двух смещаемых образов объектов при переходе от обозрения фрагмента модели, представленного на рисунке 1б, к обозрению, приведенному на рисунке 2.

Заключение

Для полноценной реализации когнитивного потенциала разработанного графического интерфейса с интерактивной визуализацией онтологических моделей потребовалось решить ряд нетривиальных геометрических и аналитических задач, обусловленных применением 3D-фигур, организующих пространство обозрения.

Разработаны методы и алгоритмы расчета параметров изображения на поверхности Инфоглобуса образов связей и объектов, а также параметров, описывающих перемещение образов объектов по этой поверхности при навигации в объектных пространствах моделей. Внедрение этих результатов позволило преодолеть существовавший разрыв между возможностями и реализацией когнитивно-ориентированной графики представления онтологических моделей в разрабатываемой в ИПУСС РАН gB -системе моделирования.

Следует отметить, что при реализации интерфейса в конкретной среде программирования набор подлежащих решению геометрических и аналитических задач, скорее всего, увеличится по сравнению тем, что упомянут в данной статье. Так будут сказываться ограничения используемой библиотеки графических функций. Подобные стеснения могут вовсе блокировать использование полезных в когнитивном аспекте графических элементов. В частности, эти замечания оказались в определенной мере справедливы и в нашем случае: используемая при разработке интерфейса приложений gB -системы векторная графика VBA в MS Excel 2003 [15] в этом смысле весьма ограничена.

Благодарности

Работа поддержана и выполнена в рамках проекта «Онтологические модели ситуаций в процессах коллективного принятия решений» Программы фундаментальных научных исследований государственных академий наук на 2008-12 гг., направление 25.

Список литературы

- [1] Ситников П.В., Смирнов С.В. Методы и средства анализа объектных информационно-логических моделей // Проблемы управления и моделирования в сложных системах: Труды X международной конф. (23-25 июня 2008 г., Самара, Россия). – Самара: СамНЦ РАН, 2008. - С. 448-456.
- [2] Виттих В.А., Ситников П.В., Смирнов С.В. Онтологический подход к построению информационно-логических моделей в процессах управления социальными системами // Вестник компьютерных и информационных технологий. 2009. № 5. - С. 45-53.
- [3] Онтологический подход к разработке и внедрению информационно-аналитической системы управления предоставлением консолидированных услуг населению (на примере сферы здравоохранения и социального развития Самарской области) // Авт.: В.А. Виттих, Е.А. Гриценко, П.В. Ситников и др. / Проблемы управления и моделирования в сложных системах: Труды XI международной конф. (22-24 июня 2009 г., Самара, Россия). – Самара: СамНЦ РАН, 2009. - С. 28-43.
- [4] Смирнов С.В. Открытая архитектура инструментальных средств моделирования сложных систем // Проблемы управления и моделирования в сложных системах: Труды международной конф. (15-17 июня 1999 г., Самара, Россия). – Самара: СамНЦ РАН, 1999. - С. 57-64.
- [5] Смирнов С.В. Онтологическая относительность и технология компьютерного моделирования сложных систем // Известия Самарского научного центра РАН. 2000. Т.2. № 1. - С. 66-71.
- [6] Гаврилова Т.А., Муромцев Д.И. Интеллектуальные технологии в менеджменте: инструменты и системы: Учеб. пособие. 2-е изд. – СПб.: Изд-во «Высшая школа менеджмента»; Издат. дом С.-Петерб. гос. ун-та, 2008.
- [7] Касьянов В.Н., Евстигнеев В.А. Графы в программировании: обработка, визуализация и применение. – СПб.: БХВ-Петербург, 2003.
- [8] Апанович З.В. Методы интерактивной визуализации информации // Проблемы управления и моделирования в сложных системах: Труды X международной конф. (23-25 июня 2008 г., Самара, Россия). – Самара: СамНЦ РАН, 2008. - С. 478-489.
- [9] Katifori A., Halatsis C., Lepouras G., Vassilakis C., Giannopoulou E. Ontology visualization methods – a survey // ACM Comput. Surv. 39, 4 (2007), 10.
- [10] Апанович З.В., Винокуров П.С. Анализ онтологии и информационного наполнения портала знаний при помощи методов визуализации информации // Проблемы управления и моделирования в сложных системах: Труды XI международной конф. (22-24 июня 2009 г., Самара, Россия). – Самара: СамНЦ РАН, 2009. - С. 556-562.
- [11] A visual exploration on mapping complex networks [сайт] [2008]. URL: <http://www.visualcomplexity.com>
- [12] Смирнов С.В. Прагматика онтологий: объектно-ориентированная модель знаний о предметной области // Одиннадцатая Национальная конф. по искусственному интеллекту с международным участием КИИ-2008 (28 сентября-03 октября 2008 г., Дубна, Россия): Труды конф. Т. 3. – М.: ЛЕНАНД, 2008. - С. 208-216.
- [13] Математический энциклопедический словарь. – М.: Сов. энциклопедия, 1988.
- [14] Аленицын А.Г., Бутиков Е.И., Кондратьев А.С. Краткий физико-математический справочник. – М.: Наука, 1990.
- [15] Уокенбах Дж. Профессиональное программирование на VBA в Excel 2003. – М.: ООО «И.Д. Вильямс», 2006.

ОБЗОР ЗАДАЧ ЕЯ-СЕРВИСОВ В ИНФОРМАЦИОННЫХ СИСТЕМАХ ПОД УПРАВЛЕНИЕМ ОНТОЛОГИИ

Е.А. Сидорова

Институт систем информатики им. А.П. Ершова СО РАН
630090, Новосибирск, ул. Лаврентьева, 6, Россия
lena@iis.nsk.su
тел: +7 (838) 330-79-91

Ключевые слова: анализ текста, лингвистический ресурс, информационная система, онтология, предметная область

Abstract

The role of ontology in the process of creation of information systems which use text analysis services for automatic filling and updating their content is discussed. There is a feedback - system knowledge and data can help to resolve ambiguous situations in the text. The text analysis services solve various problems at different stages of development of system. Its based on additional knowledge and resources, such as domain vocabulary, model for text segmentation, and fact extraction schemes that relate vocabulary items and lexical-syntactic constructions to ontology entities.

Введение

В информационных системах (ИС) онтологию используют для того, чтобы облегчить пользователю поиск информации в большом наборе ресурсов путем структурирования и систематизации знаний, создания единой иерархии понятий, унификации терминов. В задачах извлечения информации из текстов онтология системы также играет существенную роль (можно говорить, что анализ осуществляется под управлением онтологии) [1].

Онтология ИС рассматривается как трехуровневая система знаний. *Метаонтология* специфицирует онтологии более низкого уровня: онтологию предметной и проблемной области (ПО). Обычно, описывается структура таких элементов как понятие, иерархическое отношение наследования между понятиями, бинарное отношение (ассоциативное, часть-целое), тернарное отношение (атрибут–тип–значение). *Онтология ПО* описывает предметную и проблемную область ИС в терминах понятий и отношений. *Онтология нижнего уровня* включает экземпляры понятий и отношений, образующих информационное наполнение ИС [2].

Большинство ИС в той или иной форме используют ЕЯ-сервисы, предназначенные для решения различных задач, связанных с анализом текста на естественном языке. Можно выделить два типа ЕЯ-сервисов: системные сервисы, используемые для автоматического наполнения и изменения содержания системы, и пользовательские сервисы, предоставляющие пользователям ИС разнообразный доступ к информации.

В данной работе будет рассматриваться роль онтологии и ЕЯ-сервисов в процессе создания и эксплуатации информационных систем, как именно используются знания онтологии при извлечении информации из текста, и как именно результаты, полученные при обработке текста, влияют на саму онтологию.

1 Сервисы информационной системы

Разработка ИС под управлением онтологии – это обязательно итеративный процесс, поэтому можно рассматривать жизненный цикл онтологии ИС в контексте использования

средств анализа текста для ее развития (см. рисунок 1). Отметим, что задачи решаемые сервисами анализа документов на разных этапах развития ИС различны, даже типы или жанры документов, с которыми приходится работать могут быть неодинаковы.



Рисунок 1 – ЕЯ-сервисы в информационной системе

На начальном этапе разработки системы онтология играет важную роль при анализе требований и концептуальном моделировании. На данном этапе осуществляется проектирование базы знаний системы – формируется онтология верхнего уровня (метаонтология), фиксируются основные понятия предметной области.

Онтологический анализ предметной области обычно начинается с создания словаря терминов, который используется при обсуждении и исследовании характеристик объектов и процессов, составляющих рассматриваемую предметную область, также выделяются основные логические взаимосвязи между введенными понятиями (и терминами). Таким образом, встает задача *автоматического извлечения предметной терминологии*, которая включает как однословные, так и многословные термины [3-5]. Для решения этой задачи используется подборка текстов по данной тематике, на которых применяются методы обучения словаря (под обучением понимается процесс формирования словаря со статистическими показателями), включающие морфологический и поверхностный синтаксический анализ текстов. Далее, на основе статистического распределения терминов выделяются общезначимые и предметные термины, для которых, используя различные методы кластеризации (как, например, в [6]), можно автоматизировать построение иерархических отношений и сформировать списки синонимов для дальнейшего анализа. Условием применения таких методов является наличие обучающего корпуса текстов – массива текстов специальным образом размеченного.

Следующий этап – наполнение системы необходимой информацией.

В первую очередь осуществляется добавление справочно-энциклопедической информации, вид и характер которой должны найти отражение в онтологии ПО. К справочному знанию относятся, например, номенклатурные обозначения, толкования понятий ПО или заранее известный список производственных объектов, такие знания представляются в виде экземпляров понятий и отношений онтологии ПО (поэтому их можно отнести к онтологии нижнего уровня). Таким образом, рассматривается задача *автоматического добавления справочной информации* в базу знаний системы, на основе анализа имеющихся справочных ресурсов, представленных в электронном виде. Справочные ресурсы – это, как правило, хорошо структурированные тексты, поэтому использование формальной жанровой модели таких ресурсов может значительно упростить процесс анализа текста, а также ускорить его настройку [7]. Еще одной особенностью данного типа ресурсов является наличие в тексте значимых несловарных единиц, выражаемых буквенно-числовыми конструкциями (например, H_2O , 5 м/с , $103-105 \text{ км}$,

корпус 2а, изделие №4в и т.п.). Извлечение таких единиц возможно на основе лексических шаблонов, описываемых на специальном языке (подробное описание и применение так конструкций можно посмотреть в работе [8]).

Основной задачей, решаемой системным сервисом анализа документов, является извлечение значимой для пользователей информации из большого объема поступающих документов (слабо-структурированных текстовых ресурсов). Задачей данного этапа является *извлечение значимых фактов* [9]. Факт, представляя собой зафиксированное в высказывании (языковом выражении) эмпирическое знание об объектах, их свойствах и ситуациях, может быть формализован в виде когнитивной схемы, соотносящей его с понятиями и отношениями онтологии. Такая схема, предназначенная для извлечения факта из текста, должна учитывать множество языковых способов репрезентации данного факта носителями подязыка и обеспечивать их трансформацию в формальную структуру объектов и отношений. Таким образом, результат извлечения фактов из текста представляется в виде семантической сети объектов, являющихся экземплярами понятий и отношений, заданных онтологией. Данная семантическая сеть добавляется в информационное пространство ИС и, таким образом, преобразуется в знания, которыми в дальнейшем может оперировать система и ЕЯ-сервисы.

Когда в системе накоплено достаточное количество данных возникает ситуация, когда при поступлении новой информации, ее требуется согласовывать с уже имеющейся в системе. Для этого необходимо обеспечить корректность, целостность, уникальность и актуальность, полученных в результате анализа данных. Т.е. встает *задача поддержки актуальности информации*, которая включает следующие подзадачи: идентификацию или отождествление информационных объектов (обозначающих один и тот же объект действительности), поиск объекта в базе данных системы по неполному набору характеристик (критерии выбора наиболее подходящего объекта), изменение свойств объектов БД (в связи с естественным изменением характеристик обозначаемого объекта в реальном мире), обработка противоречивых данных и т.п. С другой стороны, присутствует обратная связь – знания накопленные системой могут помочь разрешить неоднозначные ситуации в тексте, уточнить неявно выраженные характеристики объектов.

Внесение изменений в онтологию предметной области (а также, в онтологию верхнего или метаонтологию) возможно либо при изменении требований к системе со стороны пользователя, либо при накоплении достаточного количества фактов сигнализирующих о наличии неполноты в системе описания онтологии. Данные факты могут извлекаться из текста по специальным правилам с обязательным требованием высокой точности.

Для доступа к информации ИС разрабатываются пользовательские ЕЯ-сервисы, такие как информационный поиск фактов или документов, содержащих определенные факты, представление кратких рефератов просматриваемых документов, структурирование информации, полученной по поисковому запросу пользователя (рубрикация, кластеризация) и т.п. Сам запрос пользователь может оформлять либо на естественном языке в виде вопроса (вопросно-ответное взаимодействие), либо по ключевой фразе, либо заполняя определенную форму (формируя тем самым структурированный в соответствии с ПО запрос), либо используя навигационные средства представляемые ИС.

2 Описание процесса обработки текста

Минимальной единицей анализа текста является словоформа (слово), число или специальная символьная конструкция. На начальном этапе каждая единица идентифицируется – слово ищется в словаре с учетом его парадигмы, для прочих символьных выражений определяется их тип. Процесс идентификации слов осуществляется на основе знаний, хранящихся в словарях. Дальнейшая обработка заключается в последовательном анализе контекста найденных лексем (и других минимальных единиц).

В первую очередь выявляются *устойчивые терминологические словосочетания* (слово-комплексы), характеризующиеся наличием определенной синтаксической связи между контактно расположенными словами и высокой частотностью в анализируемом подязыке, – для выявления контекста данного типа обычно составляются словари, которые содержат знания, необходимые для идентификации словокомплексов в тексте.

Затем рассматриваются неустойчивые словосочетания (*фразы*), для которых не задан строгий порядок слов (или которые не являются частотными в анализируемом подязыке). Фразы могут формироваться нерегулярным образом: быть разрывными, использовать различные элементы языковой отсылки (анафорические и дейктические выражения).

Следующий этап – поиск *связанных фраз*. Данный контекст характеризуется определенным линейным порядком связанных частей, использованием различных видов языковой редукции (сочинение и другие виды эллипсиса, анафора) и позволяет извлекать из текста информационные объекты со сложной структурой и связи между объектами.

На всех стадиях анализа явно или неявно используется такой тип контекста как *текст*, важным свойством которого является его жанр. Жанр играет важную роль в определении формальной структуры создаваемого автором текста, в которой, помимо логического уровня, т.е. уровня разбиения на предложения, абзацы, параграфы и т.п., выделяется уровень концептуальный, или собственно жанровый (наличие в тексте таких разделов, как заголовок, резюме, основной текст) [7].

Следующий тип рассматриваемого контекст – знания, извлеченные из *ранее обработанных документов*, которые хранятся в базе знаний ИС (как альтернатива, они могут включаться в разметку корпуса текстов, используемого в задачах автоматического обучения). Как уже было сказано выше, эти знания могут помочь уточнить неявно выраженные характеристики, заполнить недостающие параметры объектов и связей.

Технологическое описание процесса анализа включает следующие этапы: графематический анализ, лексический анализ, морфологический анализ, жанровая сегментация, поиск и извлечение фактов, идентификация информационных объектов и формирование контента документа.

В рабочем режиме обработки документа на вход модуля анализа поступает текст, который передается лексическому процессору словарного компонента. Словарный компонент осуществляет первичную сегментацию и извлечение ключевых терминов. Далее модуль сегментации осуществляет жанровую декомпозицию текста и передает все полученные результаты модулю сборки фактов, который осуществляет поиск фактов в тексте, используя упорядоченный набор схем фактов. Результатом работы модуля сборки фактов является множество информационных объектов. Модуль формирования контента документов идентифицирует и уточняет параметры полученных объектов, сравнивая их с информационными объектами, хранящимися в ИС, формирует множество информационных объектов, представляющее проанализированный документ и его контент, а также устанавливает между ними необходимые связи.

3 Лингвистические ресурсы

Очевидно, что знаний о предметной области, хранящихся в онтологии ИС, недостаточно для автоматического извлечения информации из текста – требуются дополнительные лингвистические ресурсы, которые содержат знания о языке, на котором представляется информация (правила и способы выражения, принятые в текстах по рассматриваемой тематике).

Рассмотрим подробнее, какие лингвистические ресурсы могут использоваться при анализе текста.

- 1) Корпус текстов – подборка текстов определенного жанра, тематика которых соответствует заданной предметной области. Корпус может содержать лингвистическую разметку, представляющую собой информацию, полученную автоматически при анализе текстов, либо приписанную экспертом вручную. Основное назначение корпуса – автоматизация создания других лингвистических ресурсов.
- 2) Универсальные и предметные словари, содержащие перечень минимальных единиц языка, а также устойчивых терминологических словосочетаний, используемых при описании значимой для ИС информации.
- 3) Набор универсальных и/или специфичных для заданного подязыка лингвистических знаний: морфологические классы (определение лексемы для найденной в тексте словоформы), семантико-синтаксические модели согласования терминов языка и т.п. Эти знания могут быть заданы с разной степенью подробности в зависимости от требований и возможностей разработчиков ИС.
- 4) Знания о согласовании имеющихся лингвистических знаний с предметными знаниями, заданными онтологией ИС. С этой целью термины группируются в семантические группы (или им приписываются семантические ориентации), которые в свою очередь также согласуются с элементами онтологии либо непосредственно, либо в соответствии с определенной схемой (схемой факта).
- 5) Набор описаний жанровых структур текста, соотнесенных с тем или иным типом текстовых ресурсов, хранящихся в ИС.

Дополнительно может рассматриваться прагматический контекст, и в этом случае в качестве лингвистического ресурса может фигурировать проблемная модель, зависящая от цели, которую преследуют разработчики системы анализа текста (например, модель коммуникации для вопросно-ответных диалоговых систем).

Для создания всех представленных ресурсов требуются дополнительные технологические компоненты, позволяющие автоматизировать их создание, начальное наполнение, а также обеспечить дальнейшую поддержку (внесение изменений и синхронизацию).

Задачи	Ресурсы	Средства
Извлечение терминологии	Корпуса текстов, Сем. словари	Средства разметки текстов
Обработка структурированных текстов	Лекс. шаблоны Модели документов,	Методы автоматического построения сем. словарей
Извлечение информации (фактов)	Сем.-синт. модели согласования	Средства описания шаблонов, моделей и схем фактов
Поддержка актуальности информации	Схемы извлечения фактов	Стратегии разрешения противоречий на данных

Рисунок 2 – Список ресурсов для поддержки ЕЯ-сервисов

На рисунке 2 представлены лингвистические ресурсы и их соответствия с выполняемыми сервисами задачами и дополнительными средствами, необходимыми для их создания.

Заключение

Таким образом, были рассмотрены основные задачи, решаемые сервисами автоматической обработки текстов на естественном языке в процессе создания, развития и эксплуатации ИС под управлением онтологии. Выделено два типа сервисов системные и пользовательские. Рассматриваемые далее необходимые технологические средства создания сервисов ориентированы в большей степени на первый тип, т.к. именно от качества их работы зависит скорость создания и эффективность использования (не говоря уже о качестве содержания) ИС.

Благодарности

Работа выполняется при финансовой поддержке РФФИ (проект № 09-07-00400) и Президиума РАН (Интеграционный проект СО РАН № 2/12 в рамках программы фундаментальных исследований РАН № 2).

Список литературы

- [1] Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. Учебник. СПб.: Питер, 2001.
- [2] Загорюлько Ю.А., Боровикова О.И. Подход к построению порталов научных знаний // Автотметрия. № 1, 2008, т. 44. – С. 100–110.
- [3] Сидорова Е.А. Многоцелевая словарная подсистема извлечения предметной лексики // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог 2008». – М.: РГГУ, 2008. – С. 475–481.
- [4] В.Д. Гусев, Н.В. Саломатина Алгоритм выявления устойчивых словосочетаний с учетом их вариативности (морфологической и комбинаторной) // Труды между. конференции Диалог-2004 "Компьютерная лингвистика и интеллектуальные технологии. – М.: Наука, 2004. – С. 530–535.
- [5] Лукашевич Н.В., Добров Б.В., Чуйко Д.С. Отбор словосочетаний для словаря системы автоматической обработки текстов // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог 2008». – М.: РГГУ, 2008. – С. 339–344.
- [6] Крижановский А.А. Автоматизированное построение списков семантически близких слов на основе рейтинга текстов в корпусе с гиперссылками и категориями // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции Диалог'2006 "Компьютерная лингвистика и интеллектуальные технологии". – М.: РГГУ, 2006. – С. 297–302.
- [7] Сидорова Е.А., Кононенко И.С. Представление жанровой структуры документов и ее использование в задачах обработки текста // Труды Седьмой Международной конференции памяти академика А.П.Ершова "Перспективы систем информатики". Рабочий семинар «Научное программное обеспечение». – Новосибирск: Сибирское Научное Издательство, 2009. – С. 248–254.
- [8] Загорюлько Ю.А., Кононенко И.С., Костов Ю.В., Сидорова Е.А. Подход к интеллектуализации документооборота // Информационные технологии, № 11, 2004. – С. 2–11.
- [9] Сидорова Е.А., Кононенко И.С. Подход к извлечению фактов из текста на основе онтологии // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог 2009» (Бекасово, 27–31 мая 2009 г.). Вып. 8 (15). М.: РГГУ, 2009. – С. 451–457.

ПРИМЕНЕНИЕ ЛИНГВИСТИЧЕСКИХ ТЕХНОЛОГИЙ В СПРАВОЧНО-ИНФОРМАЦИОННОЙ СИСТЕМЕ ПО КАТАЛИЗУ

И.С. Кононенко, Е.А. Сидорова

Институт систем информатики им. А.П. Ершова СО РАН
630090, Новосибирск, ул. Лаврентьева, 6, Россия
irina_k@cn.ru, lena@iis.nsk.su
тел: +7 (838) 330-79-91, факс: +7 (838) 333-83-59

Ключевые слова: тезаурус, справочно-информационная система, извлечение терминологии, фреймовый анализ

Abstract

Computational linguistic techniques for life science text are proposed that partially automate the process of development of information system on catalysis. A few aspects of terminology and thesaurus relations extraction are discussed. Information on domain situations (molecular interactions, or processes) is considered to be identified on base of special-purpose linguistic resource that involves a semantically annotated text corpus and a dictionary of predicate words with their semantic and syntactic arguments.

Введение

Опубликованная и доступная информация в области катализа достигает огромных масштабов, что является стимулом для создания справочно-информационной системы по катализу, которая могла бы предоставлять пользователю актуальный информационный срез по интересующему его вопросу из данной области.

Публикуемая информация большей частью представлена в виде научных статей, т.е. в текстовом виде, требующем дальнейшей обработки и систематизации. Работа с текстами, богатыми специальными понятиями и лексикой, требует наличия дополнительных средств представления предметной области (ПО), сопряженных с ориентированными на ПО ресурсами: словарями, тезаурусами, базами знаний, обучающими корпусами текстов. Они необходимы как в справочных целях, так и в качестве основы для создания специализированных поисковых систем.

Целью данной работы является разработка компьютерных методов анализа естественно-научной текстовой информации по катализу, обеспечивающих частичную автоматизацию решения следующих задач:

- построение предметного словаря и тезауруса, ориентированных на автоматическую обработку текстов по катализу и биотехнологиям;
- накопление и интеграция знаний о молекулярных взаимодействиях (процессах) для поддержки ситуационного поиска в справочно-информационной системе по катализу.

1 Создание предметных словарей для пополнения справочно-информационного тезауруса

В качестве основы информационно-справочной системы в настоящее время создан тезаурус по гетерогенному катализу (около 20000 терминов) и рубрикатор по гетерогенному и промышленному гетерогенному катализу (около 3000 рубрик) [1,2]. Начальное наполнение тезауруса осуществлялось на основе оглавления и предметного указателя ряда учебных и справочных пособий, что позволило составить базовый список имен понятий (дескрипторов) данной

предметной области. Однако, учитывая задачу создания системы, осуществляющей обработку новых текстов и поиск в них необходимой информации, следует отметить два существенных недостатка составленного списка – это его неполнота и неестественность языковой формы представленных в нем терминов. Так, сложные термины тезауруса (множества дескрипторов) по структуре и порядку слов не эквивалентны многословным терминам (словосочетаниям), представляющим их в реальном тексте, например, «окисление, каталитическое», «Яна-Теллера, эффект» или «Тёмкина, изотерма адсорбции» (неестественный порядок слов); «ацетилен, селективное гидрирование», «молибдат железа, катализатор» (ср. в тексте *железномолибдатный катализатор*). Таким образом, встает задача составления предметного терминологического словаря как множества слов и словосочетаний, функционирующих в реальных текстах и представляющих все возможные естественно-языковые способы именованности понятий ПО. Автоматизированное извлечение терминологии из текстов ПО позволило бы пополнить тезаурус отсутствующими в нем терминами, а также сопоставить терминам тезауруса множества соответствующих естественно-языковых выражений.

Еще одной проблемой при создании тезауруса является установление парадигматических отношений тезаурусного типа, связывающих элементы тезауруса: синонимия, гипонимия-гиперонимия (выше-ниже, с различными модификациями, например, isa-отношение). Как правило, это осуществляется вручную и требует больших затрат от экспертов-разработчиков. Частично автоматизировать эту работу позволяет автоматическое выявление и предоставление для просмотра эксперту определенных контекстов употребления терминов в тексте.

Для решения этих задач развивается технология словарной обработки текста, описанная в работе [3]. Данная технология содержит компоненты, с одной стороны, позволяющие экспертам (лингвистам, специалистам в конкретной предметной области и инженерам знаний) формировать базу знаний (словарь, правила сборки словосочетаний, таблицы признаков и т.п.) и, с другой стороны, – обеспечивающие автоматическое применение этих знаний в процессе обработки документов, в том числе и для обогащения самой базы знаний.

2 Извлечение терминологии

В процессе создания терминологического словаря можно выделить три содержательных этапа: первоначальное наполнение, сортировка (фильтрация) и экспертная оценка.

Процесс *первоначального наполнения словаря* осуществляет извлечение из массива текстов терминов-кандидатов на базе определенных лингвистических ограничений (морфологических и синтаксических моделей) с использованием технологии обучения по массиву текстов. В качестве исходного корпуса текстов использованы учебные тексты по катализу ([4]).

Морфологический анализ текстов (на основе русского морфологического анализатора группы Диалинг [5]) позволяет извлечь однословную лексику.

Поверхностный синтаксический анализ выделяет в текстовых сегментах контактные синтаксически связанные группы слов (словокомплексы) на основе фиксированного набора правил. В настоящий момент нами используется порядка 20 правил сборки словокомплексов. Наиболее распространенными структурами здесь являются сочетания существительного с прилагательным, существительного с существительным в косвенном падеже. Имеются также многословные термины, состоящие из трех и более слов.

В результате первого этапа был создан словарь, словник которого имеет объем 10236 слов и 39784 словокомплекса. Учитывая, что многие термины предметной области универсальному словарю-анализатору неизвестны, использовались средства предсказания незнакомых слов. При извлечении многословных терминов использовались правила построения именных групп (именно такую грамматическую структуру имеют большинство терминов, поэтому правила для других частей речи были исключены).

На этапе *фильтрации* полученный список терминов-кандидатов сортируется для выявления наиболее важных (статистически значимых в данной ПО) терминов с использованием накапливаемой в словаре статистической информации. Существуют методики сортировки, использующие более сложные статистические метрики, чем простой подсчет суммарной частотности [6].

Дальнейшую сортировку и предварительный отбор лексики проводит эксперт-лингвист, исходя из общих соображений о терминологичности и лингвистической корректности. Так, в процессе предварительного отбора однословных терминов учитывается, что в тезаурусе должны присутствовать лишь знаменательные лексемы четырех основных частей речи (существительные, прилагательные, глаголы и наречия) – прочие лексемы удаляются в стоп-словарь.

Многие поисковые и аналитические задачи требуют извлечения информации об именованных сущностях, описываемых с использованием лексики специальных лексико-семантических разрядов (фамилии, имена, локализации, организации), но в данной конкретной ситуации имена собственные не должны предлагаться пользователю в качестве терминов-кандидатов, поэтому лексемы, отнесенные к разряду фамилий, должны быть на этом этапе отнесены к общезначимой лексике (например, отсеяна лексема *боденштейн* – фамилия немецкого химика), в отличие от словосочетаний, содержащих такие лексемы: *квасистационарная концентрация боденштейна*, *квасистационарность боденштейна*, *концентрация боденштейна*, *принцип квазистационарности боденштейна*.

Что касается нарицательных существительных, проводится удаление всех одушевленных имен (*академик*, *механик*, *химик*, *лауреат*, *организатор* и т.п.), ряда очевидных общенаучных терминов.

В процессе фильтрации учитываются возможности грамматических и орфографических ошибок при анализе неизвестных слов методом предсказания. Из словаря должны быть удалены (отсеяны в словари удаленных терминов) явно ошибочные предсказания, в том числе

- грамматические ошибки в предсказании части речи, ошибки морфологического класса при верной части речи (неверный род, имя собственное вместо нарицательного и т.п.), ошибки нормальной формы (*квантовый-химический* вместо *квантово-химический*);
- некорректности, основанные на ошибках в тексте (*спектр*, *циклогксана*, *циклогксена*) извлеченные из содержащих ошибки текстовых фрагментов (*ттр-спектры циклогксана (а)*, *циклогксена (б)* и *циклогксадисиа (в)*, *адсорбированных на pt/al2o3 1 - десорбция циклогксана. 2 - десорбция бензола способы расчета циклогксана*), и построенные на их основе словосочетания; для выявления последних использовались средства автоматического поиска словокомплексов, содержащих данное (некорректное) слово.

В результате предварительного отбора объем словаря однословных терминов-кандидатов сократился до 7884. Чистка словаря многословных терминов проводится преимущественно на следующем этапе, т.к. должна опираться в большей степени на экспертные знания.

Сформированный словарь отфильтрованных терминов-кандидатов передается эксперту-специалисту в данной ПО для проведения *экспертной оценки*. На данном этапе необходимо осуществить сравнение полученного словаря с текущим тезаурусом. Учитывая, что имена понятий (дескрипторы) во многих случаях не имеют естественно-языковой формы словосочетания (ср. дескриптор *окисление каталитическое, парциальное*, в текстах *парциальное каталитическое окисление*), для автоматизации такого сравнения рекомендуется использовать информацию о составе словосочетания.

Работа экспертов на втором и третьем этапах поддерживается Конкордансом (рисунок 1), который дает полный индекс терминов в непосредственном и расширенном контекстах и позволяет получить все примеры употребления термина словаря вместе с его контекстами. Таким образом, конкорданс осуществляет обратную связь словаря с массивом текстов и обеспечивает своего рода лингвистическую разметку этого массива. Реализованный в системе модуль конкорданса работает с текстовыми файлами, каждый из которых может содержать не-

сколько фрагментов, размеченных метатекстовой информацией. При просмотре контекста вхождения термина пользователь может самостоятельно определять длину просматриваемого фрагмента текста (поддерживается пословное расширение контекста, просмотр абзаца или всего текста).

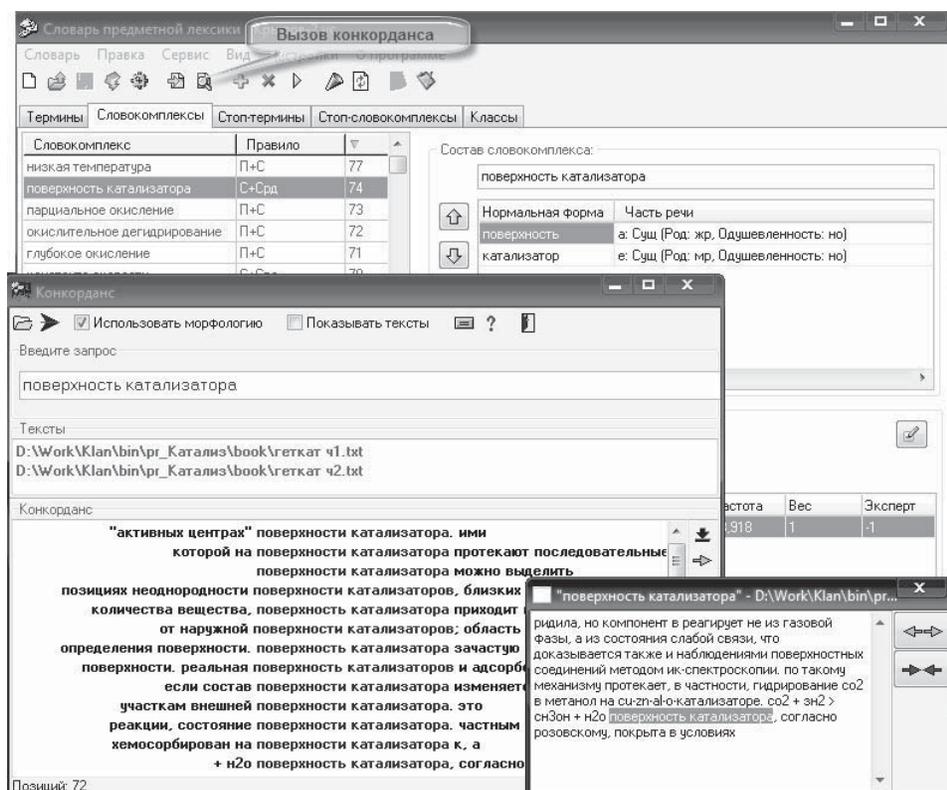


Рисунок 1 – Просмотр конкорданса в предметном словаре по катализу

3 Установление тезаурусных отношений

Задача автоматизации выявления тезаурусных связей рассматривалась на примере отношения гипонимии (и обратного ему отношения гиперонимии). Гипонимия вместе с несовместимостью, по словам Дж. Лайонза [7], «являются самыми фундаментальными парадигматическими смысловыми отношениями, посредством которых структурирован словарный состав языка», вследствие чего данному отношению уделяется особое внимание при разработке тезаурусов.

Традиционно установление этого отношения базируется на критерии Лайонза, определяющем гипонимию в терминах односторонней импликации: для пары терминов X и Y (в предположении, что Y является гиперонимом для X) строятся предложения вида “Это есть X” и “Это есть Y”, в которых данные термины являются логическими предикатами, и проверяется, что из истинности первого следует истинность второго. Этот критерий хорошо применим к текстам, содержащим дефиниции (научным статьям, терминологическим словарям, энциклопедиям и т.п.). В более общем случае, анализируя большой корпус текстов, в том числе научных текстов, содержащих новую информацию, можно выявить в нем пары терминов-кандидатов, подозреваемых на гипонимическую связанность (гипонимическое отношение). Для этого необходимо обнаружить в текстах некоторые квазигипонимические конструкции

вида "X и другие/прочие/остальные Y-и", "все Y-и, кроме/помимо/за исключением X-а", "все Y-и, включая/в том числе X", "некоторые/одни/многие Y-и, например/в частности, X" и т.п.

Для выявления соответствующих контекстов используется технология Алекс [8], предназначенная для описания и идентификации в тексте фрагментов произвольной сложности (в общем случае, разрывных) с помощью шаблонных лексических конструкций. Язык описания шаблонов позволяет определять произвольные символьные выражения, в том числе выражения, маркирующие границы извлекаемых объектов (X и Y), указывать альтернативные (синонимичные) выражения, разрывы и перестановки в лексическом составе конструкции, задавать условия, опциональные подвыражения и т.д. Шаблоны описывают основные классы конструкций, выражающих отношение между классом и подклассом, между множеством и элементом (включение, исключение, экзemplификация, чистая, либо сопровождаемая кванторным выражением). На рисунке 2 демонстрируются результаты применения технологии шаблонов: в тексте идентифицирован квазигипонимический контекст, извлечены и предоставлены на рассмотрение эксперту термины-кандидаты на гипонимическую связанность *биметаллические катализаторы* (Y), *Pd-Au/Al2O3* (X1), *Pd-Ag/Al2O3* (X2).

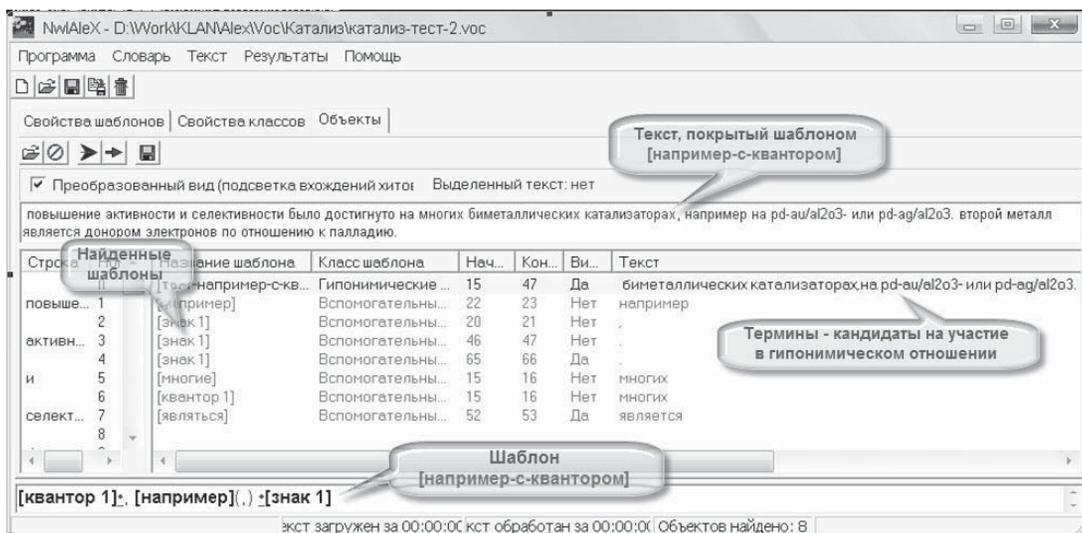


Рисунок 2 – Извлечение из текста терминов-кандидатов на гипонимическую связанность

4 Извлечение информации о молекулярных взаимодействиях

Предлагаемый подход к созданию системы извлечения, накопления и интеграции знаний о молекулярных взаимодействиях основан на методе ситуационного (или фреймового) анализа, который предполагает исследование и формальное представление типовых ситуаций в изучаемой ПО и способов их репрезентации в текстах. Стереотипная ситуация представляет собой определенную концептуализацию реальной ситуации, определяющую каркас высказывания в тексте из ПО, и включает фиксированный набор смысловых компонентов: предикатное слово, его семантический тип, аргументы (участники) ситуации и их семантические роли.

Фреймовый анализ предполагает создание языкового ресурса, включающего словарь предикатных слов с информацией об их семантических связях и способах их реализации (в том числе о предложно-падежном управлении) и множество семантически и грамматически размеченных текстов (ср. проект Framenet для английского языка [9] и идею соответствующего ресурса для русского [10]). При разработке узко специализированной информационной системы предполагается использование обучающего корпуса текстов с экспертной семантической раз-

меткой, на основе которой создается список предикатов, аргументами которых могут являться именованные сущности (названия химических элементов и соединений), и формируется набор универсальных ситуаций, описывающих процессы молекулярного взаимодействия в катализе, такие как:

вещество1 И *вещество2* «участвовали в» *процесс1* «с использованием» *вещество3* «с результатом» *вещество4*.

Здесь аргументы 1 и 2 выступают в роли Реагента, а 4 – в роли Продукта (результата процесса). Дополнительно отмечается аргумент 3 в роли Катализатора. Диагностирующим контекстом, позволяющим предположить наличие в тексте описания такой ситуации, является присутствие в нем целевого термина-обозначения химического Процесса, либо обобщенного Процесса-заместителя. Таким образом, рассматриваются следующие семантические классы терминов:

- Вещество – химический элемент или соединение (имеющее тривиальное естественно-языковое, либо систематическое номенклатурное название) может выступать в семантической роли Реагента, Продукта или Катализатора.
- Процесс – химическая реакция, характеризующаяся превращением одного или нескольких исходных веществ (реагентов) в отличающиеся от них по химическому составу или строению вещества (продукты реакции).
- Свойство вещества – рассматриваются физические и химические свойства веществ, такие как плотность, температура кипения и плавления, теплопроводность, состав ($H_2:CO = 2:1$) и т.д. Отдельно можно выделить свойства катализатора, к которым дополнительно относятся такие параметры как физический носитель (вещество), его форма (*решетка, цилиндрические гранулы, кольца, шары*), состояние (*пористость, термостойкость*), размер и т.п.
- Свойство процесса – помимо стандартных свойств реакции (*ядерная, механохимическая, сильноэндотермическая*), характеризует условия проведения или протекания процесса (давление, температура и т.п.)

На основе экспертной разметки текстов формируется словарь семантико-синтаксических описаний ситуаций (шаблонов), фрагмент которого приведен в таблице 1.

Таблица 1 – Семантико-синтаксические шаблоны ситуаций

Процессы	Семантические аргументы (роли и способы выражения) ¹			Примеры
	Реагенты	Реагент-каузатор	Продукт	
окисление	Сущ,рд	Сущ,тв *	в/во, Сущ,вн / до, Сущ,рд	<i>..при неполном окислении альдегидов в непредельные кислоты в присутствии оксидов Mo и V</i>
окисляться	Сущ ²	Сущ,тв *	в/во, Сущ,вн / до, Сущ,рд	<i>..Образующийся NO затем частично окисляется кислородом в газовой фазе до NO₂ ..полупроводники p-типа, легко окисляющиеся в высшие оксиды</i>

¹ Помета * означает, что аргумент может быть не выражен непосредственно при целевом слове, при этом указывается возможность восстановления значения аргумента по умолчанию (окисление -> Реагент-каузатор=кислород).

² Предложно-падежная форма синтаксического существительного не указывается в шаблоне, если она определяется конкретной формой реализации глагольной лексики-наименования Процесса (личный глагол, причастие, инфинитив). В этом случае идентификация аргумента происходит с помощью общих правил обработки шаблонов для глагольных предикатов.

Окончание таблицы 1

окислять	Сущ,вн	Сущ	в/во, Сущ,вн / до, Сущ,рд	..кислород окисляет восстановленную фазу (металл) со скоростью k_{rO_2} ..удобнее при сероочистке сероводород окислять прямо в серу, а не в SO_2
окислительная реакция	между, Сущ,тв Сущ,рд с, Сущ,тв		до, Сущ,рд	.. окислительная реакция между витаминами С и В12 .. Окислительная реакция кислорода с металлами .. Окислительная реакция закиси углерода до окиси углерода ..может произойти неконтролируемая окислительная реакция алюминия

Шаблон отдельной ситуации представляет собой совокупность синонимичных предикатов-наименований процесса (лексемы разных частей речи, словокомплексы), с указанием семантических ролей аргументов и соответствующих морфосинтаксических способов выражения для каждого предиката (аналогично словарю моделей управления). Заметим, что роль Катализатора и способы ее выражения универсальны (в присутствии+Сущ,рд, на+Сущ,пр), т.е. не зависят от конкретного процесса.

Все множество рассматриваемых ситуаций сводится к трем структурным типам:

- элементарные ситуации
Свойство + ХимОбъект,
где в роли Химического объекта могут выступать Процессы, Вещества:
синтез-газ состава $H_2:CO = 3:1$, сильноэндотермические реакции;
- простые **ситуации**
Процесс + Реагент^{} + Катализатор^{*} + Продукт^{^3}*,
где Процесс – это «одноступенчатая» реакция, у которой в роли Реагентов и Продуктов выступают Вещества, а роль Катализатора в общем случае факультативна:
процесс переработки метана в синтез-газ, Реакция ПКМ в промышленности протекает на никелевом катализаторе при 900-1000 °С и объемной скорости ~1000 ч;
- сложные **ситуации** соответствуют различным взаимосвязям простых ситуаций:
 - Реализация:
Окислительная конверсия метана в синтез-газ может быть осуществлена методом паровой конверсии метана.
 - Часть-целое и Следование:
Протеканию этих реакций сопутствует также паровая конверсия СО, Иногда проводят предконверсию на никелевом катализаторе при 350-550 °С.

Эти отношения формируют многоступенчатые процессы, при которых цепочка реакций образует единый Процесс с общим Продуктом.

Для автоматического поиска в тексте таких ситуаций используется аппарат семантико-синтаксических шаблонов, позволяющий представлять возможные семантические и синтаксические связи предикатов и идентифицировать их в тексте, а также формализм описания схем извлечения фактов [11], в результате применения которых строится ассоциативная семантическая сеть объектов, участвующих в типичных ситуациях молекулярного взаимодействия.

³ Обозначения: * повторитель от 0..n, ^ повторитель от 1..n.

Заключение

В работе рассмотрены возможности частичной автоматизации работы по созданию специализированной информационно-справочной системы путем применения лингвистических технологий к обработке естественнонаучных текстов.

В фокусе внимания были вопросы, связанные с наполнением тезауруса – извлечение терминологии и установление тезаурусных отношений.

Кроме того, предложен подход к разработке системы извлечения из текстов знаний о молекулярных взаимодействиях, основанный на методе фреймового анализа. Обучение такой системы планируется путем создания специализированного ресурса – корпуса текстов с экспертной семантической разметкой элементов стереотипных ситуаций из данной ПО.

Благодарности

Работа выполняется при финансовой поддержке Президиума СО РАН (Междисциплинарный интеграционный проект № 111).

Список литературы

- [1] Кузьмин А.О., Ильина Л.Ю., Гусев В.Д., Саломатина Н.В., Пармон В.Н.. Разработка тезауруса и рубрикатора для информационно-поисковой системы по катализу // Труды Седьмой Международной конференции памяти академика А.П.Ершова "Перспективы систем информатики". Рабочий семинар «Научное программное обеспечение». –Новосибирск: Сибирское Научное Издательство, 2009. – С.183-189.
- [2] Heterogeneous catalysis thesaurus (www.catalysis.ru/thezaurus/application/).
- [3] Сидорова Е.А. Многоцелевая словарная подсистема извлечения предметной лексики // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог 2008». – М.: РГГУ, 2008. Вып. 7 (14). – С. 475-481.
- [4] Крылов О.В. Гетерогенный катализ: Учебное пособие для вузов. – Академкнига, 2004.
- [5] Проект Диалинг (www.aot.ru).
- [6] Добров Б.В., Лукашевич Н.В., Сыромятников С.В. Формирование базы терминологических слово-сочетаний по текстам предметной области // Труды пятой всероссийской научной конференции "Электронные библиотеки: Перспективные методы и технологии, электронные коллекции". – СПб.: СПбГУ, 2003. – С. 201-210.
- [7] Лайонз Дж. Введение в теоретическую лингвистику, гл.10.– М.: Прогресс, 1978.
- [8] Жигалов В.А., Жигалов Д.В., Жуков А.А., Кононенко И.С., Соколова Е.Г., Толдова С.Ю. Система ALEX как средство для многоцелевой автоматизированной обработки текстов // Труды международного семинара Диалог'2002 по компьютерной лингвистике и ее приложениям. Т.2. – М.: Наука, 2002. – С.192–208.
- [9] Johnson C., Fillmore C., Petruck M, Baker C., Ellsworth M., Ruppenhofer J., and Wood E. FrameNet: Theory and Practice (framenet.icsi.berkeley.edu).
- [10] Ляшевская О.Н., Кузнецова Ю.Л. Русский фреймнет: к задаче создания корпусного словаря конструкций // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог 2009». Вып. 8 (15). М.: РГГУ, 2009. – С. 306-312.
- [11] Сидорова Е.А., Кононенко И.С. Подход к извлечению фактов из текста на основе онтологии //Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог 2009». Вып. 8 (15). М.: РГГУ, 2009. – С. 451-457.

КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ КОНТРОЛЛЕРА ОБЕСПЕЧЕННОСТИ РЕСУРСАМИ В СППР НА ОСНОВЕ ОНТОЛОГИЙ

А.Н. Гинзбург¹, М.В. Игнатьев²

¹Институт прикладной физики РАН
603005, Нижний Новгород, ул. Ульянова, 46, Россия
aginzburg@bk.ru
тел. +7 (831) 436-80-67

²Национальный исследовательский университет Высшая Школа Экономики в Нижнем Новгороде
603155, Нижний Новгород, ул. Б.Печерская, 25/12, Россия
mikhail.ignatyev@gmail.com

Ключевые слова: сложные искусственные системы, поддержка принятия решений, обеспечение ресурсами, онтологический подход, расширение функциональности информационных систем

Abstract

This paper describes the conceptual ideas about the controller of resources. Present the key features of the controller, described the decision making process using the controller. Formulated a key requirement for the system of decision support and justify the need for the controller.

Введение

Необходимость создания интеллектуальных компьютерных систем поддержки принятия решений (СППР) в сложных производственных, административных или иных искусственных системах на современном этапе развития не вызывает сомнения. Для информационных систем такого типа можно сформулировать общую цель их создания: оперативное представление различных видов информации, необходимой для принятия решения.

Методологические подходы к построению информационных систем, которые предоставляют подобного рода функциональность, достаточно широки [1]: от систематизированного таблично-графического представления информации до сложных комплексов по извлечению информации из неструктурированных и разрозненных источников, различных систем имитационного моделирования.

Современное положение дел в сфере проектирования и разработки информационно-аналитических систем поддержки принятия решений дает достаточное представление о функциональных особенностях таких систем и тенденции их развития [2].

Очередной ступенью в эволюции таких систем становятся системы, спроектированные и разработанные на основе онтологического подхода к обработке знаний [3] и холонического подхода [4] к системному анализу деятельности предприятий или организаций.

Концентрируясь на функциональных возможностях, следует обратить внимание на такую важную особенность процессов, происходящих в искусственных системах, как необходимость обеспечения ресурсами на любой стадии выполнения.

В данной работе описывается концептуальная модель части СППР, которая обеспечивает оперативную поставку различной информации о ресурсах, обрабатывая структурированный запрос о необходимости обеспечения ресурсами того или иного управленческого решения.

1 Прагматическая необходимость создания контроллера ресурсообеспеченности

Информационные системы, которые успешно функционируют в структурах искусственных систем, как правило, уже имеют возможности по работе с ресурсами. Информационные системы такого класса принято обозначать аббревиатурой ERP (Enterprise resource planning). Такие системы направлены на управление ресурсами, имеющимися у организации или, в более общем смысле, искусственной системы.

В отличие от различных подходов к управлению ресурсами контроллер обеспеченности ресурсами играет роль информатора о необходимых ресурсах. Под ресурсами в данной работе понимаются такие их типы, как: природные возобновляемые и невозобновляемые, человеческие, информационно-правовые и другие.

Если внимательно взглянуть на последствия управленческого решения, становится очевидным, что воплощение любого решения так или иначе связано с расходом ресурсов. Когда ресурс находится под контролем искусственной системы и система располагает к нему беспрепятственным доступом, тогда с этой позиции ERP системы соответствуют роли контроллера обеспечения.

Наиболее интересным объектом является ресурс, к которому у системы нет непосредственного доступа. Хорошим примером является такое достаточно частое явление, как необходимость расширения производства. Под расширением здесь понимается как расширение ассортимента выпускаемой продукции, так и территориальное расширение искусственной системы. При такой постановке задачи изначально система может иметь лишь поверхностное представление о необходимых ресурсах, а в самом трудном случае и не иметь такового. Приняв решение о целесообразности расширения производства, лицо, принимающее решения (ЛПР), вынуждено организовывать различные мероприятия по обеспечению необходимой информацией. Стоимость таких мероприятий и время их проведения должны быть минимизированы.

Контроллер обеспечения ресурсами должен оперативно и в необходимом объеме предоставить информацию о том, какие ресурсы нужны для успешного воплощения идеи, какими ресурсами располагает система, где и в каком объеме можно получить доступ к недостающим ресурсам. По нашему мнению такой контроллер должен присутствовать в системе поддержки принятия решений.

2 Концептуальная модель контроллера

Сформулируем ключевые особенности контроллера: предоставление возможности конструирования запросов от пользователя СППР; обработка запросов на получение информации о ресурсах; представление результатов запроса в необходимом пользователю виде; поиск информации о ресурсах среди знаний, накопленных в СППР.

Перечисленные выше особенности позволяют пользователям СППР получить необходимую информацию о ресурсах. Перед подробным рассмотрением функций контроллера следует определить его место в СППР на основе онтологий.

Наиболее важным средством, которое предоставляет онтологический подход к проектированию СППР, являются знания. Контроллер должен занять свое место в СППР, а, следовательно, и получить доступ к знаниям, агрегированным в СППР, среди которых присутствует важная информация о ресурсах. Помимо того что контроллер получает доступ к информации о ресурсах, необходимой видится его сопровождающая часть, которая предоставляет две возможности: поиск информации и конструирование запросов.

2.1 Место контроллера в СППР

С концептуальной точки зрения – контроллер есть отдельный программный модуль в информационной системе. Но с пользовательской точки зрения на СППР контроллер представ-

ляет собой одну из ключевых частей. С нашей точки зрения контроллер занимает свое место между знаниями, хранящимися в СППР, и пользователем, тем самым представляя собой дополнительный инструмент для обработки большого количества данных с целью извлечения необходимой информации о ресурсах, и, в конечном счете, для принятия решения с использованием полученной информации.

Следует заметить, что контроллер не должен нарушать целостность структуры СППР, а тем более заменять собой всю совокупность методов и средств, которые используются для принятия решений.

2.2 Конструирование запросов

Понятие «запрос» в контроллере это аналог понятия запрос в терминах систем управления базами данных. СППР на основе онтологий можно рассмотреть с точки зрения таких систем, считая, что наиболее удобным способом представления знаний являются реляционные таблицы.

С учетом требований к СППР и потребностей современных искусственных систем становится очевидной необходимость создания таких информационных систем, которые на архитектурном уровне, обладают способностью к полиморфизму своих возможностей. При проектировании контроллера необходимо учитывать, что СППР должна быть гибкой, оперативно реагировать на потребности пользователя.

Конструирование запросов – одна из ключевых особенностей, которой, по мнению авторов, должен обладать такой контроллер. Пользователь СППР (а в частности контроллера) должен иметь возможность манипулировать структурой запросов. Такая функция позволяет не только получать информацию о тех ресурсах, которые потребляет искусственная система, но и в любой момент иметь информацию о любых других ресурсах, что несомненно полезно при решении различного рода задач прогнозирования.

Для того чтобы решить проблему идентификации типа запроса и желаемого результата, предлагается расширить это понятие, связав с основной конструкцией запроса его семантическое значение в виде набора ключевых слов.

2.3 Обработка запросов

С помощью конструирования запросов решается проблема гибкости контроллера. Но одновременно с этим существует проблема, которая не позволяет создавать универсальные запросы с точки зрения пользователя. Каждый пользователь СППР обладает субъективным взглядом на ресурсы и склонен формулировать свои запросы с использованием своего информационного набора представлений о ресурсах. При достаточно большом количестве ЛПР, количество запросов будет расти, а на каждом оперативном уровне будет увеличиваться неопределенность.

Чтобы избежать упомянутой сложности, предлагается ввести в контроллер технику «Поиск в найденном», с использованием которой каждый пользователь при необходимости осуществляет циклический поиск среди возможных запросов. Изначально пользователь формирует условие поиска, а затем повторяет поиск среди результатов предыдущего шага с возможностью изменять условия поиска. Заметим, что поиск ведется среди описаний запросов, что дополнительно увеличивает ответственность того, кто формирует эти описания. Задача автоматического составления описания запросов, представляющая самостоятельный интерес, выходит за рамки данной статьи.

2.4 Предоставление результатов

ЛПР в каждом конкретном случае, рассчитывая на СППР, ожидает получить те результаты, которые он сможет использовать для принятия решения. Такие пожелания пользователей СППР следует учитывать. Предлагается предоставить пользователю возможность выбирать тип отображаемой информации, вид представления и другие параметры.

Следует учитывать, что вместе с увеличением периода агрегации знаний в СППР будет расти количество информации о ресурсах. Знания и информация в СППР являются платформой для функционирования, и их не рекомендуется удалять, что приводит к появлению избыточной информации. Для решения проблемы избыточности предлагается использовать различные фильтры для отображения и представления результатов. С помощью таких фильтров у пользователя появится возможность выбирать из всего объема результатов только нужные ему.

2.5 Поиск информации

СППР на основе онтологий представляет собой динамическую информационную систему. С развитием искусственной системы будет развиваться и СППР. Такой процесс развития, помимо традиционных составляющих, будет сопровождаться ростом количества источников знаний. Новые знания должны быть исследованы на наличие в них информации о ресурсах. Такой процесс исследования достаточно прост и для решения этой задачи существуют адекватные средства и алгоритмы.

3 Модель работы контроллера в СППР

Как упоминалось выше, контроллер это программный модуль в СППР. Для описания работы контроллера следует уделить внимание следующим аспектам: процесс формирования запроса; процесс исполнения запроса; процесс отображения результатов; процесс поиска необходимых ресурсов.

Процесс формирования запроса происходит под управлением пользователя СППР. Начинается этот процесс с формализации требований пользователя. Пользователь в зависимости от своей квалификации формирует запрос доступными ему средствами. Сформированный запрос включает в себя формулировку желаемого результата. С технологической стороны запрос это либо единичный запрос SQL к базе данных, либо цикл SQL-запросов.

Процесс исполнения запроса не имеет специфических особенностей по сравнению с исполнением SQL-запроса. Разница лишь в том, что после исполнения результаты подвергаются дополнительным обработкам по требованию пользователя. Сам процесс может исполняться циклически, это зависит от первоначальной конфигурации запроса.

Процесс отображения результатов предполагает достаточно быстрое выполнение. Однако все зависит от количества отображаемой информации. Во время исполнения этого процесса будут отображаться результаты запроса в той форме, которую указал пользователь. Обязательно наличие временного классифицирования результатов.

Процесс поиска необходимых ресурсов следует запускать на исполнение в случае их недостаточности в доступном для искусственной системы месте. Такой процесс является достаточно значимым для всей СППР в целом и играет одну из ключевых ролей в контроллере. Проще говоря, этот процесс отвечает на вопрос о месте положения необходимых ресурсов, их количества. Резюмируя, можно утверждать, что такой процесс должен решать задачу поиска ресурсов, упомянутую выше.

4 Процесс принятия решения с использованием контроллера

Для иллюстрации процесса принятия решения с использованием контроллера сформулируем достаточно абстрактную задачу. Некоторая организация решает выпускать новую высокотехнологичную продукцию. Обладая знаниями о технологии производства, лица, ответственные за начало производства и выпуск готовой продукции, формируют список требований для успешной реализации этого проекта.

Среди этих требований необходимо выделить требования на различные ресурсы (в общем случае разноплановые). Допустим, что первоначально требуется удовлетворить потребности в

некотором количестве высококачественной нефти. Если предприятие само производит нефть требуемого качества, то задача упрощается, и решение проблемы ложится на руководителей. В противном случае, при наличии СППР с контроллером формируется запрос на поиск информации о поставщиках нефти и качестве их продукции. Среди множества вариантов следует выбрать приемлемый и послать поставщикам уведомление о намерениях закупить нефть.

Далее будем рассматривать только те ресурсы, которые находятся вне зоны свободного доступа организации. Допустим, что производству требуется наличие высококвалифицированного персонала с опытом работы в сфере высокотехнологичной обработки нефти. Формируется очередной запрос к контроллеру и получается, например, следующий ответ: «Необходимого персонала в организации нет. Следует обратиться в отдел кадров с поручением найти соответствующий персонал».

Отметим здесь, что СППР должна предоставлять возможность каждому пользователю или ЛПР возможность самостоятельной работы с системой. Принципиальное отсутствие посредников оправдано персональной ответственностью за принятые решения и необходимостью в полноте информации для решения поставленных задач.

5 Требование к СППР

Следует иметь в виду, что в СППР должна содержаться достаточно большая информация о поставщиках различных ресурсов, такая информация должна быть упорядочена.

Здесь следует сформулировать ключевое требование к СППР. СППР должна проектироваться с расчетом на то, что в ней будет храниться информация о поставщиках ресурсов. Наиболее очевидное решение – это создание некоторого реестра поставщиков, который будет постоянно обновляться и поддерживаться в актуальном состоянии. Количество вариантов организации подобного реестра и механизмов работы с ним достаточно велико.

Очевидна взаимная заинтересованность поставщиков и потребителей ресурсов в создании такого реестра. Наличие хранилища данных с информацией о поставщиках, к которому у контроллера будет регламентированный доступ, поможет поставщикам выбрать необходимую стратегию сбыта своих ресурсов, а потребителям найти оптимальный вариант удовлетворения своих ресурсных потребностей.

Заключение

Сформулируем основные результаты работы: описана концептуальная модель и ключевые особенности контроллера; сформулировано основное требование к СППР, в которой предполагается наличие контроллера.

Создание контроллера обеспеченности ресурсами достаточно трудоемкий процесс, но конечный результат способен упростить процесс принятия решений. Необходимо помнить также, что для успешного и наиболее рационального решения, следует применять предлагаемый инструмент в совокупности с другими методами поддержки принятия решений.

Список литературы

- [1] Гаврилова Т.А., Муромцев Д.И. Интеллектуальные технологии в менеджменте: инструменты и системы // СПб.: Изд-во «Высшая школа менеджмента»; Издат. дом С.-Петербур. гос. ун-та, 2008.
- [2] Гаскаров Д.В. Интеллектуальные информационные системы. – М.: Высшая школа, 2003.
- [3] Ситников П.В. Разработка системы информационно-аналитической поддержки процессов принятия решений при управлении группой энергосервисных компаний // Проблемы управления и моделирования в сложных системах: Труды XI Международной конференции. – Самара: СамНЦ РАН, 2009. – С. 515-519.
- [4] Виттих В.А. Парадигма ограниченной рациональности принятия решений // Препринт.- Самара: ИПУСС РАН, 2009.

СПОСОБ ИССЛЕДОВАНИЯ ОНТОЛОГИЧЕСКИХ МОДЕЛЕЙ ПРОИЗВОДСТВЕННЫХ СИСТЕМ НА ОСНОВЕ ПРИМЕНЕНИЯ ГРАФОВЫХ МЕТОДИК

П.В. Ситников, Е.Г. Суворова

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
SitnikovPV@ya.ru
тел./факс: +7 (846) 333-27-70

Ключевые слова: *социотехническая система, процессы управления, онтологический анализ, онтологическая модель, семантические графы, теория графов, производственная система*

Abstract

The article describes the method of analysis of ontological models by management of complex systems based on application of the graphs theory.

Введение

В производственных системах в процессах управления зачастую оказывается недостаточным использование только информационных составляющих соответствующих информационно-аналитических систем поддержки принятия решений (ИАС). Для принятия более эффективных, качественных и оперативных решений необходимо наличие аналитического блока в ИАС.

В качестве основы использования всех аналитических методик в ИАС предлагается использовать онтологические модели, представляющие собой сети связанных объектов [1]. При расчетах различных задач, возникающих при принятии решений, информация берется из базовых онтологических моделей и с помощью специальных методик происходит поиск вариантов оптимального решения [2].

В данной работе в качестве примера таких методик приводится описание способов использования арсенала теории графов (см., например, [3-8]) для исследования онтологических моделей (по сути – графов) в ИАС при решении разнообразных задач управления сложными системами [1].

Способы применения указанных методик, предлагаемые в данной работе, реализуются на онтологических моделях информационно-аналитической системы поддержки принятия решений при управлении группой энергосервисных кампаний [9].

1 Задачи поиска кратчайших и длиннейших путей на графах

Задача о кратчайшем пути. Пусть задан ориентированный граф из $n+1$ вершины, в котором выделены две вершины – вход (нулевая вершина) и выход (вершина с номером n). Для каждой дуги заданы числа, называемые длинами дуг. Длиной пути (контур) называется сумма длин входящих в него дуг (если длины дуг не заданы, то длина пути (контур) определяется как число входящих в него дуг). Задача заключается в поиске кратчайшего пути (пути минимальной длины) от входа до выхода графа. Известно, что для существования кратчайшего пути необходимо и достаточно отсутствия в сети контуров отрицательной длины. Предположим, что в сети нет контуров. Тогда всегда можно пронумеровать вершины таким образом, что для любой дуги (i, j) имеет место $j > i$. Такая нумерация называется правильной. Легко показать, что в сети без контуров всегда существует правильная нумерация. Обозначим l_{ij} – длину дуги

$(i; j)$. Кратчайший путь в сети, имеющей правильную нумерацию, определяется следующим алгоритмом.

Алгоритм 1.

Шаг 0: Помечаем нулевую вершину индексом $\lambda_0 = 0$;

Шаг k : Помечаем вершину k индексом $\lambda_k = \min_{i < k} \lambda_i + l_{ik}$;

Индекс выхода λ_n будет равен длине кратчайшего пути.

Когда индексы (называемые в некоторых задачах *потенциалами вершин*) установятся, кратчайший путь определяется методом обратного хода от выхода к входу, т.е. кратчайшим является путь $\mu = (0; i_1; i_2; \dots; i_{n-1}; n)$, такой, что $l_{i_{n-1}n} = \lambda_n - \lambda_{i_{n-1}}$ и т.д.

Решим задачу поиска кратчайшего пути карьерного роста от директора по производству до генерального директора предприятия (граф отфильтрован из онтологической модели описания системы управления холдинга, описанной в [8]). За вершины графа примем должности, введя следующие обозначения (см. рисунок 1):

- 0 – директор по производству;
- 1 – главный инженер;
- 2 – исполнительный директор;
- 3 – начальник управления организацией производства;
- 4 – заместитель технического директора;
- 5 – технический директор;
- 6 – генеральный директор.

При данных обозначениях нумерация вершин в сети будет правильной, и мы можем применить алгоритм 1.

За длины дуг примем процент препятствий на пути к новой должности, который связан с квалификационными требованиями, предъявляемыми к лицу, замещающему определенную должность. Чем «длиннее дуга» – тем больше квалификационный разрыв между должностями в промышленной организации.

На рисунке 1 приведен пример применения алгоритма 1 для определения кратчайшего пути (числа у дуг равны длинам дуг, индексы вершин помещены в квадратные скобки, кратчайший путь выделен двойными линиями).

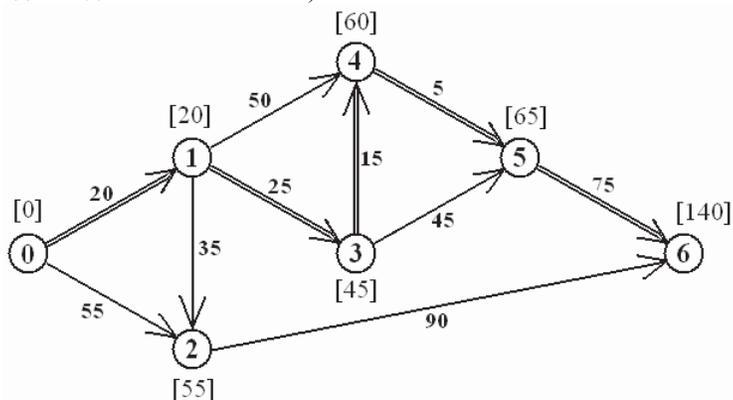


Рисунок 1 – Пример поиска кратчайшего пути на графе в задаче «о карьерном росте»

Начнем с определения индексов вершин: $\lambda_0 = 0$, $\lambda_1 = \lambda_0 + l_{01} = 20$, $\lambda_2 = \min\{\lambda_0 + l_{02}, \lambda_1 + l_{12}\} = 55$, $\lambda_3 = \min\{\lambda_0 + l_{03}, \lambda_1 + l_{13}, \lambda_2 + l_{23}\} = 45$ (не $\exists l_{03}$ и l_{23}), $\lambda_4 = 60$, $\lambda_5 = 65$, $\lambda_6 = 140$.

Длина кратчайшего пути равна индексу $\lambda_6 = 140$.

Индексы вершин определены, теперь найдем кратчайший путь $\mu = (0; i_1; i_2; \dots; i_5; 6)$.

Вершина i_5 должна быть такая, что $l_{i_5,6} = \lambda_6 - \lambda_5 = 75$, дуга с таким значением, ведущая в вершину 6, существует $\Rightarrow i_5 = 5$. Аналогично $l_{i_4,5} = \lambda_5 - \lambda_4 = 5 \Rightarrow i_4 = 4$ и $l_{i_3,4} = \lambda_4 - \lambda_3 = 15 \Rightarrow i_3 = 3$. Находим i_2 : $l_{i_2,3} = -10$, но такой дуги не существует, следовательно, и вершины i_2 в кратчайшем пути нет. Определим i_1 : $l_{i_1,2} = 35 \Rightarrow i_1 = 1$. Итак, кратчайшим является путь $\mu = (0; 1; 3; 4; 5; 6)$.

Следующий алгоритм дает возможность определять кратчайший путь в общем случае (т.е. при произвольной нумерации вершин).

Алгоритм 2 (алгоритм Форда).

Шаг 0: Помечаем нулевую вершину индексом $\lambda_0 = 0$, все остальные вершины индексами $\lambda_i^0 = +\infty, i = \overline{1, n}$;

Шаг k : Рассматриваем все дуги. Если для дуги $(i; j)$, $\lambda_j^{k-1} - \lambda_i^{k-1} > l_{ij}$, то вычисляем новое значение $\lambda_j^k = \lambda_i^{k-1} + l_{ij}$.

Индексы устанавливаются за конечное число шагов. Обозначим $\{\lambda_i^*\}$ – установившиеся значения индексов, которые обладают следующим свойством: величина λ_i^* равна длине кратчайшего пути из нулевой вершины в вершину i . Кратчайший путь из вершины 0 в вершину i определяется методом обратного хода.

Пример данного алгоритма приведен далее в задаче о назначении.

Если длины всех дуг неотрицательны, то для поиска кратчайшего пути применим алгоритм 3:

Шаг 0: Помечаем нулевую вершину индексом $\lambda_0 = 0$;

Шаг k : Пусть уже помечено некоторое множество вершин. Обозначим Q – множество непомеченных вершин, смежных с помеченными. Для каждой вершины $k \in Q$ вычисляем величину $\xi_k = \min(\lambda_i + l_{ik})$, где минимум берется по всем помеченным вершинам i , смежным с вершиной k . Помечаем вершину k , для которой величина ξ_k минимальна, индексом $\lambda_k = \xi_k$.

Подобную процедуру повторяем до тех пор, пока не будет помечена вершина n . Длина кратчайшего пути равна λ_n , а сам кратчайший путь определяется так, как это было описано выше.

Аналогично задаче о кратчайшем пути формулируется и решается задача о максимальном (длиннейшем) пути – достаточно изменить знаки дуг на противоположные и решить задачу о кратчайшем пути. Для существования решения задачи о максимальном пути необходимо и достаточно отсутствия контуров положительной длины.

В задаче поиска пути максимальной надежности длины дуг интерпретируются, например, как вероятности того, что существует связь между соответствующими двумя пунктами. Заменяя длины дуг их логарифмами, взятыми с обратными знаками, получаем, что путь максимальной надежности в исходном графе будет соответствовать кратчайшему пути в новом графе.

2 Задача о назначении

В качестве других примеров приложения теории графов для содержательного анализа онтологических моделей можно привести такие, как: поиск контура минимальной длины, поиск контура минимальной средней длины, поиск пути максимальной эффективности, задачи о максимальном потоке, транспортные задачи.

Частным случаем транспортной задачи является задача о назначении, заключающаяся в следующем: имеются n человек (работников), которые могут выполнять различные работы (занимать различные должности), число работ равно числу работников (введя фиктивные должности и/или фиктивные работы, всегда можно незамкнутую задачу привести к рассматриваемой замкнутой форме). Известны затраты s_{ij} на назначение i -го работника на j -ю должность (например, минимальная зарплата, за которую он согласится работать на этой должности). Требуется найти назначение работников на должности (каждого работника на одну и только одну должность), минимизирующее суммарные затраты (если s_{ij} интерпретируется как эффективность от работы i -го работника на j -ой должности, то оптимальное назначение должно максимизировать суммарную эффективность).

Известно множество методов решения задачи о назначении. Рассмотрим один из них.

Алгоритм 4.

Шаг 0. Назначаем каждого человека на самую дешевую для него работу, т.е. положим

$$x_{ij}^0 = \begin{cases} 1, & \text{если } s_{ij} = \min_k s_{ik} \\ 0, & \text{в противном случае} \end{cases}.$$

Если при этом назначение является допустимым (т.е. все работы выполняются), то решение получено. Если имеется «дисбаланс», т.е. не все работы выполняются ($\exists j_1 : \sum_{i=1}^n x_{ij_1}^0 > 1$), то переходим к следующему шагу.

Шаг k . Введем два подмножества множества дуг: $P_1 = \{(i; j) \mid x_{ij} = 1\}$ и $P_2 = \{(i; j) \mid x_{ij} = 0\}$.

Примем множество вершин-работ, на которых назначено несколько работников за вход сети, множество вершин-работ, которые не выполняются – за выход сети. Изменим направления дуг из множества P_1 на обратные и примем их длины равными $(-s_{ij})$, длины дуг из множества P_2 примем равными S_{ij} . Найдем путь μ^k минимальной длины в полученной сети.

$$\text{Далее полагаем } x_{ij}^k = \begin{cases} x_{ij}^{k-1}, & \text{если } (i; j) \notin \mu^k \\ 1 - x_{ij}^{k-1}, & \text{если } (i; j) \in \mu^k \end{cases}.$$

На каждом шаге число «дисбалансов» уменьшается на единицу. Следовательно, число шагов алгоритма не превышает числа «дисбалансов», которое конечно.

В качестве примера использования данного алгоритма можно привести следующую задачу, знания для решения которой отфильтрованы из созданных онтологических моделей, описывающих систему организации производственных работ по ремонту турбоагрегата на Сакмарской ТЭЦ г. Оренбурга (см. [8]). Пусть имеются $n = 3$ организации и столько же работ. Каждая организация может выполнить любую из данных работ за определенную стоимость. В целях экономии времени мы хотим сделать заказ каждой организации только на один вид работ, так чтобы наши затраты были минимальны. Введем обозначения, которые будут удобны при решении поставленной задачи.

Организации:

1 – ОАО «Оренбургэнергоремонт»;

2 – ОАО «Теплоэнергосервис»;

4 – ООО «ТД Энергосервис».

Работы:

0 – ремонт фундамента опор;

3 – модернизация ЦНД;

5 – ремонт и модернизация РНД.

Нам известна стоимость (в млн. руб.) выполнения данных работ в каждой организации, и

матрица затрат имеет вид:

	0	3	5
1	1.5	2.2	3.1
2	1.9	2.4	3.4
4	2.1	2	3.5

Заказываем каждой организации самую дешевую для нее работу (заказы выделены на рисунке 2 тонкими дугами), получаем:

$$\begin{aligned}
 x_{10}^0 &= 1 & x_{20}^0 &= 1 & x_{40}^0 &= 0 \\
 x_{13}^0 &= 0 & x_{23}^0 &= 0 & x_{43}^0 &= 1 \\
 x_{15}^0 &= 0 & x_{25}^0 &= 0 & x_{45}^0 &= 0
 \end{aligned}$$

Имеется «дисбаланс», ($\exists j_1 = 0 : \sum_{i=1}^n x_{ij_1}^0 = 2 > 1$), т.е. не все работы выполняются, пере-

ходим к следующему шагу. $P_1 = \{(1; 0), (2; 0), (4; 3)\}$, P_2 – множество всех остальных дуг, примем за вход сети вершину-работу 0, а за выход сети – вершину-работу 5. Изменим направления дуг из множества P_1 на обратные и примем их длины равными $(-s_{ij})$, длины дуг из множества P_2 примем равными s_{ij} (см. рисунок 3).

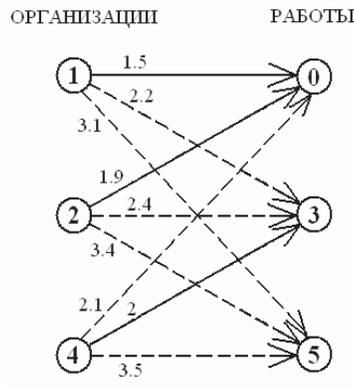


Рисунок 2 – Пример задачи о назначении, решаемой на основе фрагмента онтологической модели организации системы производственных работ по ремонту турбоагрегата на Сакмарской ТЭЦ г. Оренбурга

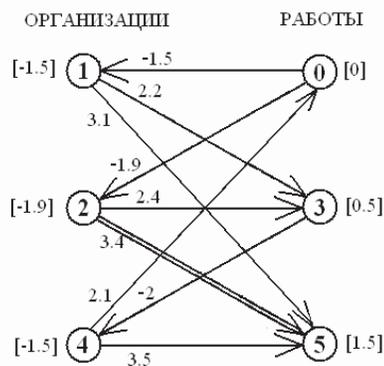


Рисунок 3 – Пример кратчайшего пути в задаче о назначении, решаемой на основе фрагмента онтологической модели организации системы производственных работ по ремонту турбоагрегата на Сакмарской ТЭЦ г. Оренбурга

Найдем путь μ^1 минимальной длины в полученной сети (потенциалы вершин, вычисляемые при нахождении кратчайшего пути в рассматриваемом примере, приведены в квадратных скобках).

Для нахождения кратчайшего пути воспользуемся алгоритмом Форда. Вычислим индексы вершин.

Шаг 0: $\lambda_0^0 = 0, \lambda_i^0 = +\infty, i = \overline{1,5}$.

Шаг 1: $\lambda_1^1 = -1.5, \lambda_2^1 = -1.9, \lambda_i^1 = +\infty, i = \overline{3,5}$.

Шаг 2: $\lambda_3^2 = 0.5, \lambda_5^2 = 1.5$, остальные индексы не изменятся.

Шаг 3: $\lambda_4^3 = -1.5$, остальные индексы не изменятся.

Теперь известны потенциалы всех вершин, находим кратчайший путь $\mu^1 = (0; i_1; \dots; i_4; 5)$ методом обратного хода: $l_{i_4,5} = 3$ и вершины i_4 не существует, $l_{i_5,5} = 1 \Rightarrow$ вершины i_3 не существует, $l_{i_2,5} = 3.4 \Rightarrow i_2 = 2, l_{i_1,2} = -0.4 \Rightarrow$ не существует вершины i_1 , и кратчайший путь имеет вид: $\mu^1 = (0; 2; 5)$.

В результате в рассматриваемом примере за один шаг получим оптимальное назначение, отличающееся от найденного на нулевом шаге тем, что организации 2 заказывается работа 5 (см. дугу, обозначенную двойными линиями на рисунке 3).

Аналогичным способом можно решить любую транспортную задачу (искать кратчайший путь из множества объектов, в которые доставили товара больше, чем требуется, во множество объектов, где товара не хватает).

3 Задачи календарно-сетевого планирования и управления

Онтологические модели могут быть естественным информационным базисом задач календарно-сетевого планирования и управления [10-12].

Рассмотрим производственный проект, состоящий из набора операций (работ). Технологическая зависимость между операциями задается в виде сети (сетевого графика). При этом дуги сети соответствуют операциям, а вершины – событиям (моментам окончания одной или нескольких операций). Для каждой операции $(i; j)$ задана ее продолжительность t_{ij} .

Для примера рассмотрим задачу определения продолжительности проекта (управление временем). Проиллюстрируем возможность использования онтологических моделей как информационного базиса задачи определения продолжительности проекта.

Продолжительность проекта определяется путем максимальной длины, называемым критическим путем. Методы поиска пути максимальной длины описаны выше. Операции, принадлежащие критическому пути, называются критическими. Остальные (некритические) операции имеют резерв времени, характеризуемый максимальной задержкой операции, при которой продолжительность проекта не изменяется. Критические операции имеют нулевой резерв. Приведем соответствующие формулы.

Алгоритм 5.

Предположим, что выполнение комплекса операций (проекта) начинается в нулевой момент времени. Обозначим Q_0 – множество событий, не требующих выполнения ни одной из операций, т.е. входы сети с правильной нумерацией; Q_i – множество событий, непосредственно предшествующих событию i , т.е. множество вершин j сети, для которых существует дуга $(j; i)$.

Положим

$$(1) t_i^- = \max_{j \in Q_0} t_{ji}, \quad t_i^- = \max_{j \in Q_i} (t_j^- + t_{ji}).$$

Величина t_i^- называется ранним моментом (временем) свершения i -го события и характеризует время, раньше которого это событие произойти не может.

Длина критического пути

$$(2) T = \max_i t_i^-$$

определяется ранним временем свершения конечного события, т.е. события, заключающегося в завершении всех операций.

Поздним моментом t_i^+ свершения события называется максимальное время его наступления, не изменяющее продолжительности проекта. Обозначим R_i – множество событий, непосредственно следующих за событием i , т.е. множество вершин j сети, для которых существует дуга $(i; j)$. Вычислим для каждой вершины-события i длину l_i максимального пути от этой вершины до выхода сети – события, заключающегося в завершении всего комплекса операций:

$$(3) l_i = \max_{j \in R_i} (l_j + t_{ij}).$$

Положим $t_i^+ = T - l_i, i = \overline{1, n}$.

Для завершения проекта за время T необходимо и достаточно, чтобы событие i произошло не позднее момента $t_i^+, i = \overline{1, n}$.

Полным резервом Δt_i события i называется разность между его поздним и ранним моментами свершения, т.е.

$$(4) \Delta t_i = t_i^+ - t_i^-, i = \overline{1, n}.$$

Полный резерв критических событий (событий, принадлежащих критическому пути) равен нулю.

Приведем порядок применения описанного алгоритма для расчета продолжительности проекта по ремонту турбоагрегата на Сакмарской ТЭЦ Оренбургской области (на основе сетевой график, отфильтрованного из соответствующей онтологической модели [8]).

Вершины графа обозначают этапы выполнения работ. Ведем следующие обозначения дуг:

- (0,1) обозначает подготовку ТЗ по заключению договора генподряда «под ключ»;
- (1,2) – проведение торгов на выбор генподрядчика;
- (2,3) – заключение договора «под ключ» с генподрядчиком;
- (3,4) – разработка ТЭО (проект);
- (3,5) – поставка подогревателей низкого давления;
- (3,6) – поставка ЦВД в сборе;
- (4,6) – модернизация ЦНД, проточной части и РНД;

- (4,7) – экспертиза, согласование и утверждение проекта;
- (5,6) – поставка подогревателя сетевой воды;
- (6,7) – разработка и поставка АСУ ТП;
- (7,8) – строительные-монтажные работы;
- (8,9) – пуско-наладочные работы;
- (8,10) – обучение персонала;
- (9,10) – комплексное опробование оборудования.

Нам известно время выполнения (в днях) каждой из этих работ (рисунок 4), определим продолжительность всего проекта, воспользуемся для этого алгоритмом 5. Данная нумерация вершин является правильной, значит, применять алгоритм можно.

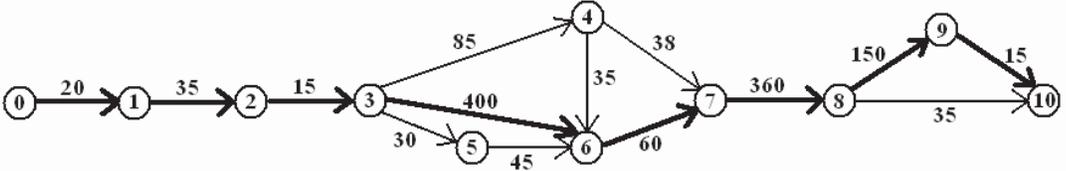


Рисунок 4 – Задача определения продолжительности проекта

Найдем множество $Q_i, i = \overline{0,10}$:

$$Q_0 = \{0\}, Q_1 = \{0\}, Q_2 = \{1\}, Q_3 = \{2\}, Q_4 = \{3\}, Q_5 = \{3\},$$

$$Q_6 = \{3,4,5\}, Q_7 = \{4,6\}, Q_8 = \{7\}, Q_9 = \{8\}, Q_{10} = \{8,9\}.$$

Подсчитаем ранние моменты свершения i -х событий, т.е. время, раньше которого эти события произойти не могут:

$$t_1^- = t_{01} = 20, t_2^- = t_1^- + t_{12} = 55, t_3^- = 70, t_4^- = 155, t_5^- = 100,$$

$$t_6^- = \max\{t_3^- + t_{36}, t_4^- + t_{46}, t_5^- + t_{56}\} = 470, t_7^- = 530, t_8^- = 890,$$

$$t_9^- = 1040, t_{10}^- = 1055.$$

Длина критического пути $T = \max_i t_i^- = 1055$.

Найдем множество $R_i, i = \overline{0,9}$:

$$R_0 = \{1\}, R_1 = \{2\}, R_2 = \{3\}, R_3 = \{4,5,6\}, R_4 = \{6,7\},$$

$$R_5 = \{6\}, R_6 = \{7\}, R_7 = \{8\}, R_8 = \{9,10\}, R_9 = \{10\}.$$

Вычислим для каждой вершины i длину l_i максимального пути от этой вершины до выхода сети – события, заключающегося в завершении всего комплекса операций:

$$l_{10} = 0, l_9 = 15, l_8 = 165, l_7 = 525, l_6 = 585, l_5 = 630,$$

$$l_4 = 620, l_3 = 985, l_2 = 1000, l_1 = 1035, l_0 = 1055.$$

Теперь найдем поздние моменты свершения i -х событий, т.е. максимальное время их наступления, не изменяющее продолжительность проекта:

$$t_1^+ = T - l_1 = 20, t_2^+ = 55, t_3^+ = 70, t_4^+ = 435, t_5^+ = 425,$$

$$t_6^+ = 470, t_7^+ = 530, t_8^+ = 890, t_9^+ = 1040, t_{10}^+ = 1055.$$

Для завершения проекта за время $T = 1055$ необходимо и достаточно, чтобы событие i произошло не позднее момента $t_i^+, i = \overline{1,10}$.

Полные резервы Δt_i событий i равны соответственно:

$$\Delta t_1 = t_1^+ - t_1^- = 0, \quad \Delta t_2 = 0, \quad \Delta t_3 = 0, \quad \Delta t_4 = 280, \quad \Delta t_5 = 325,$$

$$\Delta t_6 = 0, \quad \Delta t_7 = 0, \quad \Delta t_8 = 0, \quad \Delta t_9 = 0, \quad \Delta t_{10} = 0.$$

Полные резервы событий, принадлежащих критическому пути (он выделен на рисунке жирными линиями) равны нулю.

Заключение

Разнообразные информационно-аналитические системы производственных систем нуждаются в аналитическом блоке для обеспечения лиц, принимающих решения релевантной и оперативной информацией.

Все аналитические методики и программные приложения предложено реализовывать используя онтологические информационно-логические модели.

В работе предложен один из возможных аналитических способов исследования онтологических моделей, основанный на теории графов, показана возможность и работоспособность его использования в группе производственных энергосервисных предприятий.

Список литературы

- [1] Виттих В.А., Ситников П.В., Смирнов С.В. Онтологический подход к построению информационно-логических моделей в процессах управления социальными системами // Вестник компьютерных и информационных технологий. – 2009. – №5. – С. 45-53.
- [2] Ситников П.В., Смирнов С.В. Методы и средства анализа объектных информационно-логических моделей // Проблемы управления и моделирования в сложных системах: Труды X Международной конференции. – Самара: СамНЦ РАН, 2008. – С. 448-456.
- [3] Апанович З.В. Средства для работы с графами большого объема: построение и оптимизация компоновочных планов // Системная информатика: Сб. науч. тр. Вып. 10: Методы и модели современного программирования. – Новосибирск: Изд-во СО РАН, 2006. – С. 7-58.
- [4] Дилигенский Н.В., Цапенко М.В. Многокритериальная оценка сравнительной эффективности организационных систем управления // Проблемы управления и моделирования в сложных системах: Труды VIII Международной конференции. – Самара: СамНЦ РАН, 2006. – С. 66-72.
- [5] Касьянов В.Н., Евстигнеев В.А. Графы в программировании: обработка, визуализация и применение. – СПб.: БХВ-Петербург, 2003. – 1104 с.
- [6] Кирсанов М.Н. Математический центр московского метро // Exponenta Pro. – 2003. – №4. – С. 60-62.
- [7] Попов И.В., Фролкина Н.А. Анализ связанных объектов и визуализация результатов // Труды международной конференции «Диалог'2004 «Компьютерная лингвистика и интеллектуальные технологии»: [Электронный документ]. – 2004. (<http://www.dialog-21.ru/Archive/2004/Popov.htm>). Дата обращения: 11.03.2010.
- [8] Тульчинский В.Г., Ющенко А.К., Ющенко Р.А. Графовые запросы для интеграции данных посредством XML // Кибернетика и системный анализ. – 2008. – Т. 44. – №2. – С. 171-183.
- [9] Ситников П.В. Разработка системы информационно-аналитической поддержки процессов принятия решений при управлении группой энергосервисных компаний // Проблемы управления и моделирования в сложных системах: Труды XI Международной конференции. – Самара: СамНЦ РАН, 2009. – С. 515-519.
- [10] Анфилатов В.С., Емельянов А.А., Кукушкин А.А. Системный анализ в управлении. – М.: Финансы и статистика, 2005. – 368 с.
- [11] Попов В.Н., Касьянов В.С., Савченко И.П. Системный анализ в менеджменте. – М.: КНОРУС, 2007. – 304 с.
- [12] Смирнов С.В. Эффективный подход к моделированию продолжительности работ в сетевом планировании и управлении // Проблемы управления и моделирования в сложных системах: Труды VII Международной конференции. – Самара: СамНЦ РАН, 2005. – С. 117-126.

ОСНОВНЫЕ АСПЕКТЫ ФОРМИРОВАНИЯ ВНЕШНЕЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ СОВРЕМЕННОЙ КОМПАНИИ

Е.Г. Серова

Санкт-Петербургский Государственный Университет
Высшая школа менеджмента
199155, Санкт-Петербург, пер. Декабристов, 16, Россия
serovah@gmail.com
тел: +7 (812) 350-81-55, факс: +7 (812) 350-04-06

***Ключевые слова:** менеджмент, компьютерное моделирование, информационная инфраструктура компании, системы управления взаимоотношениями с клиентами, мультиагентные системы*

Abstract

Companies applying information, communication technologies and computer simulation modeling tools are paying more and more attention to the ability to adaptive and hybrid architectures when building and developing information infrastructure. This paper studies an approach to forming an external information infrastructure for a company with a developed clientele by joining opportunities of CRM technologies and agents. Companies can succeed in carrying out their development strategies by following modern trends, adding intellectual information tools to CRM systems and creating an adaptive hybrid for external infrastructure based on the multi-agent approach.

Введение

На современном этапе развития России использование информационно-коммуникационных технологий (ИКТ) в управлении, в том числе методов и инструментов компьютерного моделирования, выступает ключевым фактором повышения эффективности основного бизнеса: улучшения качества продукции и услуг, экономии трудовых и материальных затрат, роста производительности труда, совершенствования организации производства. В промышленно развитых странах, несмотря на значительно более благоприятные параметры информатизации бизнеса, вопросы выбора и внедрения современных информационных систем и инструментальных средств бизнес-моделирования, отвечающих требованиям рынка и стратегическим задачам бизнеса, также находятся в центре внимания. Необходимо также отметить тот факт, что акцент смещается в сторону развития внешней информационной инфраструктуры (далее сокращение ИИ) компаний, улучшения взаимоотношений с клиентами (CRM - Customer Relationship Management), и продолжает усовершенствоваться класс интеллектуальных информационных технологий (далее – ИТ).

Основное назначение использования средств моделирования при решении бизнес-задач – обеспечение взаимопонимания на всех уровнях организации, преодоление разрыва между стратегическим видением бизнеса и его реализацией на практике. Одним из возможных путей решения этой задачи является применение мультиагентных систем (MAS – MultiAgent Systems), получивших стремительное развитие за последние десять лет. С этой целью в современных средствах моделирования, применяемых в бизнесе, используются специальное программное обеспечение, языки и системы, с помощью которых разрабатываются модели, демонстрирующие, как построены в компании бизнес-процессы, как организовано взаимодействие между людьми и, что необходимо изменить для оптимизации архитектуры организации в целом.

Исследовательский характер данной работы заключается в изучении применения технологии мобильных агентов для расширения возможности CRM в условиях развития клиентской сети и реализации адаптивной информационной инфраструктуры на основе комплексного подхода. Необходимо иметь в виду, что если клиентоориентированный подход ведения бизнеса распространен и нередко используется как элемент стратегии компании, концепция программных агентов относится к числу развивающихся и инновационных, а сама идея интеграции методологий MAS и CRM достаточно нова.

Актуальность изучения данной проблемы не представляет сомнения, так как в настоящее время все больше компаний сталкиваются с необходимостью совершенствования управления клиентской сетью и осознают необходимость использования для этих целей современных информационно-коммуникационных технологий и подходов.

Цель проведенного исследования – проанализировать применимость и адаптируемость совокупности технологий CRM и MAS, а также возможность формирования внешней информационной инфраструктуры компании на основе комплексного подхода в условиях развития клиентской сети.

Для достижения поставленной цели одной из основных задач проведенного исследования является изучение понятия внешней информационной инфраструктуры компании, современных технологий ее реализации и возможность построения адаптивной инфраструктуры с использованием интеграции мультиагентных систем и систем управления взаимоотношениями с клиентами.

1 Информационная инфраструктура и ее роль в поддержке деятельности компании

В современной экономике информационные технологии действительно являются органичной частью передовых систем управления производством, товарно-материальными запасами и коммуникациями. Кроме того, эти технологии содействуют глобализации бизнеса, поскольку обеспечивают с минимальными затратами быстрый доступ к работникам компании по всему миру, а также координацию глобальных взаимодействий фирм, находящихся на разных звеньях цепочки создания ценности [4].

ИКТ не просто ведут к повышению эффективности деятельности компании, а могут расцениваться как значимый нематериальный актив, причем максимальный эффект будет достигнут в том случае, если они успешно интегрируются в единую информационную структуру компании. ИИ должна не только отвечать целям бизнеса, формироваться в соответствии с основными задачами его реализации, но и развиваться в соответствии с развитием бизнес-процессов организации и даже опережать бизнес в развитии.

К основным аспектам управления, поддерживаемым информационной инфраструктурой относятся:

- Информационная поддержка принятия управленческих решений;
- Автоматизация и оптимизация производственных процессов;
- Эффективное использование информации;
- Взаимная интеграция информационных подсистем;
- Взаимодействие с заказчиками и поставщиками;
- Выход на новые рынки сбыта.

Развитая внутренняя информационная инфраструктура более характерна для производственных предприятий, нацеленных на качество создаваемого продукта, и ее создание представляет собой больше техническую и технологическую проблему.

Создание же внешней инфраструктуры, по сравнению с внутренней, является в первую очередь стратегической, социальной и организационной проблемой. Внешняя инфраструктура

характеризуется большей динамикой, подвижностью, стихийностью, риском и непредсказуемостью и более характерна для предприятий, оказывающих услуги и продающих товары. Такие предприятия, как правило, имеют большую клиентскую сеть и для них актуальны проблемы, связанные с управлением взаимоотношениями с клиентами.

Следует выделить три основных аспекта построения внешней информационной инфраструктуры: стратегический, социальный и организационный.

Формируя ИИ, компания должна в первую очередь учитывать стратегическое направление своего развития, исходя из специфики деятельности и того влияния, которое она оказывает на общество. Только затем определяется технологическое решение, отвечающее поставленным целям и задачам. Стратегический аспект построения внешней информационной инфраструктуры заключается в том, что ее формирование должно быть направлено не на автоматизацию отдельных бизнес-процессов и решение текущих оперативных задач, а должно рассматриваться как комплексная и долгосрочная стратегия развития компании. Одним из основных направлений развития стратегии фирмы может стать клиентоориентированный подход, реализуемый с помощью технологий CRM.

Социальный аспект построения внешней информационной инфраструктуры выражается в том, что рассматриваемая инфраструктура может использоваться как сотрудниками организации, так и оказывать целенаправленное влияние на клиентов – не только реальных, но и потенциальных, а, следовательно, и на общество, как потребителя продуктов и услуг компании. В качестве простого и очевидного примера уместно привести влияние деятельности на социальные аспекты жизни общества таких известных, ориентированных на массового пользователя компаний, как IBM, Oracle, Microsoft. Продукция этих компаний глубоко вошла в жизнь общества и, следовательно, построение информационной инфраструктуры носит социальный характер или, другими словами, имеет место социо-инфраструктура.

Организационный аспект построения внешней информационной инфраструктуры во многом связан со стратегическим и с социальным (например, ответственность перед сотрудниками за принимаемые решения) и заключается в том, что инфраструктура должна быть интегрирующим звеном в цепи взаимоотношений компании, ее сотрудников и общества.

Для менеджеров высшего и среднего уровней управления в качестве целей формирования ИИ служат повышение эффективности бизнеса, надежность принятия решений, обеспечение оперативной связи с поставщиками и клиентами, внедрение и использование информационных систем и информационно-коммуникационных технологий, организация системы информационных услуг и обеспечение максимального ROI. Для сотрудников компании на оперативном транзакционном уровне управления эффективная информационная система должна обеспечивать постоянный доступ к необходимой корпоративной и персональной информации, взаимодействие с БД и приложениями в масштабе реального времени (порталы подразделений, intranet), мобильный доступ и удобство работы.

Создание информационной инфраструктуры, использующей ИКТ, как правило, включает в себя следующий комплекс мероприятий:

- Организационные мероприятия (определение правил разработки, ответственности, финансирования, структур документов);
- Технологические мероприятия (выбор средств проектирования, установка программного обеспечения, организация маршрутизации, безопасности и сопровождения);
- Технические мероприятия (приобретение, установка и техническое обеспечение эксплуатации программно-аппаратных средств и оборудования);
- Кадровые мероприятия (обучение сотрудников).

Наиболее часто встречаемый в России путь создания ИИ – это часто изолированное информационное обеспечение отдельных процессов и видов деятельности. Современный и эффективный подход формирования информационной инфраструктуры – принятие комплексно-

го ИТ-решения, создание адаптивной ИИ на базе высокотехнологичных средств визуализации, обработки, хранения данных и информационных ресурсов.

Развитие ИТ все острее ставит на повестку дня вопрос создания новых моделей управления информационной инфраструктурой. Многие специалисты сходятся во мнении, что использование сервисориентированной модели, основанной на автоматическом контроле сложных сетей и систем, обнаружении новых устройств и предсказаний конфликтов – способно существенно повысить эффективность управления бизнес-процессами¹.

Следует также обратить внимание на то, что в основе ИИ лежат информационные технологии, которые по мере развития компании должны постоянно совершенствоваться, динамично и адекватно реагировать на изменения внешней среды. Это позволит информационной инфраструктуре стать фактором конкурентного преимущества, позволяющим добиться наилучших результатов.

2 Адаптивная информационная инфраструктура

Успех динамично развивающейся компании во многом зависит от того, в какой степени ее информационная инфраструктура обладает свойством адаптивности.

Адаптивная инфраструктура – это методология реализации более гибкой и «разумной» инфраструктуры, настраиваемой на потребности конкретного предприятия и позволяющей организациям любого размера оперативно реагировать на изменения рынка и информационных потоков.

ИИ компании должна способствовать развитию корпоративной информационной системы, интеграции новых приложений и ИТ-решений без коренных перестроек. Изменения информационной системы не должны приводить к необходимости заново создавать инфраструктуру компании или коренным образом изменять ее, должна существовать возможность максимально эффективного использования существующих информационно-технологических ресурсов (ИТР). В работе [3] ИТР определяются как совокупность элементов информационно-коммуникационных технологий, приобретенных, адаптированных и используемых в операционной деятельности организаций, учреждений и домашних хозяйств.

Понятие адаптивной информационной инфраструктуры тесно связано с концепцией *Adaptive Enterprise*, предполагающей возможность построения гибкой ИТ-инфраструктуры компании, способной адаптироваться к постоянно меняющимся бизнес-задачам и основанной на технологиях виртуализации, Grid и SOA (*Service-Oriented Architecture*).

Adaptive Enterprise позволяет синхронизировать бизнес и ИТ, эффективно, быстро и легко использовать информационную инфраструктуру предприятия для поддержки бизнес-решений².

Концепция *Adaptive Enterprise* служит для обеспечения:

- автоматического интеллектуального управления всеми распределенными инфраструктурными компонентами информационной системы – серверами и системами хранения данных;
- высокой надежности и готовности инфраструктуры;
- динамической оптимизации и перераспределения вычислительных ресурсов системы.

Adaptive Enterprise базируется на четырех основных принципах: простота, стандартизация, модульность и интеграция.

¹ Рожной А., Управление ИТ-инфраструктурой: горячая тема года [Электронный ресурс] Режим доступа: <http://www.cnews.ru/reviews/?2005/12/13/193169>, свободный. – Заглавие с экрана.

² *Adaptive Enterprise* [электронный ресурс] // Бизнес Компьютер Центр. - Режим доступа: <http://www.bcc.ru/ae/>, свободный. – Заглавие с экрана.

Простота инфраструктуры подразумевает уменьшение ИТ-элементов в информационной инфраструктуре и позволяет легко модифицировать ее для любых целей. *Стандартизация* обеспечивает взаимодействие с любой инфраструктурой за счет использования стандартных технологий, процессов, интерфейсов и многократно используемых элементов. *Модульность* обеспечивает адекватное отражение изменения одного элемента на всей сети. И, наконец, *интеграция* подразумевает создание динамических связей между бизнесом и ИТ, организацию связей между приложениями и процессами, организацию работы всего со всем.³

Одним из лидеров в разработке этого направления является корпорация HP.

Adaptive Enterprise в том виде, как его представляет Hewlett-Packard, образуется из трех составляющих:

Архитектурная платформа для корпоративных и межкорпоративных приложений, поддерживающая бизнес-процессы. В основе подобной платформы лежит, прежде всего, ориентация на стандартные интерфейсы. Ключевое его преимущество – простота реализации систем из стандартных компонентов.

Модель UDC. В основе Utility Data Center (усиленная версия HP-UX Workload Manager) лежит ПО Virtual Server Environment, которое позволяет управлять нагрузкой Unix-систем.

Программная платформа. В адаптивной платформе используется проверенное временем ПО HP OpenView, являющееся основой для компонентного архитектурного подхода динамического управления ресурсами⁴.

Таким образом, концепция Adaptive Enterprise - это и стратегия, и архитектура, и изменение взгляда на использование компьютерных систем на предприятиях. Все это подразумевает создание инфраструктуры, в которой аппаратные и программные средства и услуги можно интегрировать для достижения наилучшего результата за счет обеспечения большей адаптивности ИТ-среды. Создается особая инфраструктура, состоящая из управленческого программного и аппаратного обеспечения, поддерживающего функции интеллектуального управления и услуги для клиентов. Эта концепция подкрепляется стратегическим партнерством мировых лидеров отрасли ИТ, таких как HP, Microsoft, BEA, Siebel, Oracle, Peoplesoft, SAP, CiscoSystems. Стратегия Adaptive Enterprise позволяет динамично реагировать на изменения рынка, более полно удовлетворять потребности клиентов, сокращать расходы компании и оптимизировать ИТ-инфраструктуры с расчетом на будущее.

Анализируя предложения вендоров ИТ-рынка, можно сделать вывод о том, что адаптивность, по сути, является сегодня определяющим механизмом, глубоко затрагивающим все бизнес-процессы компании. Адаптивной должна быть не только информационная инфраструктура компании, но и менеджмент, управленческие структуры.

3 Интеграция технологии CRM и мультиагентного подхода при формировании информационной инфраструктуры компании с развитой клиентской сетью

Сети, формирование которых в начале 1980-х годов началось во многом стихийно, в настоящее время все больше превращаются в инструмент целенаправленной координации усилий всех участников. В соответствии с этим возникает настоятельная необходимость как изучения теоретических принципов формирования сетей, так и выделения совокупности конкретных управленческих навыков, которые требуются для успешного их функционирования, путем разработки стратегии управления отношениями фирмы с другими участниками сети [6].

К основным принципам управления клиентской сетью можно отнести:

- 1) Системный подход к работе с клиентами;

³ Концепция Adaptive Enterprise [электронный ресурс] // Бизнес Компьютер Центр. - Режим доступа: <http://www.bcc.ru/ae/aeldea>, свободный. – Заглавие с экрана.

⁴ Черняк Л. Естественные технологии // Computerworld.-2003.- 4 авг.(№28-29).- Режим доступа: <http://www.osp.ru/cw/2003/28-29/66621/>, свободный. – Заглавие с экрана.

- 2) Постоянную деятельность по привлечению новых клиентов и работе с потенциальными клиентами;
- 3) Проектный подход к построению и развитию отношений с клиентами;
- 4) Технологический аспект при построении отношений с клиентами, под которым в первую очередь понимается необходимость сегментации клиентов, их потребностей, возможностей и ожиданий. Определение и пересмотр статуса клиента и планирование индивидуальной работы с ним. Своевременная передача информации, как основа для построения отношений с клиентом и успешной работы компании.

Как уже упоминалось выше, характерным для компаний, имеющих развитую клиентскую сеть, ориентированных на оказание услуг и продажи, является наличие развитой внешней информационной инфраструктуры. Существует общепринятое мнение, что технологии CRM наиболее приемлемы для таких предприятий. Однако можно предположить, что такие технологии применимы и для промышленных предприятий, которые также заинтересованы в реализации произведенной продукции, хотя, естественно, функции планирования ресурсов и повышения качества имеют достаточно высокий приоритет.

Говоря о реализации клиентоориентированной стратегии предприятия, необходимо акцентировать внимание на том, что CRM это не только набор технологических инструментов, программные средства выступают здесь лишь как одна из составляющих данного подхода. В первую очередь – реализация концепции CRM – это определенный системный подход к организации деятельности компании [7].

В целом стратегия внедрения CRM должна опираться на следующие взаимосвязанные положения:

- 1) Клиентоориентированная политика компании должна служить основой внедрения и последующей адаптации системы;
- 2) Уровень технологической поддержки определяется бизнес-целями;
- 3) Структура ИТ-ресурсов и этапы внедрения задаются детализированными бизнес-задачами, решение которых распределено во времени.

Развернутая CRM-система может включать в себя самые различные модули или использовать информацию из других приложений и баз данных, имеющихся в компании. Однако максимизация ценности каждого клиента достигается в основном за счет трех элементов маркетинговой стратегии: организации системы интегрированных каналов маркетинговых коммуникаций, разработки программ стимулирования спроса и создания продуктов в соответствии с потребностями клиентов. Соответственно, наиболее важные технологические компоненты CRM-системы — это подсистемы интерактивного взаимодействия с клиентами, контроля текущих операций, а также базы данных по продуктам и аналитические модули.

Одной из основных тенденций развития концепции CRM является то, что в настоящее время большинство CRM-продуктов рассматриваются как дополнительный элемент в информационной инфраструктуре предприятия, реализующий ту или иную функцию и интегрированный в нее. Стоит подчеркнуть еще один очень важный момент: нельзя рассматривать CRM только в качестве проекта по внедрению ИТ-технологии. Недооценивание роли правильной организации бизнес-процессов и изменения культурной среды организации неизбежно приведет к неудачам внедрения. Компания должна определить – что необходимо сделать, чтобы бизнес-процессы соответствовали принципам CRM [7].

Еще одной тенденцией развития CRM с точки зрения функциональных возможностей является смещение акцентов с операционного ПО к аналитическому и далее к коллаборационному. Растет спрос на аналитические возможности систем, например, такие, как маркетинговая аналитика: анализ профиля потребителя, моделирование покупательского поведения, планирование рекламных компаний и т.д.

Общемировые тенденции свидетельствуют, что технология CRM становится все более популярной и основной акцент в процессе развития информационной инфраструктуры ком-

пании смещается в сторону делегирования полномочий по формированию требований самому клиенту. Однако следует отметить, что чаще всего при формировании системы управления взаимоотношениями с клиентами используются стандартные решения на основе модулей CRM в качестве механизмов для связи ERP-системы предприятия с внешней средой. Технологии CRM по своей сути не могут быть полноценным инструментом формирования информационной инфраструктуры компании, реализующей информационные взаимодействия во внешней среде. Поэтому исследование возможности построения внешней информационной инфраструктуры с целью организации клиентских сетей на базе комплексного подхода - использование широко применяемой технологией CRM в совокупности с достаточно редко пока реализуемым, но очень перспективным подходом, основанном на мультиагентных системах (MAS), представляется достаточно интересным и перспективным. Эта область менеджмента ещё мало изучена, но быстро развивается.

Интеграция методологий клиентоориентированного подхода организации деятельности компании и мультиагентного подхода, или, другими словами, реализация стратегии CRM с использованием мультиагентных систем может заключаться в следующем [8]:

- Моделирование и прогнозирование поведения клиентов, как постоянных, так и потенциальных;
- Координация работы дилеров и удаленных подразделений с помощью мультиагентной системы;
- Автоматизация и усовершенствование процесса Customer Support в рамках концепции CRM;
- Накопление знаний и навыков сотрудников отделов маркетинга и продаж в базах знаний соответствующих агентов;
- Разработка интегрированного мультиагентного Интернет-портала для ведения агентом личного контента пользователя;
- Создание поискового агента для мониторинга внешней информации;
- Организация портала дистанционного обучения.

Технологии CRM и MAS могут дополнять друг друга. Обе предлагают определенный подход к построению бизнеса, а не просто способ автоматизации отдельных процессов. Так, мультиагентные системы представляют несколько радикальную концепцию, открывающую эру сетевых организаций с коллективным взаимодействием интеллектуальных роботов, предлагая смену мощных централизованных систем полностью децентрализованными, в которых иерархическая структура уступает место сетевой организации, жесткое бюрократическое управление “сверху-вниз” (на основе команд от начальников к подчиненным) - переговорам, план – гибким договоренностям. Результат применения этой концепции – рост объемов производства компаний, повышение прибыльности и конкурентоспособности продукции, мобильности на рынке. Технология CRM также предлагает гибкий подход к построению бизнеса компании. Методология CRM должна рассматриваться не только как концепция взаимодействия с клиентами компании, а скорее как система, помогающая выстроить долгосрочный клиентоориентированный бизнес.

Заключение

В настоящее время бизнес и информационно-коммуникационные технологии пересекаются все теснее. В России, как и во всем мире, ИТ становятся критичным элементом цепочки создания продукта или услуги и получения прибыли. Доходы компаний растут не только за счет существенного сокращения издержек и оптимизации управления, но и благодаря увеличению количества клиентов за счет клиентоориентированной стратегии компании. Немаловажным является тот факт, что основное внимание при построении ИИ компании начинает уделяться ее адаптивности и гибридным архитектурам. Компания может успешно реализовать

стратегию развития, основываясь на этих современных тенденциях, дополняя CRM интеллектуальными информационными средствами и создавая гибридную адаптивную внешнюю инфраструктуру на базе мультиагентного подхода.

Список литературы

- [1] Гаврилова Т.А., Муромцев Д.И. Интеллектуальные технологии в менеджменте: инструменты и системы. Учебное пособие. СПб.: Изд. «Высшая школа менеджмента» Издат. дом С.-Петерб.гос.ун-та, 2007.
- [2] Городецкий В.И., Карсаев О.В., Конюший О.В., Самойлов В.В. Моделирование процессов управления воздушным движением на основе многоагентных технологий. Научный вестник МГТУГА, серия Навигация и УВД, 2008.
- [3] Казанцев А.К., Серова Е.Г., Серова Л.С., Руденко Е.А. Информационно-технологические ресурсы российской экономики. – СПб.: Издат. дом С.-Петерб. гос. ун-та, 2007.
- [4] Катькало В.С. Эволюция теории стратегического управления.– СПб.: Издат.дом С.-Петерб.гос.ун-та, 2006.
- [5] Клейнер Г.Б. Стратегия предприятия. - М.: Издательство «Дело», 2008.
- [6] Куш С.П. Сетевой подход в маркетинге: российский опыт. Вестник СПбГУ. Сер. Менеджмент вып.1 №8, 2002. - С. 81-107.
- [7] Серова Е.Г. Концепция CRM и ее место в общей стратегии развития компании // Современный менеджмент: проблемы и перспективы: Труды IV научно-практической конференции. СПбГИЭУ. (16 апреля 2009 г., Санкт-Петербург, Россия). – Санкт-Петербург: СПбГИЭУ 2009. - С. 301-306.
- [8] Serova Elena. The role of Multi-agent Approach in Building Information Infrastructure for a Modern Company and Carrying Out Management Tasks. – International book series “Information Science and Computing”, Intelligent Information and Engineering Systems, 2009. Vol.3. - P. 130-136.

Biography

Elena G. Serova - PhD in Economics, Senior Lecturer, Deputy Chair of Information Technologies in Management Department, GSOM Saint-Petersburg State University

Research Interests

- Information – technological resources of Russian Economy,
- Managing and Using Information System. A Strategic Approach. Enterprise Information Systems. Marketing Information Systems. CRM.
- Business Process Management,
- Information Management,
- Computer modeling for Business Solutions,

Total number of publications - 33

Internships Abroad

- 2009 - Managing IT for Increasing Business Value. EFMD, Brussels (Belgium).
- 2008 - Strategic Management Education Training Program. UC Berkeley, Haas School of Business, USA.
- 2007 - Financial and Managerial Accounting. Faculty Development Program, Columbia Business School — GSOM SPbSU, St.Petersburg.
- 2007 - Best Practices in Finland. Technical Research Centre (VTT). Helsinki (Finland).
- 2005, 2008 – SAP AG Information service center.

КОРПОРАТИВНАЯ АВТОМАТИЗИРОВАННАЯ СИСТЕМА ОПЕРАТИВНОГО ОБМЕНА ИНФОРМАЦИЕЙ НА ОСНОВЕ СЕРВИСОРИЕНТИРОВАННОГО ПОДХОДА

И.Н. Мараканов

Институт проблем управления РАН
117997, Москва, Профсоюзная, 65, Россия
inm@ipu.rssi.ru
тел/факс: +7 (095) 334-75-90

***Ключевые слова:** информационная система, обмен информацией, центры обмена данными, сервисориентированный подход, web-службы*

Abstract

Principles of construction and software of corporate automated system of information exchange are developed. System work is based on use of the data exchange centers. The main elements of data exchange centers are specialized Web-services, which provide storage, reliable and rapid delivery of information. The description of the main functions and basic components of data exchange centers is given. The problem of optimization of the size of the transferred data is considered. As criterion of optimality the estimation of average time of delivery is used. An adaptive algorithm for choosing the size of the fragment transmitted messages is described. Description and a method for solving the problem of multi-user sending data are considered. The management subsystem provides access of users and blocking of the module of delivery of messages. Lock management function performs a special program. This program verifies the status of this module and provides lock and unlock function. Results of practical use and space-time characteristics of the developed system are represented.

Введение

Одной из важных функций современных территориально распределенных корпоративных информационных систем является обмен данными между пользователями и между приложениями. При этом от таких систем часто требуется обеспечить гарантированную доставку информации (данных, документов) в гетерогенной сетевой среде в сжатые сроки. В обычных средствах электронной почты (например, Outlook Express) такие возможности отсутствуют, и требуется разработка и использование специального программного обеспечения.

Одним из примеров систем, обеспечивающих интеграцию приложений, является использование Брокеров Сообщений и средств Messaging Middleware. В системах интеграции корпоративных информационных систем одним из подходов для такой интеграции стал подход, основанный на использовании программного обеспечения (ПО) гарантированной пересылки сообщений между приложениями. Примером такого ПО является Microsoft BizTalk Server. Основная идея этого ПО следующая. Имеется несколько приложений, связанных некоторой коммуникационной средой, возможно не очень надежной. Одно приложение (например, система документооборота А) должно переслать информацию/документ другому приложению (системе документооборота В). Система А передает документ серверу пересылки сообщений и "забывает" о нем. Сервер пересылки сообщений обеспечивает гарантированную доставку информации до системы В. Брокер сообщений при этом выполняет не только маршрутизацию сообщений между информационными системами, но и выполняет также задачу трансформации этих сообщений для обеспечения совместимости информационных систем. Брокер сообщений интегрирует гетерогенные приложения и данные и предоставляет три типа служб: пе-

рессылка сообщений и перемещение данных обеспечивает физический транспорт доставки сообщений между приложениями; интеллектуальная маршрутизация, которая определяет для каждого сообщения то, к какому приложению оно должно попасть; трансформирование обеспечивает определение соответствия данных между различными семантиками одного приложения или разных приложений.

Цель создания рассматриваемой в настоящей статье системы заключается в обеспечении надежного и оперативного процесса обмена данными между корпоративными пользователями с использованием центров обмена данными (ЦОД) (рисунок 1).



Рисунок 1 – Схема функционирования ЦОД

Среди функций ЦОД можно выделить следующие: обеспечение обмена информацией в рамках локальной сети организации и между пользователями разных организаций; ведение базы данных передаваемых сообщений; контроль входящих и исходящих документов; поиск информации в базе данных; синхронизация справочников и версии программного обеспечения. Основными компонентами ЦОД являются специализированные Web-службы, обеспечивающие надежное хранение и оперативную доставку информации.

1 Структура автоматизированной системы обмена информацией

При обмене информацией возможны три основных типа взаимодействия пользователей: 'каждый с каждым', 'кольцо' и 'звезда'. Обмен информацией по схеме 'каждый с каждым' приводит к большому числу связей, а именно $N(N - 1)$, и усложняет систему. Использование схемы типа 'кольцо' вносит дополнительные задержки при доставке информации. Наиболее подходящей схемой взаимодействия для рассматриваемой системы является звездообразная схема. Данная схема требует использования дополнительного центрального сервера системы, но, с другой стороны, она обеспечивает эффективный процесс обмена информацией. Следует также отметить, что многие территориально распределенные корпоративные системы для решения своих задач используют аналогичную топологию аппаратных средств, поэтому на практике дополнительные затраты на техническое обеспечение рассматриваемой автоматизированной системы могут оказаться минимальными.

На рисунке. 2 показана структура разработанной системы. Сформированные пользователями данные (файлы) хранятся в специальном каталоге ЦОД организации.

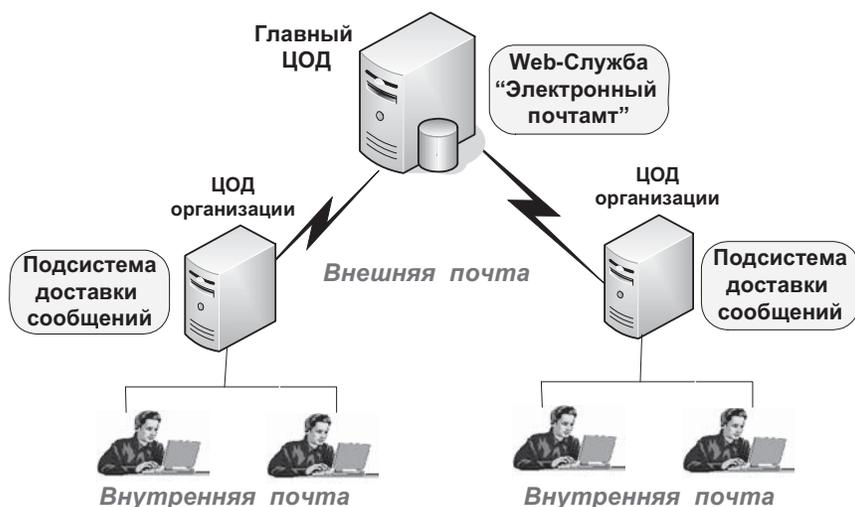


Рисунок 2 – Структура автоматизированной системы обмена корпоративной информацией

Подсистема доставки сообщений обеспечивает передачу этих файлов в главный ЦОД системы и прием от главного ЦОД файлов, предназначенных для данной организации. В главном ЦОД имеется Web-служба, которая выполняет функции 'Электронного почтама', т.е. принимает от ЦОД организаций данные и размещает полученные файлы в соответствующих дисковых каталогах, а также осуществляет отбор файлов, предназначенных данной организации и передачу их в ЦОД организации. Использование такой сервисориентированной архитектуры позволяет обеспечить оперативный обмен данными между указанной Web-службой главного ЦОД и обратившимся к ней клиентским приложением, и реализовать режим гарантированной доставки информации в режиме 'on-line'. Разработанное программное обеспечение системы позволяет передавать файлы заданными фрагментами и, в случае сбоя, осуществлять повторный обмен информацией, начиная с того фрагмента файла, на котором произошел сбой. Таким образом, при обмене информацией возникает задача выбора оптимального размера фрагмента файла, обеспечивающего минимизацию заданного критерия, например, среднего времени доставки информации. В общем случае аналитическая постановка данной задачи достаточно сложна, поэтому рассмотрим ее упрощенный вариант, который, тем не менее, позволяет получить необходимые для практического использования оценки.

2 Оптимизация размеров передаваемых фрагментов данных

Если предположить, что сбой во время доставки фрагмента файла может происходить не более одного раза, средний размер переданной до сбоя части фрагмента составляет половину фрагмента и скорость передачи данных постоянная, то среднее время доставки файла размера R можно записать в следующем виде:

$$T(r) = \frac{R}{r} \left\{ P(r) \left(\frac{r}{s} + \frac{1}{2} \cdot \frac{r}{s} + t \right) + (1 - P(r)) \cdot \left(\frac{r}{s} + \tau \right) \right\}.$$

Здесь введены следующие обозначения:

$T(r)$ – среднее время передачи данных (сек.),

r – размер фрагмента передаваемых данных (байт),

R – общий размер передаваемых данных (байт),

$P(r)$ – вероятность сбоя при передаче данных размера 'r',

s – скорость передачи данных (байт/сек.),

t – время переналадки после сбоя (сек.),

τ – время переналадки после нормального завершения передачи данных (сек.).

Задача заключается в отыскании такого значения r^* , для которого среднее время $T(r)$ минимально, т.е. выполняется равенство

$$T(r^*) = \min T(r), r \leq R,$$

где r^* – оптимальный размер фрагмента файла.

Для качественного анализа поведения $T(r)$ и для оценки значения r^* в качестве $P(r)$ можно использовать известные виды функций, обладающие свойствами монотонного неубывания и вогнутости. Если, например, в качестве зависимостей $P(r)$ использовать зависимости, показанные в левой части рисунка 3, то соответствующие им зависимости $T(r)$ примут вид, показанный в правой части рисунка 3. На практике зависимость $P(r)$ может быть задана в некоторых фиксированных точках и аппроксимирована кусочно-линейной функцией. Соответственно в этих же точках могут быть вычислены значения функции $T(r)$ и получена оценка значения r^* .

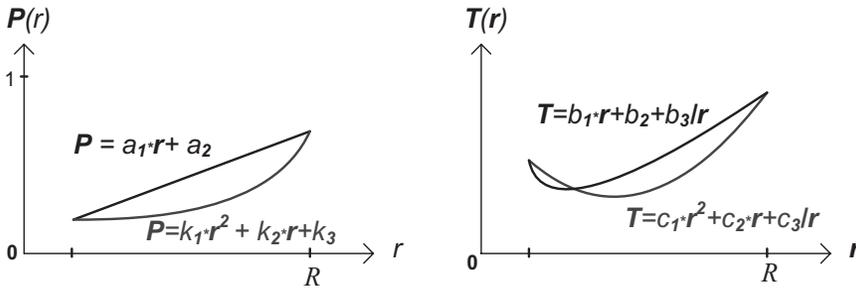


Рисунок 3 – Примеры зависимостей $P(r)$ и $T(r)$

В разработанной системе реализован адаптивный алгоритм формирования функции $P(r)$ и выбора размера оптимального фрагмента передаваемых файлов. На начальной стадии функционирования системы используются значения $P(r)$, полученные экспертным путем. По мере функционирования системы зависимость $P(r)$ корректируется в зависимости от полученных результатов эксплуатации системы. Зависимость оптимального размера фрагмента r_{opt} от размера файла R показана на рисунке 4. Через R_{max} на рисунке обозначен максимальный размер передаваемого файла.

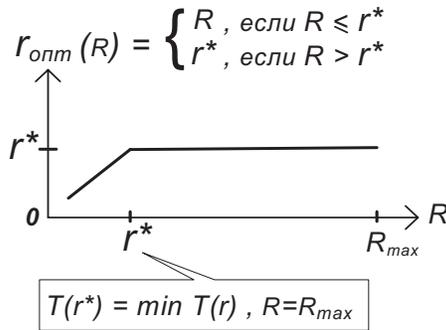


Рисунок 4 – Вид зависимости оптимального размера фрагмента от размера файла

Таким образом, вычислив по имеющимся данным значение r^* , формируется функция r_{opt} и для заданного значения R определяется оптимальный размер фрагмента файла.

3 Управление многопользовательской отправкой данных

При подготовке и отправке файлов группой пользователей возникает необходимость управления доступом к общему дисковому каталогу, в котором хранятся подготовленные пользователями файлы, и их отправкой получателям. Общая схема доставки информации показана на рисунке 5.



Рисунок 5 – Схема доставки информации

Задача управления многопользовательской отправкой данных заключается в следующем. В процессе работы пользователей, сформированные ими файлы с передаваемыми данными сохраняются в общем каталоге на сервере организации. Каждый пользователь, если он обладает необходимыми правами, может вызвать на исполнение модуль доставки сообщений, который осуществляет обмен данными между сервером (ЦОД) организации и главным сервером (главным ЦОД). Система должна управлять процессом обмена информацией, разрешая активизировать указанный модуль доставки только одному из пользователей, запрещая в это время работу с ним другим пользователям. Таким образом, возникает задача управления блокировками общего ресурса, которым в данном случае является общий дисковый каталог доставки сообщений. В разработанной системе функцию управления блокировками выполняет специальная программа-диспетчер, к которой обращаются модули доставки сообщений. Программа-диспетчер обеспечивает чтение, анализ и формирование специальных метаданных, проверку состояния (занят/свободен) общего ресурса, получение информации о том пользователе, который в данный момент времени выполняет обмен данными, если ресурс занят, блокировку и разблокирование общего ресурса.

Блок-схема разработанной двухуровневой системы управления многопользовательской доставкой сообщений на примере выполнения запроса приложения на доступ к общему ресурсу показана на рисунке 6. На первом уровне обеспечивается доступ к общему ресурсу в каждый момент времени только одного из обратившихся пользователей. На втором уровне осуществляется чтение и анализ метаданных, и формирование разрешения или отказа на использование модуля доставки.



Рисунок 6 – Блок-схема выполнения запроса на использование общего ресурса

4 Результаты практического использования системы

На рисунке 7 показана главная форма разработанной системы доставки сообщений. В нижней части формы располагается список готовых к отправке сообщений, а верхней части формы – список доставленных сообщений. В двух правых столбцах фиксируется входящий номер получателя сообщений, а также дата и время доставки. По мере доставки сообщений они переходят из нижнего списка в верхний. В случае сбоев или ошибок, возникающих в процессе доставки сообщений, они фиксируются в специальном файле и могут быть просмотрены. Оператор имеет возможность контролировать как сам процесс доставки, так и его результаты.

№	Получатель	Номер док-та	Дата	Файлов	Размер (кб)	Содержание сообщения	Входящий ГВЦ	Дата доставки
1	ГВЦ	154	27.05.2009	6	342	описание-1		
2	ГВЦ	160	02.09.2009	9	15745	КРАТКОЕ СОДЕРЖАНИЕ 123456789	239 Дубл...	03.09.2009 14...
3	ГВЦ	154	27.05.2009	6	342	описание-1	237 Дубл...	03.09.2009 14...
4	ГВЦ	160	02.09.2009	9	15745	КРАТКОЕ СОДЕРЖАНИЕ 123456789	239	03.09.2009 14...
5	ГВЦ	900159	02.06.2009	2	1	КРАТКОЕ СОДЕРЖАНИЕ 123456789	238	03.09.2009 13...

№	Получатель	Номер док-та	Дата	Файлов	Размер (кб)	Содержание сообщения	Входящий ГВЦ	Дата доставки
1	ГВЦ	154	27.05.2009	6	342	описание-1	237 Дубл...	03.09.2009 15...
2	ГВЦ	160	02.09.2009	9	15745	КРАТКОЕ СОДЕРЖАНИЕ 123456789		
3	ГВЦ	9	02.06.2009	5	109	КРАТКОЕ СОДЕРЖАНИЕ 123456789		

Рисунок 7 – Главная форма системы доставки сообщений

Одной из важных задач, которую удастся эффективно решить с помощью рассматриваемой системы, является регистрация исходящих из филиалов сообщений в главном офисе системы. Пользователи в филиалах в режиме реального времени получают информацию о факте регистрации в главном офисе и входящем регистрационном номере сообщения (электронного документа), направленного в адрес главного офиса. При этом система позволяет вести оперативный контроль и учет зарегистрированных сообщений и практически исключить ошибки, связанные, например, с повторной регистрацией или потерей передаваемых сообщений. Аналогично решается и обратная задача регистрации исходящих из главного офиса сообщений, направленных в адрес филиалов.

Практическое использование данной системы подтвердило ее надежность и оперативность работы. Так, например, передача файлов, имеющих размеры в несколько килобайт, выполнялась в пределах одной секунды, а файлы в несколько мегабайт передавались в течение нескольких секунд.

Заключение

В статье рассмотрены принципы построения корпоративной автоматизированной системы обмена информацией. Система базируется на использовании сервисориентированной архитектуры и специализированных Web-служб. Важной характеристикой системы является гарантированность и оперативность доставки информации получателям. Предложенные подходы и методы решения рассмотренных задач использованы при разработке программного обеспечения системы. Разработанная система обеспечивает оперативный и надежный обмен данными между корпоративными пользователями.

ИСПОЛЬЗОВАНИЕ МЕТОДОВ СТАТИСТИЧЕСКОЙ ОПТИМИЗАЦИИ ДЛЯ УЛУЧШЕНИЯ ЭФФЕКТИВНОСТИ ПОКАЗА ИНТЕРНЕТ РЕКЛАМЫ

Д.А. Ёлкин¹, И.А. Минаков², С.И. Вольман¹

¹ООО НПК «Маджента Девелопмент»
443125, Самара, ул. Ново-Садовая 349, Россия
elkin@magenta-technology.ru
тел: ++7 (846) 342-51-74; 342-51-75

²Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
cscmp@iccs.ru
тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

Ключевые слова: интернет реклама, рекламная кампания, статистические методы, показатели эффективности рекламы

Abstract

The article describes main concepts of online advertising management, main approaches for Internet ad optimization and main statistical rules for ad optimization.

Введение

В настоящее время Интернет все более плотно входит в нашу повседневную жизнь, а вместе с ним в нашу жизнь входит и Интернет реклама. Основные вопросы, которые возникают при показе рекламы - где, когда и кому показывать рекламу, чтобы она была наиболее эффективна. В случае, если имеется большое количество рекламных кампаний и ограниченное количество площадок для показа – вопрос эффективности показа рекламы становится особенно актуальным. А значит, становится актуальным вопрос оптимизации Интернет рекламы [1].

Основная задача оптимизации – это достижение максимального количества реакций на рекламу (кликов, покупок товаров, некоторых специфических действий) при условии того, что количество показов рекламы, оговоренное в контракте, будет выполнено целиком и в срок. Среди основных показателей эффективности хода рекламных кампаний наиболее важными являются следующие:

- Прогресс / доставка кампании – это количество показов, кликов или иных реакций на рекламу, указанных для компании, которые рекламная кампания уже набрала.
- CTR (Click through rate) – это отношение кликов к показам данной кампании, которое, как правило, выражается в процентах. По данному показателю чаще всего и оценивают эффективность хода (показа) рекламной кампании.

Существуют различные подходы по оптимизации Интернет – рекламы. Это и использование методов линейного программирования, и эвристические правила, основанные на опыте экспертов по оптимизации рекламы, и статистические методы оптимизации, каждый из которых обладает своими достоинствами и недостатками. Кратко каждый из подходов может быть охарактеризован следующим образом:

- применение эвристических правил. Для каждого рекламного сервера существует набор специальных параметров, исторической информации по которым не существует, т.е. это управляющие параметры кампаний, такие как приоритет или вес кампаний, частота показа конкретной рекламы конкретному пользователю и т.д. Изменение этих параметров

позволяет эффективно управлять ходом кампании. Кроме того, к эвристическим методам могут быть отнесены методы контекстной оптимизации и методы, основанные на анализе поведения пользователя.

- анализ статистической информации. Этот метод по праву можно считать одним из основных, и именно на нем будет сделан главный акцент в данной статье. Такой тип анализа позволяет понять, на каких рекламных площадках, в какие часы, дни, в каких странах и т.д. реклама шла наиболее эффективно, а где эффективность была значительно хуже. А это означает, что мы можем удалить менее эффективные элементы.
- контекстная оптимизация, суть которой заключается в том, что анализируется содержимое (контент) рекламных площадок, где может быть размещена реклама и сама рекламная кампания либо некоторое ее текстовое описание. Другими словами, на конкретных рекламных площадках отображается та реклама, которая наиболее подходит по тематике данным рекламным площадкам (сайтам, страницам).
- анализ поведения пользователей. Суть данного подхода заключается в том, что когда пользователь посещает различные Интернет ресурсы, информация относительно мест посещения сохраняется (информация, как правило, хранится в cookie браузера пользователя). Таким образом, зная список страниц, которые посетил пользователь, мы можем догадываться о его предпочтениях, а зная его предпочтения, мы можем отображать рекламу более персонализировано для конкретного пользователя.

Более полный обзор существующих методов оптимизации Интернет рекламы был сделан ранее в соответствующей статье [2]. Статистические методы оптимизации являются наиболее действенными, а значит и актуальными для их более подробного описания и изучения, поскольку опираются на статистическую информацию и напрямую учитывают предпочтения пользователей по той или иной рекламе, накопленные с течением времени. Поэтому данная статья подробно описывает общие принципы статистической оптимизации и конкретные варианты стратегий оптимизации, основанных на анализе статистических данных.

Классификация методов статистической оптимизации Интернет рекламы

При использовании статистической оптимизации важно учитывать не только совокупность параметров, по которым должна производиться оптимизация, но и временные интервалы, за которые необходимо рассматривать данную информацию. Если анализировать статистическую оптимизацию с точки зрения временных интервалов, то следует выделить два основных типа оптимизации:

- оперативная оптимизация (рассматриваемый временной интервал не превышает двух недель от текущей даты);
- историческая оптимизация (рассматриваемый временной интервал от двух недель до всего срока жизни кампании).

Оперативная оптимизация:

- позволяет учитывать только актуальные данные, которые не успели устареть, и в которых нет шума, вносимого состояниями кампании, которые на текущий момент потеряли свою актуальность.
- позволяет оценить, насколько эффективны последние изменения, сделанные для кампании. В долгосрочном периоде объем информации, полученный за счет последних изменений, может быть не существенным, в то время как в краткосрочном периоде его достаточно, чтобы сделать выводы о положительной или отрицательной динамике хода кампании.

Историческая оптимизация:

- позволяет оценить динамику кампании за более длительный период, позволяет выявить подъемы либо спады кампании по тем или иным параметрам, которые могут быть не видны при краткосрочном анализе.

- позволяет снизить шумы, которые могут вноситься «сезонностью» при краткосрочном анализе.

Таким образом, для качественной оптимизации Интернет кампаний необходимо грамотно сочетать методы как оперативной, так и исторической оптимизации.

Отдельные методы

Как описывалось ранее в [2], каждая рекламная кампания в процессе своей доставки имеет несколько основных целей: она должна быть доставлена в соответствии с заданным контрактом, т.е. должна быть показана заданное количество раз, либо должна получить заданное количество пользовательских реакций, и второе - иметь максимальный отклик у пользователей, т.е. чтобы пользователь как можно чаще реагировал (совершал некоторые действия) на данную рекламную кампанию. Данные цели в самом общем случае могут быть условно приоритизированы таким образом, что высший приоритет имеет цель по доставке кампании, т.е. чтобы кампания была доставлена вовремя и максимально полно, а уже при условии соблюдения этой цели, желательно, чтобы кампания имела максимальный отклик у пользователей. Используя различные статистические методы оптимизации, и комбинируя их между собой, можно удовлетворить каждую из вышеописанных целей.

Среди перечня всех статистических методов, описанных в статье, посвященной обзору методов оптимизации [2], рассмотрим наиболее эффективные методы для улучшения реакции пользователей на данную рекламу, в частности для улучшения CTR кампании. Это методы оперативной и исторической оптимизации. Среди этих методов выделим две ключевые стратегии оптимизации:

- Безопасное удаление элемента.
- Грубое (рискованное) удаление элемента.

Стратегия безопасного удаления элементов

Рассмотрим сначала алгоритм, используемый согласно безопасной стратегии удаления. Данная стратегия справедлива для любых элементов таргетинга рекламной кампании, как например, страницы, сайты, элементы гео-информации, браузеры, операционные системы и т.д. Общий принцип этой стратегии гласит - удали из кампании наименее эффективные элементы, удаление которых будет не навредит кампании, т.е. ее доставка не пострадает, и в то же время приведет к максимальному приросту CTR. Ключевыми факторами этой стратегии являются безопасность и эффективность удаления. Вся информация, на основе которой делаются выводы об эффективности и безопасности удаления, является статистической информацией, накопленной в течение определенного количества времени.

Безопасность удаления подразумевает, что, удалив от кампании некоторый неэффективный элемент (например, с минимальным CTR), мы сводим к минимуму то обстоятельство, что в результате этого удаления CTR вместо того, чтобы увеличиться, потенциально может уменьшиться. *Эффективность удаления* подразумевает, что мы удаляем не некоторые случайные элементы, а элементы, при удалении которых показатели эффективности рекламы, а именно, CTR возрастет.

В данной стратегии *безопасность* достигается за счет следующих действий:

- 1) Мы рассчитывает среднее значение CTR для всех однотипных элементов, на которых идет данная кампания. Делаем это, используя формулу:

$$(1) CTR_{average} = \frac{\sum_{i=1}^n clicks_i}{\sum_{i=1}^n imps_i}, \text{ где}$$

n – это количество однотипных элементов, на которых идет кампания,
 $clicks_i$ – это количество кликов, которое набрала данная кампания на данном элементе за рассматриваемый период, по сути это мат. ожидание кампании на данных элементах,
 $imps_i$ – аналогичное число показов на данном элементе.

После того как мы нашли величину среднего CTR для данной кампании за определенный период, мы отсекаем все элементы, на которых данная кампания имеет величину CTR за аналогичный период, выше или равную данной. Так как это небезопасно и заведомо неэффективно удалять не самые плохие элементы.

- 2) Следующий шаг, это понижения уровня шума, который вносится наличием у кампании недостаточной статистики. Для этого делаем предположение, что кампания за тот же период набрала на каждом элементе на m кликов больше, чем в действительности (как правило, $m=2$ является достаточным), и снова вычисляем CTR с учетом нового числа кликов на каждом элементе. Если статистики было немного, а именно число показов мало, то добавленные нами клики, существенно поднимут CTR элемента, а значит, его новый CTR будет выше ранее рассмотренного среднего CTR кампании на данном элементе, а значит удалять этот элемент не безопасно. Если величина CTR данного элемента существенно не меняется, значит, на нем было достаточно показов, чтобы судить о достоверности CTR, которое имеет кампания на данном элементе.
- 3) Далее нам необходимо проанализировать, насколько величина CTR, вычисленная в пункте 2, отличается от величины среднего CTR. Экспертным путем было установлено, что наиболее безопасно удалять элементы, величина CTR на которых на 10-15% меньше среднего CTR, вычисленного в пункте 1.
- 4) Суммарное количество показов, которое удаляет данная стратегия не должно превосходить 15-20% от общего числа показов, которое набрала кампания на данных элементах за рассматриваемый период.

Т.о. если элементы удовлетворяют всем пунктам, описанным выше, то считается, что их удаление будет безопасно, т.е. не понизит CTR кампании.

Эффективность данной стратегии достигается за счет выполнения следующих факторов:

- 1) Не имеет смысла удалять те элементы, на которых недостаточно статистики. Считается, что статистики достаточно, если число показов на данном элементе не меньше величины $imps_{min}$, которая вычисляется по формуле:

$$(2) \quad imps_{min} = K \cdot \frac{1}{CTR_{average}}, \text{ где}$$

$CTR_{average}$ – это среднее CTR кампании,

а само отношение $\frac{1}{CTR_{average}}$ – это число показов, за которое в среднем достигается один клик для данной кампании,

K – это корректирующий коэффициент для каждой конкретной рекламной сети, т.е. в самом простом случае он равен 1, если мы хотим подстраховаться, то его можно увеличить до 2 или 3.

- 2) После того как были отсечены элементы с недостаточной статистикой, среди всех оставшихся элементов и комбинаций из этих оставшихся элементов, мы выбираем те, значение метрики эффективности которых будет лучше. В данном случае метрика эффективности $metric$ вычисляется по формуле:

$$(3) \quad metric = \frac{ctrRate_i}{impsRate_i}, \text{ где}$$

$ctrRate_i = \frac{PotenCTR_i}{CTR_{average}} - 1$ – величина $ctrRate_i$ описывает потенциальный прирост CTR

после удаления i -го элемента,

$$PotenCTR_i = \frac{\left(\sum_{j=1}^n click_j \right) - click_i}{\left(\sum_{j=1}^n imps_j \right) - imps_i} - \text{CTR после удаления } i\text{-го элемента.}$$

$impsRate_i = \frac{imps_i}{\sum_{j=1}^n imps_j}$ – величина, характеризующая количество удаляемых показов.

Таким образом, величина метрики будет больше, если прирост CTR после удаление элементов будет выше, а потеря показов будет как можно ниже. По данной метрике выбирается несколько наиболее эффективных элементов, которые рекомендует удалить данная стратегия.

Стратегия грубого удаления элементов

Данная стратегия является менее безопасной, по сравнению со стратегией, рассмотренной выше. Основным критерий, который учитывается в ней – это эффективность удаления, причем эффективность больше основывается не на математических расчетах, а на экспертной оценке пользователей, использующих данную стратегию оптимизации. Безопасность здесь учитывается только в том плане, что стратегия не позволяет удалить элементы, если на них суммарно более чем 30% от общего числа показов, которое набирает кампания на данных элементах.

Эффективность достигается за счет того, что удаляются элементы, на которых данная кампания набрала число показов больше, чем $minImps$ (для большинства рекламных сетей величина $minImps$, равная 1500, является достаточной), и при этом на данном элементе нет ни одного клика. Таким образом, данная стратегия удаляет заведомо плохие элементы, но при этом никак не учитывает погрешность, вносимую шумом, из-за отсутствия достаточной статистики. Как уже было написано выше, данная стратегия является более рискованной, чем предыдущая стратегия, поэтому на пользователей, которые ее применяют, возлагается дополнительная ответственность.

Результаты применения статистических методов, использующих стратегии удаления, описанные выше, можно увидеть на рисунке 1. Данная диаграмма показывает, как CTR кампании (показатель реакции пользователя на данную рекламу) начал увеличиваться с момента применения к данной кампании методов статистической оптимизации. Сам график показывает изменение/рост CTR, а вертикальные линии – моменты, когда были применены стратегии по удалению элементов. В результате CTR кампании, увеличился с 0,19% до 0,29%, что в относительных величинах дает прирост около 50%, что является существенным. Средний же прирост CTR при использовании данных стратегий составляет порядка 20-30%.

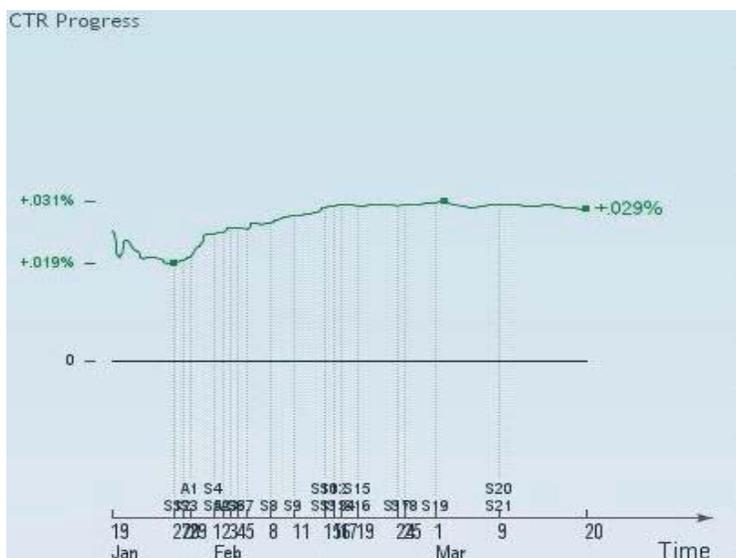


Рисунок 1 – Результаты использования методов статистической оптимизации для улучшения CTR

Стратегия оптимального добавления элементов

Кроме того, что статистические методы оптимизации используются для улучшения CTR рекламных кампаний, они хорошо подходят для улучшения доставки кампании. Существуют различные стратегии добавления элементов, основанные на статистических методах. Стратегию, которая получила наиболее широкое применение на практике, описана ниже.

В основе данной стратегии лежит анализ статистической информации, относительно того, какое количество свободных показов может получить данная рекламная кампания при добавлении к ней некоторого элемента (например, страницы, сайта и т.д.). Свободные показы – это показы, которые на предполагаемых добавляемых элементах набирали кампании более низкого приоритета, по сравнению с анализируемой кампанией. Данная информация анализируется за фиксированные промежутки времени, как правило, это две-три недели, так как за этот промежуток успевают накопиться достаточное количество информации, и при этом она еще не утратила своей актуальности. Данная стратегия применяется к рекламным кампаниям в том случае, когда у кампаний возникают проблемы с доставкой, а именно, при наборе показов они существенно отстают от планового значения, которое вычисляется на текущий момент времени. Плановое значение вычисляется на основе того, сколько кампания должна была бы набрать показов на текущий момент, если бы она набирала все показы равномерно день за днем. Т.е. для вычисления планового значения в самом общем случае можно использовать формулу:

$$(4) \quad \textit{imps}_{plan} = \frac{\textit{imps}_{goal}}{t_{total}} \cdot t_{elapsed}, \text{ где}$$

\textit{imps}_{goal} – общее количество показов, которое должна набрать кампания,

t_{total} – общее количество времени, которое идет кампания от начала до конца,

$t_{elapsed}$ – общее количество времени, которое прошло от начала кампании.

Для того чтобы знать, какой именно элемент надо добавить, необходимо знать какое потенциальное количество показов может принести добавление данного элемента кампании. Под элементом в данном случае понимается страница, секция, сайт, браузер и т.д., т.е. по сути,

любой элемент таргетинга кампании, на котором она на текущий момент не идет. Потенциальное количество показов – это такое количество показов, которое кампания сможет отобрать у других рекламных кампаний более низкого или равного с ней приоритета при добавлении данного или совокупности данных элементов. Для того чтобы определить потенциальное количество показов, которое получит исходная анализируемая кампания при добавлении элементов, необходимо пройти по каждому элементу, и выяснить какое количество кампаний более низкого или равного приоритета идет на данном элементе, и какое количество показов они набирают в день, а соответственно, сколько показов искомая кампания сможет отобрать.

Следующей задачей после определения потенциального количества показов, которое может принести каждый элемент, является определение того, сколько элементов, а главное, каких добавлять. Здесь нужно помнить, что важно добавить не те элементы, на которых больше всего потенциальных показов, а те, на которых показов соизмеримое количество с нуждами кампании, и при этом данный элемент обладает неплохим CTR, т.е. при его добавлении CTR кампании сильно просесть не должен.

Нужды кампании можно определить как разность между количеством показов, которое кампания должна была набрать по плану на текущий момент, и реальным количеством показов, которое на данный момент кампания уже набрала. Как правило, полученные нужды кампании корректируют, умножая на коэффициент, который находится в диапазоне [1..3]. Корректировка необходима от того, что количество потенциальных показов, которое может принести каждый элемент, как правило, точно определить нельзя, так как для этого необходимо оценивать все срезы кампании по всем возможным элементам таргетинга. Кроме того, что эти вычисления могут быть не точными, они могут занимать больших вычислительных ресурсов, так как для анализа требуется большое количество статистической информации.

После того как нужды кампании определены с учетом корректирующего коэффициента, остается среди всех элементов, пригодных для добавления, выбрать те, которые будут наиболее оптимальными. Так как в данной задаче главное – это улучшить доставку кампании, то первый приоритет при добавлении элементов имеет потенциальное количество показов, которое принесет элемент, а второе – это CTR этого элемента. Но, несмотря на эти приоритеты, при выборе элемента важно соблюдать некоторую грань разумности. А именно, если один элемент имеет количество потенциальных показов на 100 больше, чем другой, и при этом второй элемент обладает CTR вдвое лучшим, чем первый элемент, то надо понимать, для исходной кампании эти 100 лишних показов критичны, для того чтобы быть успешно доставленной или нет. Если критичны, то целесообразнее пожертвовать CTR и добавить первый элемент с большим числом показов, если не так критичным (например, величина нужд исходной кампании на порядок, а то и на два выше этих 100 лишних показов), то лучше добавить второй элемент с лучшим CTR.

На рисунке 2 показан эффект от применения статистических методов оптимизации рекламных кампаний. На этой диаграмме видно, что до применения советов кампания имела проблемы с доставкой, после добавления элементов (страниц, секций) ситуация с доставкой выровнялась, и в итоге кампания была доставлена вовремя и целиком.

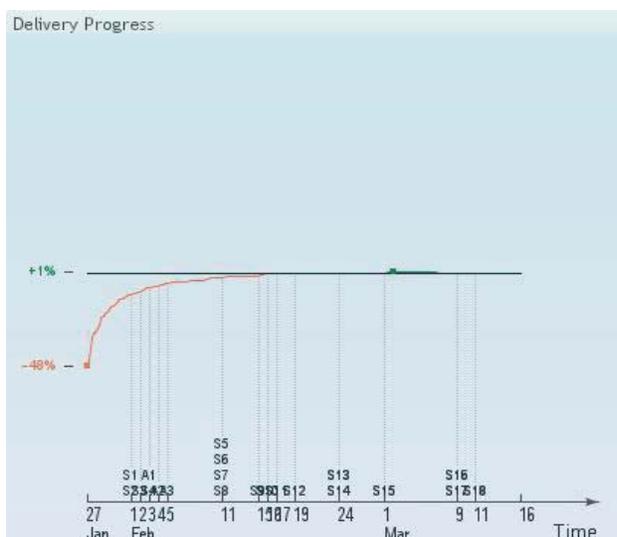


Рисунок 2 – Результаты использования методов статистической оптимизации для улучшения доставки кампании

Для сравнения описанных методов был произведен анализ эффективности применения данных стратегий для улучшения доставки и CTR кампаний. Результаты анализа представлены в таблице 1.

Таблица 1

Название стратегии / метода	Средний прирост CTR	Изменение прогресса / доставки кампании
Безопасное удаление элемента (статистический метод)	15-20%	Прогресс существенно не меняется
Грубое удаление элемента (статистический метод)	20-25% (достигается за счет того, что удаляются с большим числом показов, в результате чего большее количество показов может перейти на более эффективные элементы)	Прогресс существенно не меняется, либо может быть снижен на 3-5%
Оптимальное добавление элемента (статистический метод)	CTR существенно не меняется	Улучшение доставки на 5-15%
Эвристические методы	5-15%	Улучшение доставки на 10-15%
Методы линейной оптимизации (использование транспортной задачи [2])	1-5% (трудно достичь оптимального плана в реальности)	Прогресс существенно не меняется

Данные, представленные в таблице 1, были получены в результате наблюдения более 50 рекламных кампаний. Срок наблюдения за кампаниями составил месяц (время жизни большинства рекламных кампаний равен месяцу). В результате данного наблюдения становится очевидным, что методы статистической оптимизации дают несколько больший

прирост CTR, по сравнению с прочими методами, и при этом не ухудшают доставку кампаний. Кроме того, результаты, которые каждый раз предлагаются статистическими методами, обладают более высокой степенью доверия, так как анализируется реальная статистика предпочтений пользователя.

Заключение

В данной статье описаны основные статистические методы и подходы, которые используются для оптимизации Интернет рекламы. Методы, основанные на анализе статистической информации, позволяют учитывать исторические закономерности и, тем самым, принимать грамотные оптимизационные решения, добавляя эффективные и удаляя малоэффективные элементы. Оперативный статистический анализ позволяет учитывать только актуальные данные, которые не успели устареть, а также позволяет оценить, насколько эффективны последние изменения, сделанные для кампании. В более длинном временном интервале объем информации, полученный за счет последних изменений, может быть не существенным, в то время как в более коротком интервале его достаточно, чтобы сделать выводы о положительной или отрицательной динамике хода кампании.

Исторический статистический анализ позволяет оценивать динамику кампании за длительный период, которая может быть не видна при оперативном анализе, кроме того, позволяет сглаживать шумы, которые могут вноситься «сезонностью» при оперативном анализе. Таким образом, для качественного управления кампаниями необходимо грамотно сочетать методы оперативной и исторической оптимизации.

В статье приведены результаты применения на реальных рекламных кампаниях различных стратегий оптимизации, в том числе стратегий по добавлению и удалению элементов, основанные на статистических методах. По данным результатам можно судить, что использование статистических методов и величин позволяет существенно повысить эффективность хода рекламных кампаний, что для реальных пользователей таких методов и систем, использующих данные методы, напрямую проецируется в денежную прибыль.

Список литературы

- [1] Plummer Joe, Rappaport Steve, Hall Taddy, Barocci Robert The Online advertising playbook. //Wiley 2007, John Wiley & Sons, Inc, 2007, p. 28-32.
- [2] Ёлкин Д.А., Минаков И.А., Вольман С.И., Обзор подходов оптимизации Интернет рекламы // Проблемы управления и моделирования в сложных системах: Тр. XI Междунар. конф. С.62-68.
- [3] Минаков И.А. Онлайн-анализ пользователей Интернет-портала продажи «горящих» авиабилетов // Информационные технологии, 2006. № 1. С.62-68.
- [4] Якушин А.В., Вольман С.И., Минаков И.А. Разработка системы поддержки принятия решений при оптимизации хода рекламных кампаний в сети Интернет // Проблемы управления и моделирования в сложных системах: Тр. XI Междунар. конф. С.68-72.

РАЗРАБОТКА СИСТЕМЫ ПРОГНОЗИРОВАНИЯ ПОВЕДЕНЧЕСКИХ ХАРАКТЕРИСТИК ПОЛЬЗОВАТЕЛЯ В СЕТИ ИНТЕРНЕТ

А.В. Якушин¹, А.В. Кочуров¹, С.И. Вольман¹, И.А. Минаков²

¹ООО НПК «Маджента Девелопмент»

443125, Самара, ул. Ново-Садовая 349, Россия

yakoushin@magenta-technology.ru

тел: +7 (8462) 342-51-74; 342-51-75; 342-51-76; 342-51-77

²Институт проблем управления сложными системами РАН

443020, Самара, ул. Садовая, 61, Россия

cscmp@iccs.ru

тел: +7 (846) 332-39-27, факс: +7 (846) 333-27-70

***Ключевые слова:** рекламная кампания, прогнозирование, частотные ограничения, поведенческие характеристики пользователя*

Abstract

The paper discusses new approach to forecasting of users behavior characteristics related to online advertisement. The relationship between the campaign goals and users behavior is considered. The approach allows forecasting the behaviors of users, depending on the situation of the ad network as a whole. Main principles of the approach, implementation, the result of research on existing networks and perspectives are described.

Введение

Современные маркетинговые кампании немислимы без рекламы в сети Интернет. Огромное количество информации, размещаемое на новостных и развлекательных порталах, на сайтах известных издательств и телевизионных компаний привлекает массы посетителей. И почти всегда сопровождается рекламой важных событий, новостей, товаров, услуг.

Любая рекламная кампания становится результативной только в том случае, если она отвечает поставленной перед ней цели. Основными же целями любого рекламного проекта являются внедрение на рынок новых товаров и услуг, а также вовлечение потенциальных потребителей и заключение с ним сделки [4].

Существующие на данный момент сервисы доставки рекламы, такие как Google DoubleClick, 24/7 Realmedia OAS, AdTech Helios, Google AdSense, Yandex.Direct, Begun и пр. специализируются на выполнении целей рекламных кампаний, стараясь использовать все имеющиеся ресурсы интернет страниц. Эти сервисы предоставляют оператору определённый набор инструментов, позволяющих управлять рекламными кампаниями и принимать решения по оптимизации их хода. Имеющиеся статистические данные дают возможность проводить детальный анализ сложившейся ситуации.

Для достижения целей рекламных кампаний сервисы доставки рекламы позволяют настраивать их на различные поведенческие характеристики посетителей сайтов. В связи с этим возникает необходимость оценки и анализа поведенческих характеристик пользователя, просматривающих рекламу, а так же их прогнозирование при изменении рекламной сети. Однако наличие человеческого фактора, зашумлённости данных, а также высокая степень зависимости хода рекламных кампаний друг от друга не позволяет операторам вручную принимать эффективные решения по управлению и снижению рисков.

Данная статья посвящена вопросам прогнозирования поведенческих характеристик пользователя в рамках рекламы в сети Интернет, а так же является продолжением предыдущей статьи [1].

Анализ предметной области

Одной из важнейших задач оптимизации хода рекламных кампаний является прогнозирование динамики поведения пользователя в сети Интернет. Прогнозирование может быть проведено различными методами, так или иначе базирующихся на моделировании аспектов поведения пользователя и влияющих на него факторов [1]. Это не только поведенческие характеристики пользователя, но и содержимое рекламной сети в виде совокупности связанных интернет страниц, а также набор рекламных кампаний и алгоритмы AdServer'a по доставки рекламных кампаний до пользователя. Зависимость различных составляющих интернет рекламы показана на рисунке 1.



Рисунок 1 – Зависимость составляющих интернет рекламы

Владельцы сайта размещают контент страницы и предоставляют места для рекламы, на которые AdServer доставляет рекламные баннеры. Медийные агентства заказывают рекламные кампании на AdServer'e. Посещение пользователем определённых страниц способствует достижению целей рекламных кампаний. Это приведёт к их автоматической перенастройке, что в конечном итоге отразится на характере показываемой рекламы [2, 3].

Для анализа поведения пользователя используются данные AdServer'a, представленные, как правило, в виде статистики, например, отчёт о количестве пользователей, показов и кликов, набранных кампаниями на страницах. Следовательно, о поведении конкретных пользователей можно судить лишь по усреднённым косвенным показателям: объём показов, доля показов на одного пользователя, вероятность клика на кампанию и т. п. С другой стороны, задачей прогнозирования является также получение аналогичных косвенных показателей.

Введём понятие «пользователь». Пользователем – это посетитель интернет страницы, которого можно идентифицировать по некоторым ключевым параметрам: ip адрес, как основной параметр, и используемый браузер, содержащий cookie данные пользователя. Следствием является одна из важных особенностей предметной области: реальный человек может быть представлен в виде нескольких пользователей и наоборот.

Пользователь, как любая другая составляющая исследуемого объекта, обладает рядом характеристик. Эти характеристики можно разделить прямые и косвенные.

К прямым характеристикам относятся те, которые можно непосредственно получить, используя систему отчётов AdServer'a:

- распределение количества показов рекламы по дням недели;
- распределение количества показов рекламы по часам;
- количество уникальных пользователей, просмотревших рекламу;
- распределение количества совершённых просмотров рекламы одним пользователем.

К косвенным характеристикам относятся те, которые можно получить с помощью некоторых эвристик:

- новизна рекламы для пользователя;
- перенасыщенность рекламой;
- продолжительность сессии пользователя;
- плотность просмотра рекламы;

Рекламная кампания содержит ряд настроек, чётко определяющих целевую аудиторию. Эти настройки: географическое положение, допустимые браузеры, допустимые операционные системы, виды Интернет-соединения – все они основаны на использовании ключевых параметров пользователя. Другие настройки кампании уже ограничивают показ выбранной аудитории. Именно эти параметры определяют то, как поведение пользователя влияет на ход рекламной кампании. К таким настройкам относятся:

- ограничения по дневные и часовые цели кампании по набору показов и кликов;
- ограничение по дням недели и часам;
- приоритет и вес кампании;
- частотное ограничение по показам / кликам для конкретного пользователя.

Под частотными ограничениями понимается ограничение на количество просмотров рекламных баннеров одной кампании (или кликов на них) в заданный интервал времени. Интервалы могут быть: сессия, час, день, неделя, месяц, весь ход кампании.

Взаимосвязь этих настроек рекламной кампании и поведенческих характеристик пользователя будет рассмотрено в этой статье.

Постановка проблемы

Прогнозирование поведенческих характеристик пользователя имеет ряд особенностей. Во-первых, пользователь в процессе просмотра рекламы ведёт себя трудно предсказуемо. Не имеется возможности с достаточной уверенностью утверждать, что конкретный пользователь будет посещать определённый ряд страниц и посмотрит определённое количество рекламы. Во-вторых, техническая реализация моделирования поведения каждого пользователя в принципе возможна, однако требуемые временные рамки для расчёта прогноза не позволяют проводить моделирование такого рода. В-третьих, существующие системы сбора статистики не позволяют получать все необходимые данные по техническим и правовым причинам.

В силу этих особенностей, системы прогнозирования поведенческих характеристик пользователя вынуждены иметь дело не с набором уникальных по поведению пользователей, а с множеством абстрактных пользователей. Таким образом, в процессе моделирования рассматривается не отдельный пользователь, а некая группа пользователей.

Из описанных выше настроек кампании наибольшей сложными в моделировании являются частотные ограничения. Задание частотных ограничений для кампании означает, что конкретный пользователь при заходе на страницу будет видеть баннеры кампании не каждый раз. Следовательно, часть показов рекламы, где кампания могла бы потенциально быть показана, достанутся другим кампаниям. И для каждого пользователя, в силу уникальности его поведения, эффект от использования частотных ограничений будет различным. Таким образом, задача прогнозирования эффекта использования частотных ограничений в силу допущения о группах пользователей, становится трудноразрешимой.

Существующие системы прогнозирования поведенческих характеристик пользователя частично решают эту задачу. В основе существующих решений лежит метод расчёта коррекционного коэффициента, показывающего, в сколько раз изменится возможное количество показов, набираемых кампанией в день, при изменении частотных ограничений рекламной кампании. Обычно такие коэффициенты задаются экспертом. Единственным достоинством такого метода является простота реализации. Недостатком метода является неадекватность модели,

лежащая в его основе, не учитывающая взаимное влияние рекламных кампаний. Помимо этого, не учитывается влияние ограничений на временные рамки, влияющие на используемые коэффициенты.

Выявленные недостатки заставляют оператора отказываться от использования эффективных способов настройки кампании.

Предлагаемый подход

Основной идеей предлагаемого подхода является разработка модели поведения пользователя наиболее близкой к реальному поведению, и дальнейшее использование её для прогнозирования поведенческих характеристик [1].

В модели имеется ряд допущений. В первую очередь предполагается, что пользователь посещает случайные страницы. Поэтому пользователь характеризуется законом распределения посещаемых страниц. Вторым допущением будет являться то, что на определённом наборе страниц существует несколько групп пользователей, имеющих свои определённые характеристики. В предлагаемой модели предлагается рассматривать:

- $B(u)$ – начало первой сессии в течение дня;
- $E(u)$ – окончание последней сессии;
- $C(u)$ – количество сессий в течение дня;
- $D(u)$ – продолжительность сессии;
- $A(u)$ – количество просматриваемой рекламы за час.

Здесь u – это пользователь.

Такая модель позволяет прогнозировать частотные характеристики непосредственно с прогнозированием временных характеристик – плотности прихода пользователей на сайт.

Процесс показа рекламы группе пользователей является случайным процессом. Пусть $f(x)$ – плотность распределения пользователей по числу показов рекламы за день на пользователя, получаемая из вышеописанных параметров модели поведения пользователя и статистических данных отчётов, предоставляемых AdServer'ом. Расчёт производится градиентным методом с использованием эвристических методов поиска начального приближения.

Лимит числа просмотров кампании i для одного пользователя для кампаний равны L_i . Будем полагать, что список кампаний упорядочен по возрастанию ограничений.

В силу того, что количество показов рекламы достаточно велико и может достигать десятков миллионов показов в день, можно перейти к рассмотрению типичных последовательностей. Перейдем также к рассмотрению мат. ожиданий. Это допустимо, поскольку для задачи прогнозирования представляет интерес в первую очередь математическое ожидание числа показов кампании, а не точное распределение, получить которое невозможно вследствие отсутствия необходимой информации. При таком рассмотрении число выпавших на кампанию показов будем считать равным его математическому ожиданию. Другими словами, если g показов разыгрывается между кампаниями, то количество показов равно:

$$(1) \quad I_i(r) = r \frac{w_i}{\sum_k w_k},$$

где i – номер кампании, w_i – вес кампании.

Формула (1) применима для большинства случаев, и используется во всех системах динамики прогнозирования поведения пользователя. Однако, наличие хотя бы одной кампании с частотным ограничением меняет реакцию группы пользователей.

Рассмотрим поведение отдельного пользователя. Пользователь просматривает страницы содержащие рекламу. При этом каждая кампания, показываемая к моменту просмотра пользователем g реклам, характеризуется набранным количеством показов $I_i(r)$, а так же возможностью показываться данному пользователю, определяемой неравенством:

$$(2) \quad I_i(r) \leq L_i.$$

Тогда количество показов кампании к моменту показа дополнительных Δr :

$$(3) \quad I_i(r + \Delta r) = I_i(r) + \Delta r \cdot \frac{w_i}{\sum_{k: R_k \leq L_k} w_k},$$

при условии, что $I_i(r + \Delta r) \leq L_i$.

Заметим, что в любой момент времени по отношению к отдельному пользователю существует два вида кампаний: достигшие предела L_i и продолжающие делить оставшиеся показы пользователя пропорционально весам.

При переходе на математическое ожидание из (1) получаем:

$$(4) \quad MI_i = \int_0^{+\infty} I_i(r) f(r) dr.$$

После разделения на отрезки, на которых набор кампаний, достигших предела не изменяется, из (4) получим следующее:

$$(5) \quad MI_i = \sum_{k=0}^N \int_{r_k}^{r_{k+1}} I_i(r) f(r) dr,$$

где N – количество отрезков, r_k – количество показов в начале отрезка k . Отметим, что в общем случае количество отрезков равно количеству кампаний. Однако в случае, когда более одной кампании в один момент выбирают свой лимит, отрезок вырождается в точку.

С учётом (3) формулу (5) можно переписать в виде:

$$(6) \quad MI_i = \sum_{k=0}^i MI_i^k + L_i \int_{r_i}^{\infty} f(r) dr,$$

где r_i – количество показов рекламы пользователю, при котором останавливается кампания i .

При этом:

$$(7) \quad MI_i^k = \frac{w_i}{\sum_{j=k}^N w_j} \int_{r_k}^{r_{k+1}} \left(r - \sum_{j=1}^{k-1} L_j \right) f(r) dr$$

Для нахождения r_k предлагается следующий итеративный алгоритм:

- 1) Полагаем, что на предыдущих шагах вычислены точки r_k для $k < k_0$, где k_0 - номер текущего шага.
- 2) r_k – минимальное количество показов, когда сменяется набор кампаний. Другими словами – минимальное количество показов, когда одна из кампаний выберет свой лимит L_k .

Таким образом

$$(8) \quad \left(r - \sum_{j=1}^{k-1} L_j \right) \frac{w_i}{\sum_{j=k}^N w_j} = L_i, \quad i \geq k_0$$

- 3) Так как кампании должны быть отсортированы по возрастанию L_i , кампанию, для которой был получен этот минимум, меняем местами с кампанией k .

Полученные в итоге величины MI_i представляют собой среднее число показов на пользователя для кампании i .

Анализ результатов

В качестве проверки для анализа качества работы программы был проведен тестовый эксперимент, сравнивающий существующую и предлагаемую системы. Было взято 100 рекламных кампаний, имеющих различные параметры и находящихся в различных комбинациях с конкурирующими на страницах кампаниями. Эксперимент проводился для трёх различных рекламных сетей. Проверялось относительное отклонение прогноза по показам на следующую неделю. Допустимым отклонением считалось 20%. В таблице 1 отражены результаты тестирования. Показана доля кампаний, имеющих допустимый прогноз.

Таблица 1 – Результаты тестирования допустимых отклонений прогноза при размещении рекламы на страницах различного типа

	Существующая система	Предлагаемая система
Сеть 1	62%	83%
Сеть 2	36%	69%
Сеть 3	48%	74%

Отдельно были рассмотрены кампании, на которых прогноз имел значительные отклонения. Исследования показали, что поведение пользователей может сильно отличаться не только в различных рекламных сетях, но и в рамках одной сети.

Было рассмотрено 20 кампаний высокого приоритета, имеющих частотные ограничения и идущих на страницах различного типа: домашняя страница, новостная лента, остальные страницы. Сравнивались объём набранных показы и объём доступных к набору показов для кампании за неделю. Результаты исследования показаны в таблице 2.

Таблица 2 – Результаты тестирования допустимых отклонений прогноза при размещении рекламы на страницах различного вида

Тип страницы	Доля набранных показов
Домашняя	89%
Обычная	60%
Новостная лента	20%

Исследования показали, что для частотных ограничений наименьший эффект проявляется на домашних, а наибольший – на новостных лентах. Из этого можно сделать вывод, что пользователи не задерживаются на домашней странице, и не получают большое количество показов рекламы. С другой стороны, на новостной ленте пользователи могут долго задерживаться, просматривая новость за новостью, генерируя большие объёмы показов рекламы.

Также в процессе тестирования было высказано предположение об эффекте перемешивания пользователей для кампаний разного приоритета. Может сложиться ситуация, когда из-за жестких частотных ограничений кампания более высокого приоритета не может быть показана, и ранее нее показывается рекламная кампания более низкого приоритета. В этом случае у кампании низкого приоритета создаётся наведённый эффект, так, будто эта кампания также имеет частотные ограничения.

Описанные выше случаи могут быть учтены в более сложной и перспективной модели потока пользователей, которая будет рассмотрена в последующих статьях.

Заключение

Предлагаемый подход к прогнозированию поведенческих характеристик пользователя позволяет добиться ряда преимуществ, включая такой важный фактор как взаимное влияние кампаний, учёт которого позволяет проводить анализ всей сети в целом. Предложенная модель поведенческих характеристик является самодостаточной и может быть использована как отдельный инструмент, так и быть интегрированной в другую модель. Предложенный метод позволяет решать поставленные задачи в приемлемый срок с достаточной точностью.

Направление, связанное с моделирование поведения пользователей в сети Интернет по отношению к рекламе, представляется весьма перспективным, и работа в этом направлении будет продолжаться.

Список литературы

- [1] И.А. Минаков, А.В. Якушин, А.В. Кочуров, А.Р. Хайрутдинов, С.И. Вольман Разработка системы моделирования динамики поведения пользователей для оптимизации рекламных кампаний в сети Интернет // Проблемы управления и моделирования в сложных системах: Тр. XI Междунар. конф., Самара, 22 июня – 24 июня 2009. Самара: СНЦ РАН, 2009. С.644-651.
- [2] <http://www.247realmedia.com>
- [3] <http://heliosiq.adtech.de>
- [4] <http://www.reinfo.ru/>

ЗАДАЧИ ПРОГРАММНОГО ИНЖИНИРИНГА СЛОЖНЫХ СИСТЕМ НА ОСНОВЕ КРИТЕРИЯ ЖИЗНЕСПОСОБНОСТИ ПРОЕКТА

А.Е. Колоденкова

Уфимский государственный авиационный технический университет
450000, Уфа, ул.К. Марка, 12, Россия
anna82_42@mail.ru
тел: +7 (347) 272-74-65, факс: +7 (347) 272-29-18

Ключевые слова: критерий жизнеспособности программных проектов, задачи распределения ресурсов в управлении проектами, статистическая оценка реалистичности проекта

Abstract

The formal approach is offered to estimate program projects realness based on the statistic method. The approach is based on comparison of various project realisation variants.

Введение

Для снижения рисков, возникающих на ранних стадиях разработки тех или иных проектов, необходимо в первую очередь проводить тщательную оценку жизнеспособности проекта. Это связано с тем, что проект, реализуемый для одной сложной системы, может оказаться нереализуемым для другой в силу ряда причин, таких как масштаб и сложность проекта, уровень квалификации исполнителей проекта и т.п.

Оценка жизнеспособности проекта необходима для выявления ключевых факторов успеха и не допущения провала проекта, выявления надежности, окупаемости, результативности проекта и т.п. Смысл оценки жизнеспособности проекта заключается в представлении всей информации о проекте в форме, позволяющей лицу, принимающему решение, сделать заключение о целесообразности или нецелесообразности реализации проекта.

Настоящая работа посвящена обсуждению задач программного инжиниринга сложных систем на основе критерия жизнеспособности программного проекта и решению задачи оценки его реалистичности вероятностно-статистическим методом.

1 Особенности программных проектов сложных систем

Объективная потребность контролировать процесс разработки программных проектов, прогнозировать и гарантировать стоимость, а также сроки и качество результатов разработки привела к появлению *программной инженерии* (software engineering) [1] – области компьютерных наук и технологий, связанной с индустриальным производством надежных, экономичных и высококачественных программных продуктов для решения определенной проблемы в виде комплекса взаимосвязанных программ и сопроводительной технической документации, необходимой для его эксплуатации.

Несмотря на определенные успехи, достигнутые программной инженерией на современном этапе развития, перманентный «кризис программного обеспечения» продолжается и прямо сказывается на программных проектах [2, 3]. В результате, все больше проектов становятся проблемными: они характеризуются низкой производительностью, существенным отставанием от графика или превышением запланированной сметы расходов, не в полном объеме удовлетворяют выдвигаемым требованиям заказчика как по качеству, так и по функциональным возможностям. Более того, все чаще проекты оказываются просто проваленными, т.е. не дове-

денными до конца из-за нарушения сроков, перерасхода бюджета и не соблюдения качества конечных результатов проекта.

В современной государственной научно-технической политике важную роль играют так называемые инновационные проекты [4], которые реализуются в виде крупных межотраслевых проектов по созданию, освоению и распространению технологий, способствующих кардинальным изменениям в технологическом базисе экономики, а также по развитию фундаментальных исследований, научно-техническому обеспечению социальных программ, международного сотрудничества. В настоящей работе под *программным проектом сложной системы* понимается разновидность инновационного проекта, рассматриваемая как интеллектуальная деятельность коллектива разработчиков, направленная на достижение заранее определенного результата/цели при заданных ресурсных и временных ограничениях с учетом требований к качеству. При этом результатом программного проекта является создание уникального инновационного программного продукта, который можно запускать, тестировать, исправлять и развивать.

Важнейшей особенностью программных проектов является их выполнение в условиях риска и неопределенности. При этом возможность риска наиболее велика на стадиях выработки концепции, планирования и начала работ над проектом. Именно на ранних стадиях выполнения проекта имеется возможность либо минимизировать влияние риска, либо вообще избежать его. Для снижения рисков при разработке программного проекта необходимо в первую очередь провести тщательную оценку его реалистичности. Сложность оценки реалистичности программных проектов обусловлена, следующими факторами [5]:

- во-первых, неполнотой информации, обусловленной игнорированием имеющегося большого опыта проектных организаций: реальные технико-экономические показатели удачно выполненных программных проектов не обобщаются и не используются в качестве исходной базы для управления новыми программными проектами;
- во-вторых, отсутствием полного представления о требованиях к проекту, что в свою очередь увеличивает вероятность того, что проект будет выполнен с невысоким качеством, а цели не будут достигнуты;
- в-третьих, слабой структурированностью теоретических и фактических знаний о программных проектах.

В литературе по программному инжинирингу (см., например, [5, 6]) подчеркивается, что одним из условий повышения эффективности оценки реалистичности программного проекта является использование формально математических методов, без которых дать удовлетворительный прогноз развития проекта на достаточно большом промежутке времени, опираясь только лишь на собственный опыт и интуицию практически невозможно. В настоящей работе для оценки реалистичности программных проектов предложено использовать вероятностно-статистический подход, основанный на широко распространенной в современной науке функциональной парадигме (все необходимые компоненты разрабатываемых программных средств представляют собой «черные ящики», выполняющие сформулированные в требованиях задачи функции при отсутствии описания их внутренней реализации).

2 Критерий жизнеспособности проектов

Существует множество критериев жизнеспособности проектов, позволяющих дать оценку степени обоснованности проекта с точки зрения возможностей его выполнения и воплощения: реалистичность проекта (соответствие целей и задач проекта реальной ситуации, уровень обеспеченности проекта разного рода ресурсами); реализуемость проекта (оценка возможностей конкретного исполнителя выполнить комплекс работ проекта); рентабельность проекта (величина полученного результата должна быть выше, чем размер вложенных затрат),

оптимальность проекта (лучший выбор из целесообразных вариантов в рамках существующих ограничений) и др.

В настоящей работе в качестве оценки жизнеспособности проектов рассматривается реалистичность проекта с точки зрения предполагаемых временных и ресурсных затрат на его осуществление. Судя по литературным источникам [5-8], именно недостаточный анализ реалистичности проектов, сопряженный с серьезными трудностями, является причиной малой эффективности реализации инновационных и, в частности, программных проектов. Действительно, исполнители не всегда владеют полной информацией о проекте и поэтому не в состоянии дать точную оценку его реалистичности и, прежде всего, сроков его выполнения. При этом недостаточная осведомленность исполнителей часто сопровождается изменением требований, сроков и объема ресурсов, выделяемых на проект. Данные обстоятельства, а также связанные с ними риски и последствия, естественно приводят к неизбежным ошибкам при планировании проекта, из-за которых предварительные оценки его реалистичности теряют свою актуальность.

Для оценки реалистичности проекта разрабатываются различные варианты реализации проекта. Однако это имеет смысл только в том случае, когда существует уверенность, что предложенные варианты могут в той или иной мере быть осуществлены, что проекты носят реалистичный, выполнимый характер. Здесь важную роль играет число рассмотренных вариантов реализации проекта и чем их больше, тем выше будет обоснованность принимаемых проектных решений.

Разработка вариантов реализации проекта осуществляется ответственными исполнителями под наблюдением руководителя-менеджера. Его деятельность является определяющей при уточнении критериев отбора лучшего варианта. Выбор из различных вариантов реализации программного проекта самого жизнеспособного представляет собой один из наиболее ответственных этапов разработки проекта. Ключевым здесь является критерий такого выбора, который может отражать экономичность, производительность, надежность, качество и прочие показатели проекта.

В настоящей работе в качестве критерия реалистичности программных проектов рассматривается вероятность выполнения совокупности работ за выделенное время.

3 Задачи распределения ресурсов в управлении проектами

На разных этапах жизненного цикла программного проекта возникает широкий спектр задач, которые являются следствием манифеста программной инженерии: «максимизация качества проекта при минимуме затрат». Ясно, что успешное решение данных задач обеспечивает успешное выполнение проекта. Можно выделить следующий класс широко распространенных задач программного инжиниринга, которые стоят перед менеджером современных инновационных программных проектов:

- *Оптимизации проектов по стоимости и срокам.* После проведения анализа стоимости проекта, принимается решение относительно сокращения или увеличения затрат на задачи или ресурсы определенного типа [9-12].
- *Минимизация затрат при формировании команды исполнителей проекта.* Поскольку скорость вывода новых программных продуктов на рынок в современных условиях является конкурентным преимуществом, то естественно возникает задача формирования такой команды исполнителей проекта, которая бы осуществила выполнение всех работ в заданные сроки, и при этом затраты на оплату труда были минимальны [13, 14].
- *Минимизация упущенной выгоды.* Поскольку в условиях дефицита финансовых ресурсов нет возможности одновременно реализовать все проекты календарного плана программы, то необходимо выбрать первоочередные проекты, реализация которых обеспечивает наибольший эффект. Остальные проекты отодвигаются на более поздние сроки, и ставится

задача построения графика финансирования проектов минимизирующей упущенную выгоду [15, 16].

- *Выбор наиболее приоритетных работ при заданном ограничении на время выполнения общего набора работ* [17].
- *Управление стоимостью компании.* Здесь рассматривается задача выбора подмножества проектов из множества возможных, причем в качестве ограничения могут выступать: бюджет, производственные мощности, количество свободных сотрудников для осуществления проектов, наличие помещений для ведения проектов и т.д. [18].

Решение данных задач может быть положено в основу для принятия решений по значениям выходных параметров программного проекта.

Следует отметить, что основой оптимального распределения ресурсов является их ограниченность, что требует их использования (соответственно распределения) с учетом критерия оптимальности (времени выполнения проекта, затрат, риска и др.). Проблема оптимального распределения ресурсов решается с помощью математических моделей (линейного и нелинейного программирования, динамического программирования и т.д.).

В настоящей работе рассмотрена одна из актуальных задач программного инжиниринга на основе критерия жизнеспособности программного проекта - задача о реалистичности проекта при заданных ограничениях на объем работ по проекту, на время выполнения проекта, на производительность исполнителей.

В основу предлагаемого подхода положено использование системной модели оценки реалистичности программного проекта, представленной на рисунке 1.

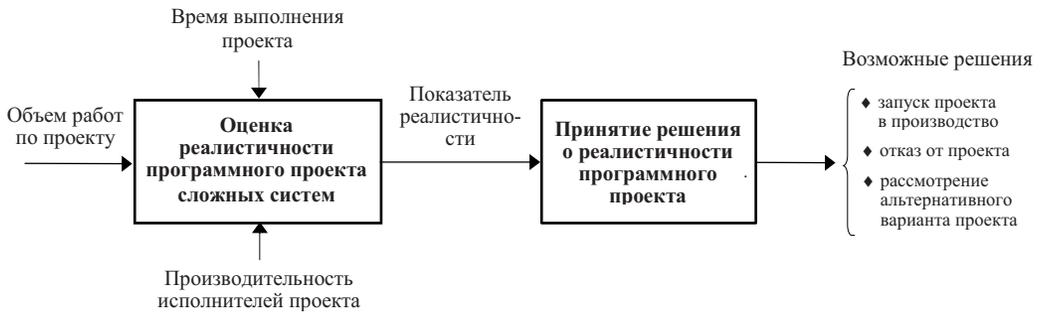


Рисунок 1 – Системная модель оценки реалистичности программного проекта

При постановке задач ресурсного планирования предполагается, что проект описывается в виде сетевого графика выполнения работ. При этом в силу неоднозначности выделения работ одному и тому же проекту можно поставить в соответствии различные варианты сетевых графиков. Предполагается, что любой такой сетевой график можно представить в виде последовательно-параллельных схем выполнения работ. Работы, выделенные на сетевых графиках равнозначны, а правила преобразования сетевых графиков последовательно-параллельных схем описаны в работе [19].

Предлагаемый подход к оценке реалистичности программного проекта основан на формировании специальной скалярной характеристики $\alpha^{(i)}$, значение которой определяется особенностями программного проекта:

$$\alpha^{(i)} = \min_i \{p_j^{(i)}\}, \quad i = \overline{1; m}, \quad j = \overline{1; k_i},$$

где m – число вариантов реализации проекта, k_i – количество видов работ в i -м варианте реализации проекта, $p_j^{(i)}$ – вероятность выполнения j -й работы в i -м варианте реализации проекта, представленных на i -м сетевом графике при заданных ограничениях на время $T_j^{(i)}$ выполнения каждой работы.

Полагаем, что вероятность $p_j^{(i)}$ зависит от времени $T_j^{(i)}$, выделяемого на выполнение j -й работы при i -м варианте реализации проекта, а также от производительности исполнителей. В результате, на ранней стадии реализации программного проекта возникает задача, связанная с проведением сравнительного анализа альтернативных вариантов сетевых графиков выполнения работ, позволяющего принять решение об отказе от данного варианта проекта или о запуске его в производство, либо о генерации его новых альтернативных вариантов.

4 Вероятностно-статистический метод оценки реалистичности программных проектов

Пусть задана производительность исполнителей $t_n = \varphi_{j,n}(V_j^{(i)})$, а также варианты сетевых графиков выполнения работ с указанием объема для каждой работы $V_j^{(i)}$ ($i = \overline{1;m}, j = \overline{1;k_i}$), где m - число вариантов реализации проекта, k_i - количество видов работ. Ставится задача расчета скалярной характеристики $\alpha^{(i)}$, характеризующей реалистичность программного проекта [20]. Предлагается алгоритм решения данной задачи, который для i -го варианта реализации программного проекта содержит следующие четыре этапа:

- *на первом этапе* рассчитываются допустимые времена выполнения работ $T_j^{(i)}$ для различных вариантов сетевых графиков выполнения работ с указанными объемами работ $V_j^{(i)}$ и производительностями исполнителей.
- *на втором этапе* рассчитываются вероятности выполнения работ $p_j^{(i)}$ для различных вариантов сетевых графиков с допустимыми временами выполнения работ $T_j^{(i)}$.
- *на третьем этапе* формируется скалярная характеристика реалистичности программного проекта $\alpha^{(i)}$ для каждого альтернативного варианта реализации программного проекта, которая определяется минимальной вероятностью выполнения работ $\alpha^{(i)} = \min_i \{p_j^{(i)}\}$.
- *на четвертом этапе* осуществляется принятие решений (одно из решений, выделенных на рисунке) на основе анализа множества характеристик $\alpha^{(i)}$.

Здесь первые два этапа алгоритма реализуют предложенный автором вероятностно-статистический метод «Сдвига» [20], позволяющий рассчитать вероятности для каждой работы и заключающийся в постановке и решении обратной задачи, по отношению к известной прямой задаче оценивания закона распределения функции случайного аргумента [21].

В контексте рассматриваемой задачи роль времени выполнения работ $T_j^{(i)}$ играет случайная величина X ; объемов работ $V_j^{(i)}$ - случайная величина Y , а производительности исполнителей $t_n = \varphi_{j,n}(V_j^{(i)})$ - непрерывная, строго монотонная возрастающая функциональная зависимость $y = \varphi(x)$. Тогда, если имеются выборочные значения случайной величины Y , возможные значения которой принадлежат интервалу $\{a^y, b^y\}$, а также зависимость $y = \varphi(x)$, то ставится и решается задача нахождения закона распределения $F_1(x)$ случайной величины X , возможные значения которой принадлежат интервалу $\{a^x, b^x\}$.

Заключение

Для решения задач программного инжиниринга можно использовать достижения одного из наиболее эффективных интеллектуальных инструментов менеджера – методов и технологий принятия решений в условиях неопределенности [22], включая методы, альтернативные

вероятностно-статистическому: интервальные, нечеткие и возможностные методы с привлечением нестохастических моделей неопределенности - субъективные вероятности Л. Севеджа и А. Демпстера, правдоподобие и доверие Ж. Шеффера, емкость Ж. Шоке и др.

Благодарности

Автор выражает благодарность В.Е. Гвоздеву. Работа выполнена в рамках исследований, проводимых при поддержке РФФИ (грант РФФИ 10-08-00359-а).

Список литературы

- [1] Липаев В.В. Программная инженерия. Методологические основы: Учебник. М.: ГУ-ВШЭ, ТЕИС. 2006.
- [2] Карпенко С.Н. Введение в программную инженерию. Н.Новгород, 2007 (<http://www.unn.ru/pages/issues/aids/2007/16.pdf>).
- [3] Архипенков С.Я. Лекции по управлению программными проектами. М., 2009 (http://www.arkhipenkov.ru/resources/sw_project_management.pdf).
- [4] Гордова И.Б. Управление инновационными процессами. Кемерово: КемТИПП, 2005.
- [5] Макконенелл С. Сколько стоит программный проект. М.: Русская редакция, СПб.: Питер, 2007.
- [6] Беркун С. Искусство управления IT-проектами. СПб.: Питер, 2007.
- [7] Бэбьюли Ф. Управление проектом. М.: ФАИР-ПРЕСС, 2002.
- [8] Липаев В.В. Проектирование программных средств. М.: Высшая школа, 1990.
- [9] Баркалов П.С., Буркова И.В., Глаголев А.В., Колпачев В.Н. Задачи распределения ресурсов в управлении проектами. М.: ИПУ РАН, 2002.
- [10] Маркаров Д.А. Частная задача оптимизации сроков при управлении проектами // Управление большими системами. М.: ИПУ РАН, 2005. С.53-59.
- [11] Румянцев В.П., Низаметдинов Т.Ш. Оптимизационная задача оценки объемов, бюджета и сроков программных проектов // Системы проектирования, технологической подготовки производства и управления этапами жизненного цикла промышленного продукта CAD/CAM/PDM: IX Международная конференция. Москва, 2007.
- [12] Уандыков Б.К. Оптимизация программ по стоимости // Теория активных систем: международная научно-практическая конференция. Москва, 2003.
- [13] Галинска Е.А., Половинкина А.И., Рентеева Е.Л. Модель минимизации затрат при формировании команды информационного проекта // Управление большими системами. Выпуск 14. Воронеж: ВГАСУ, 2006. С.29-33.
- [14] Беляев Ю.А., Врублевская С.С., Половинкина А.И., Рентеева Е.Л. Модель минимизации затрат при формировании команды проекта // Сборник научных трудов СевКавГТУ, 2007. № 3.
- [15] Бурков В.Н., Квон О.Ф., Цитович Л.А. Модели и методы мультипроектного управления. М.: ИПУ РАН, 1998.
- [16] Баркалов С.А., Бурков В.Н. Минимизация упущенной выгоды в задачах управления проектами. М.: ИПУ РАН, 2001.
- [17] Бекмурзаев В.А., Шеховцов А.А. Использование метода динамического программирования и его оптимизация при решении задач управления проектами // Журнал научных публикаций аспирантов и докторантов. 2008 (<http://www.jurnal.org/articles/2008/inf18.html>).
- [18] Романов В.С. Задача управления стоимостью компании – дискретный случай // Управление большими системами. М.: ИПУ РАН, 2006. - С. 142-152.
- [19] Барлоу Р., Прошан Ф. Математическая теория надежности. М.: Советское радио, 1969.
- [20] Колоденкова А.Е. Статистический подход к оценке реалистичности программных проектов для автоматизированных информационно-управляющих систем // Мехатроника, автоматизация, управление. 2010. № 4. - С. 52-60.
- [21] Пугачев В.С. Теория вероятностей и математическая статистика. М.: ФИЗМАТЛИТ, 2002.
- [22] Орлов А.И. Теория принятия решений. М.: Изд-во «Март», 2004.

К ВОПРОСУ О ПРОГНОЗИРОВАНИИ КОЛИЧЕСТВЕННЫХ ПАРАМЕТРОВ СЛОЖНЫХ СИСТЕМ ПРИ ИСПОЛЬЗОВАНИИ НЕЧЁТКОЙ ИНФОРМАЦИИ

Г.И. Горемыкина, М.А. Пономарева

Балашовский институт Саратовского государственного университета
412300, Балашов, ул. К. Маркса, 19, Россия
mig0109@mail.ru
тел: +7 (84545) 4-24-10

Ключевые слова: прогнозирование, нечёткая информация, гибридная сеть, экспертная оценка, нечеткая кластеризация, гранулированность информации

Abstract

In the article authors describe prediction of quantitative parameters of complex systems under the use of fuzzy information. The result was used as input data developed by the expert system for managing business processes.

Введение

Прогнозирование количественных параметров сложных систем является значимой проблемой, как на этапе создания таких систем, так и на этапе их последующего функционирования. Задача усложняется, когда решение приходится принимать при использовании информации высокой степени гранулированности. Такая ситуация возникает, например, когда отсутствует возможность получения достаточного объёма релевантной информации о прогнозируемых параметрах или решение принимается в быстро изменяющихся условиях. Если осуществляется прогнозирование экономических параметров, то информационная гранулированности может быть индуцирована нестабильной экономической ситуацией.

Описанная проблема возникла перед авторами при создании интеллектуальной системы поддержки принятия решения об управлении бизнес-процессами при использовании нечёткой информации. Указанная система, учитывая стратегию инновационного поведения предприятия, реализует принятие решения об управлении инвестиционными проектами на основе их оценки риска и степени привлекательности. Процесс разработки системы был реализован в среде Mat Lab с использованием пакета Fuzzy Logic Toolbox (более подробно см. [1]).

Для прогнозирования входных параметров был применен алгоритм нейронных сетей, что обусловлено резкими колебаниями цен и издержек, а также наличием неявных тенденций в динамике изменения объёма реализации.

1 Обучение нейронечеткой сети ANFIS для прогнозирования количественных параметров интеллектуальной системы

Рассмотрим применение нейронечеткой сети *ANFIS* для прогнозирования цены. В качестве исходных данных была использована информация о ценах на муку за последние 5 лет из архива предприятия. Как показывает практика, существуют сезонные колебания цены на муку: цена падает с августа по январь и повышается с февраля по июль. Поэтому прогнозирование цены осуществлялось двумя способами: посезонно и помесечно. Вследствие этого для обучения нейронной сети авторами были сформированы две выборки, каждая из которых отражает ту или иную тенденцию изменения цены. Кроме того при использовании двух выборок повышается точность прогноза.

Таблица 1 – Цена (тыс. руб. за 1 тонну) муки пшеничной высшего сорта: выборка I

Месяц Год	Февраль	Март	Апрель	Май	Июнь	Июль
2003	11,42	11,49	11,56	11,60	11,63	11,69
2004	11,52	11,55	11,56	11,59	11,62	11,78
2005	11,65	11,69	11,70	11,71	11,71	11,73
2006	11,78	11,81	11,82	11,83	11,89	12,05
2007	11,92	11,96	11,98	11,99	12,05	12,09
2008	11,90	11,91	11,98	11,99	12,06	12,10

Таблица 2 – Цена (тыс. руб. за 1 тонну) муки пшеничной высшего сорта: выборка II

Месяц Год	Август	Сентябрь	Октябрь	Ноябрь	Декабрь	Январь
2002-2003	11,53	11,51	11,46	11,40	11,33	11,33
2003-2004	11,68	11,66	11,65	11,60	11,55	11,50
2004-2005	11,78	11,76	11,76	11,75	11,70	11,62
2005-2006	11,99	12,00	11,80	11,70	11,69	11,70
2006-2007	12,06	12,02	12,01	11,90	11,80	11,79
2007-2008	12,10	11,99	11,92	11,90	11,86	11,85

Входными переменными в первой выборке являются цены на муку в каждом месяце с февраля по июнь, а выходной переменной — цена за июль. Для генерации нейронной сети был выбран режим *Grid partitions* (носители выбираются независимо для каждого входа выборки). Для каждой входной переменной были заданы по три функции принадлежности типа *trimf*, в качестве выходной — линейная функция (*linear*), выбор последнего пункта позволяет разработать модель методом Сугено первого порядка [3].

Для обучения нейронной сети авторами был выбран гибридный способ. Целевое значение ошибки было установлено равным нулю, а количество циклов — трем.

Структура построенной нечеткой модели после обучения гибридной сети представлена на рисунке 1. Недостаточная графическая наглядность построенной модели объясняется большим количеством правил (равным 243) в разработанной системе нейронечеткого вывода.

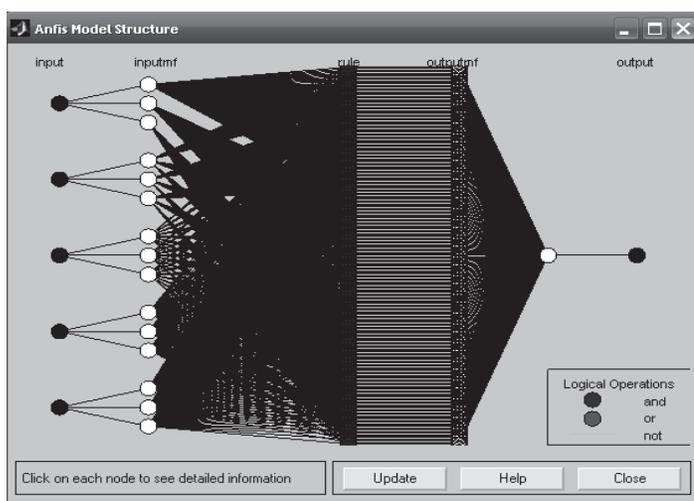
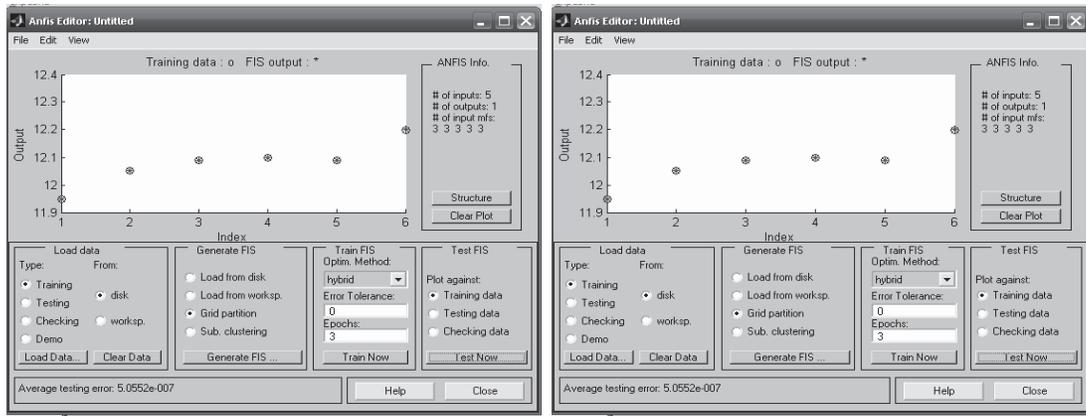


Рисунок 1 – Структура гибридной системы класса ANFIS для прогнозирования цены

Результаты аппроксимаций загруженных данных с помощью гибридного алгоритма изображены на рисунке 2. Здесь и далее: * – реальные значения цены, * – прогнозируемые значения цены.



а)

б)

Рисунок 2 – Результат аппроксимации загруженных данных с помощью гибридного алгоритма: а – выборка I; б – выборка II

После проведения прогнозирования цен на муку с августа 2008 по июль 2009 было проведено также прогнозирование цены каждого месяца исследуемого года с учетом динамики цен в каждом месяце предыдущих лет (2002-2007 гг.). Это было сделано с целью попытки выявить, как влияет, например, на цены февраля 2009 года цены в этом же месяце, но в 2002-2007 гг. Полученная с помощью прогнозирования динамика цен на август 2008 – июль 2009 представлена на рисунке 3.

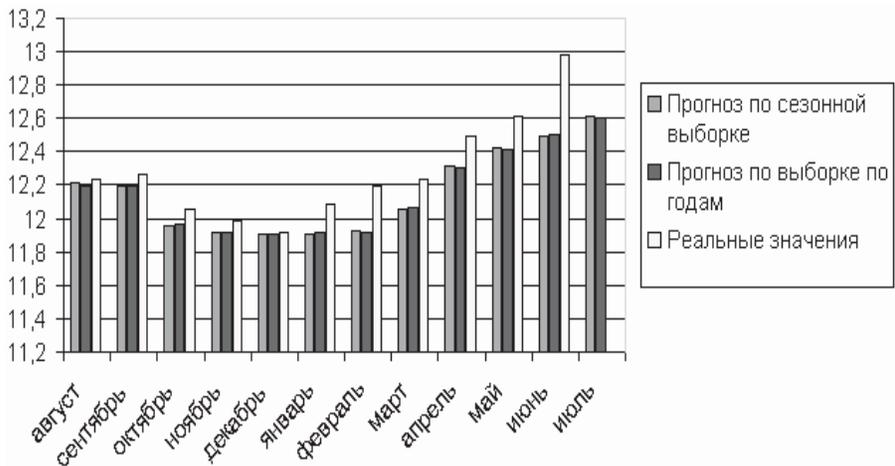


Рисунок 3 – Прогноз цен на муку на август 2008 - июль 2009 гг.

В период окончания реконструкции и начала выпуска продукции (когда риск проекта максимален) был сделан прогноз на цены января двумя способами: с помощью сезонной выборки и выборки по одному и тому же месяцу 2002-2007 гг. Полученные прогнозные значения показаны соответственно на рисунках 4,а и 4,б.

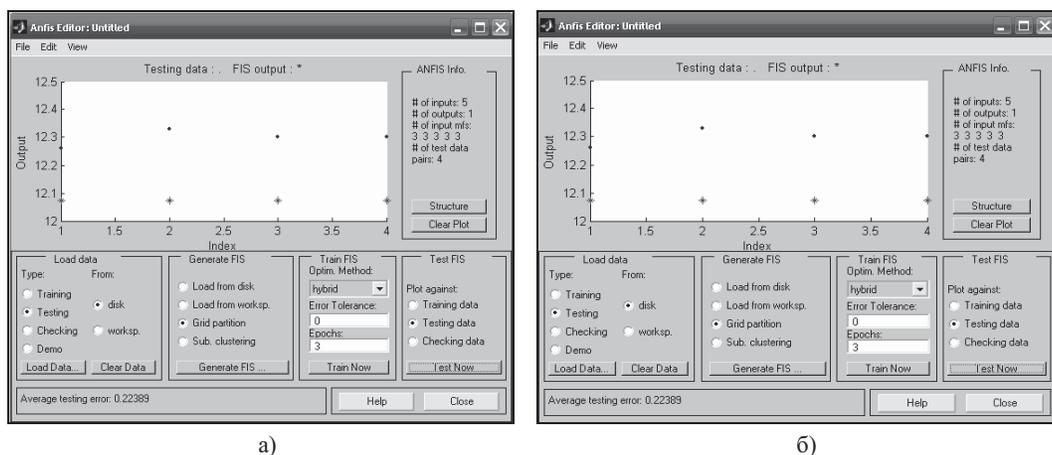


Рисунок 4 – Тестирующая выборка цен за январь с учетом:
 а – сезонного прогноза, б – прогноза по годам

Анализируя различные подходы к прогнозированию цен, стоит отметить следующее:

- результаты прогнозирования по сезонам и по годам практически не отличаются;
- оба прогноза сильно отличаются от реальных значений цен.

Первое, по мнению авторов, объясняется тем, что нейронная сеть адекватно отражает скрытые закономерности динамики цен в период с 2002 по 2007 гг. Второе же означает, что, несмотря на наметившуюся тенденцию, нейронной сети не хватает накопленной информации для прогнозирования цен в условиях нестабильной экономической ситуации. Из-за невозможности расширения объёма обучающей выборки (полное отсутствие данных в условиях сложившейся нестабильной экономической ситуации) нельзя уловить в полной мере те скрытые закономерности, которые позволяют отыскивать в базах данных нейросетевые технологии. В связи с этим было принято решение о привлечении группы квалифицированных экспертов. Однако эксперты по-разному оценили значения цены и затруднились задать ее точно. Поэтому было решено применить теорию нечетких множеств, которая позволяет задать значение цены в виде нечеткого числа. Понятия «нечёткое множество» и «нечеткое число» употребляется в соответствии с терминологией [3,4].

2 Применение нейросетевых технологий для исследования влияния мнения эксперта на спрогнозированный обобщенный параметр

Существуют методики проведения экспертной оценки, при которых применяется ранжирование экспертов – каждому из экспертов присваивается весовой коэффициент. Нередко ЛПР присваивает больший весовой коэффициент эксперту с позиции занимаемой этим экспертом должности, а не с точки зрения точности даваемых им прогнозов. Поэтому авторы приняли решение использовать нейросетевые технологии, с помощью которых эта ситуация элиминируется.

Для оценки качества прогнозов, даваемых тем или иным экспертом, был проведен ретроспективный анализ прогнозирования цен по данным из имеющегося на предприятии архива прогнозов. В качестве входных данных для построения нейронной сети были заданы прогнозируемые экспертами цены, выходной переменной является реальная цена. При этом решено было ограничиться лишь выборкой по годам. В результате обучения нейронная сеть определила степень влияния входных переменных (прогноз каждого эксперта) на величину реальной цены.

Для обучения нейронной сети был использован алгоритм нечеткой кластеризации [3], который позволил сгруппировать прогнозы экспертов для наиболее точного прогнозирования цены с учетом мнения всех экспертов с их весовыми коэффициентами. Это, в свою очередь, позволило уменьшить количество правил вывода и свести их всего к четырем (рисунок 5). Структура построенной гибридной сети представлена на рисунке 6.

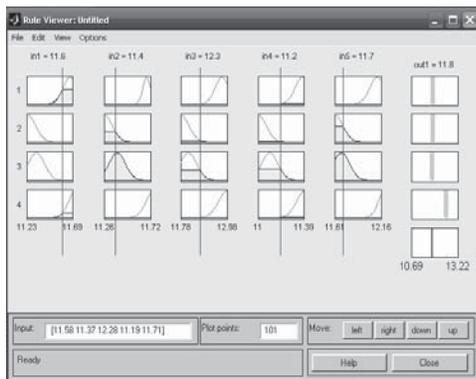


Рисунок 5 – Окно правил вывода, сформированных после применения нечеткой кластеризации

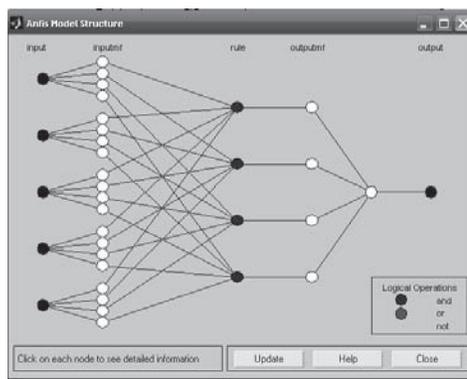


Рисунок 6 – Структура нечеткой гибридной сети прогнозирования цены на основе экспертной оценки

Для прогнозирования цены на февраль-июль 2009г. экспертам было предложено задать цену треугольным числом A , указав его характеристические параметры, см. рисунок 7 и таблицу 3.

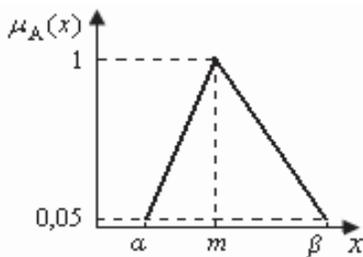


Рисунок 7 – Характеристические параметры нечеткого числа A = «примерно m »

Выбор кусочно-линейной функции принадлежности был в значительной степени обусловлен малым объемом имеющейся информации о прогнозируемых параметрах, а также тем, что решение принималось в быстро изменяющихся условиях. Чтобы получить прогноз, отражающий мнение каждого из экспертов с учетом его веса, и получить треугольное значение цены, необходимое для работы экспертной системы управления инвестиционными проектами авторы использовали построенную нейронную сеть, структура которой изображена на рисунке 6.

Таблица 3 – Прогнозируемая цена (треугольное значение) на февраль-июль 2009 года

№ эксперта	1	2	3	4	5	ANFIS
α	11,35	11,39	12,36	11,17	11,79	11,81
m	11,43	11,45	12,51	11,20	11,94	11,90
β	11,50	11,59	12,71	11,22	12,01	12,16

Руководство предприятия, на котором проводились исследования, заинтересовалось влиянием мнения каждого из экспертов на прогнозирование итоговой цены. Поэтому была представлена графическая зависимость обобщенной спрогнозированной цены от мнения каждого из экспертов (рисунок 8).

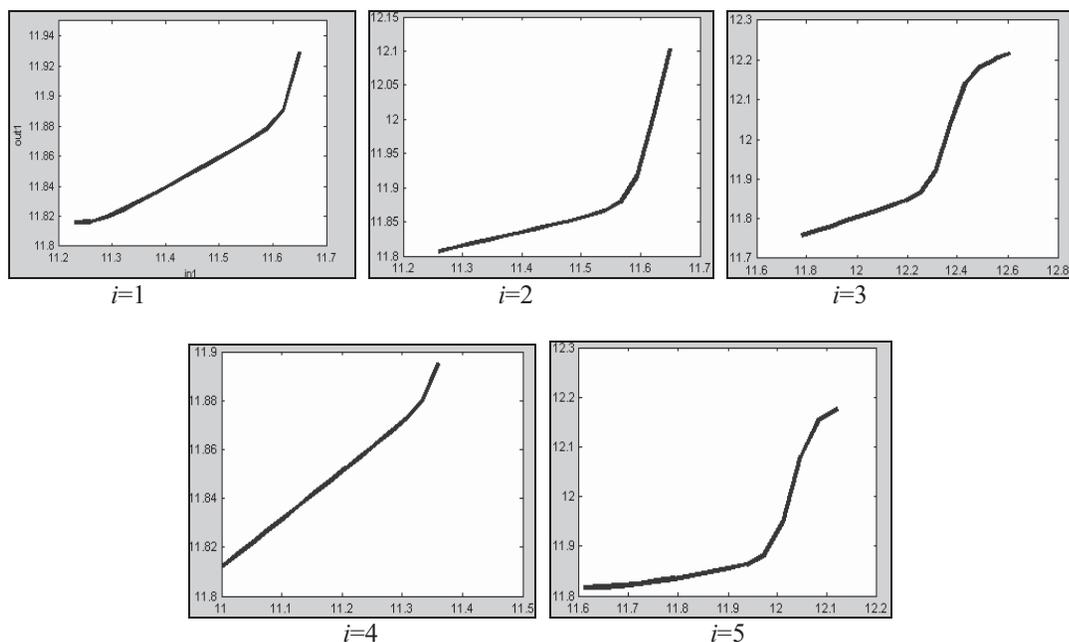


Рисунок 8 – Графики зависимости обобщенной спрогнозированной цены от мнения i -го эксперта ($i=1,2,3,4,5$)

По графикам, представленным на рисунке 8, нетрудно определить, мнение каких экспертов наиболее весомо для прогнозируемой итоговой цены. Проанализировав каждую из зависимостей, можно сказать, что наиболее весомым является вклад в прогнозирование пятого эксперта, а наименьшим весом обладает четвертый эксперт. В связи с этим интересно было бы посмотреть влияние комбинаций мнений экспертов с разными весовыми коэффициентами на обобщенную цену (рисунок 9). На рисунке 9,*а* представлена комбинация «слабый»-«слабый», 9,*б* – «слабый»-«сильный», 9,*в* — «средний»-«средний», 9,*г* — «сильный»-«сильный», где слова «слабый», «средний», «сильный» указывают на эксперта соответственно с низкой, средней и высокой степенью влияния на прогнозирование цены (что, в свою очередь, соответствует низкому, среднему и высокому весовым коэффициентам).

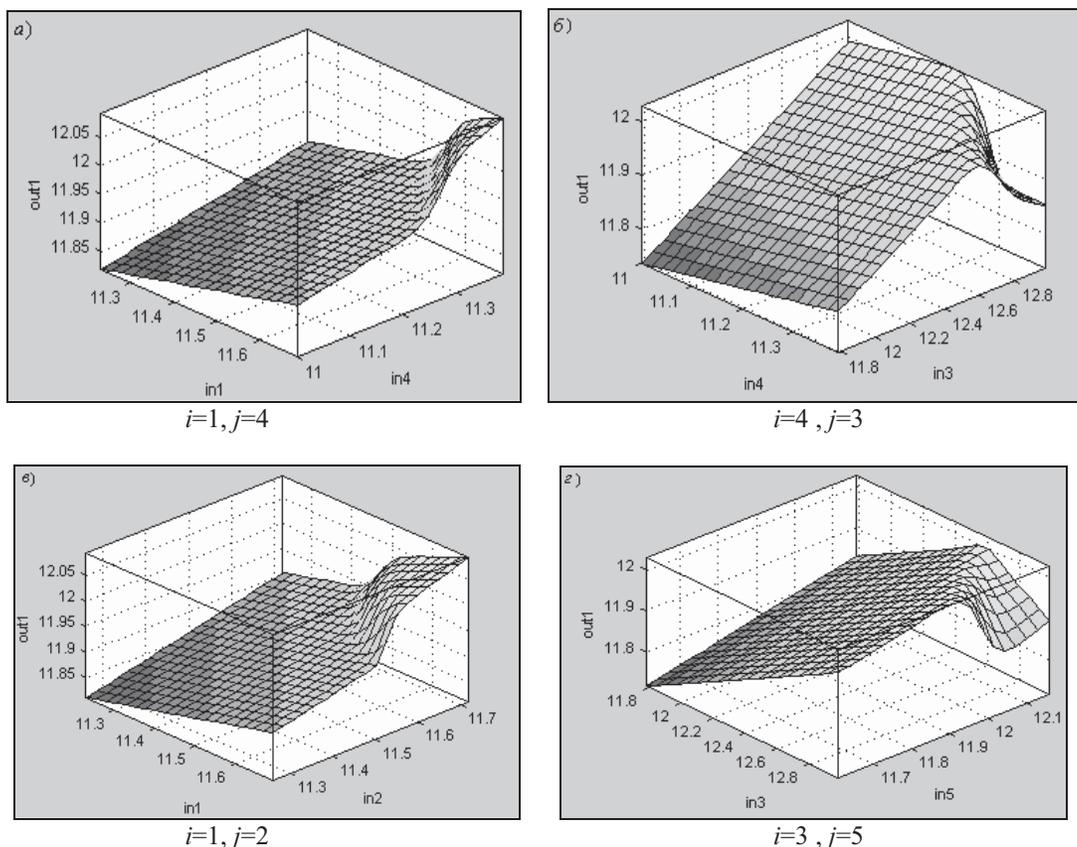


Рисунок 9 – Графики зависимости обобщенной спрогнозированной цены от комбинаций прогнозов i -го и j -го экспертов ($i, j = 1, 2, 3, 4, 5$)

Аналогично прогнозированию цены были спрогнозированы переменные издержки на единицу продукции, а также объём реализации. Полученные таким образом данные были использованы в качестве входных данных разработанной системы.

Заключение

Проверка построенной нечеткой модели гибридной сети показала достаточно высокую степень ее адекватности реальным исходным данным, что позволило использовать ее для прогнозирования входных параметров, построенной интеллектуальной информационной системы.

Список литературы

- [1] Горемыкина, Г.И., Пономарева, М.А. К вопросу о применении имитационного моделирования в управлении бизнес-процессами в условиях нестабильной экономической ситуации// Проблемы управления и моделирования в сложных системах: Труды XX международной конференции (22-24 июня 2009 г., Самара, Россия). – Самара: Самарский НИЦ РАН, 2009. - С. 503-508.
- [2] Пегат, А. нечеткое моделирование и управление. М.: Лаборатория знаний, 2009.- 798с.
- [3] Ярушкина, Н.Г. Основы теории нечетких и гибридных систем. М.: Финансы и статистика, 2004. – 320с.
- [4] Zadeh L.A. Fuzzy sets // Inform. Contr. 1965. V.8. P. 338—353.

ПОДХОД К УПРАВЛЕНИЮ РИСКАМИ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ НА ОСНОВЕ ПРОДУКЦИОННО-ФРЕЙМОВОЙ МОДЕЛИ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ

А.С. Потапов, И.В. Аникин

Казанский государственный технический университет им. А.Н. Туполева
420111, Казань, ул. К.Маркса, 10, Республика Татарстан, Россия
anikinigor@evm.kstu-kai.ru
тел: +7 (843) 231-00-56

Ключевые слова: управление рисками, оценка рисков, фрейм, правила продукции, экспертные системы

Abstract

In this clause I described a risk management approach based on a knowledge representation model which consisted of frames and production rules. New risk management techniques based on representation model were proposed. This techniques were used in a proof-of-concept consulting program. Using this approach various criteria of selection of countermeasures to reduce information security threats were chosen. Besides that a selection algorithm was implemented that allows to choose the most efficient countermeasure corresponding to required criteria.

Введение

Управление рисками — это процедуры и действия, которые позволяют выявлять, оценивать, отслеживать и устранять риски до или во время их превращения в проблемы.

В рамках процесса оценки рисков особое место занимают экспертные системы, применение которых позволяет ускорить и упростить данный процесс. При разработке экспертных систем ключевыми являются вопросы выбора модели представления знаний и работы аппарата логического вывода.

По завершению процесса оценки в рамках управления рисков необходимо снизить, либо устранить найденные риски. Для этого необходимо решить задачу выбора наиболее эффективного набора контрмер по некоторым критериям.

Продукционно-фреймовая модель представления знаний

Риск – это вероятный ущерб, который зависит от защищенности системы. В рамках оценки рисков активно используются процедуры экспертной оценки. Однако недопустимо использовать только «ручной» труд эксперта, так как в этом случае поставленную задачу оценки рисков невозможно будет решить в приемлемые сроки с приемлемым бюджетом. Поэтому для более быстрой и качественной оценки применяются разнообразные системы автоматизации. Оценка рисков ИБ представляет собой трудно формализуемую задачу, поэтому для ее решение актуально использование технологии экспертных систем.

Основными компонентами любой экспертной системы являются база знаний и аппарат логического вывода. База знаний должна быть наиболее наглядной и позволять строить структурные модели информационной системы, модели угроз и рисков, связанных с отдельными составляющими информационной системы. Это позволит выявить те элементы и объекты информационной системы, риск нарушения безопасности которых, является неприемлемым.

Была сформирована онтология атак на корпоративную информационную сеть, после чего сформирована база знаний экспертной систем. В современных экспертных системах существует 4 основных способа представления знаний в базе знаний [1]:

- логические модели;
- производственные модели;
- семантическая сеть;
- фреймовые модели.

Наиболее подходящим представлением с точки зрения наглядности и структурированности системы является представление с помощью фреймовых моделей.

Фреймы имеют вид структурированных компонентов ситуаций, называемых слотами. [2] Слот может указывать на другой фрейм, устанавливая, таким образом, связь между двумя фреймами. С каждым фреймом ассоциируется разнообразная информация.

Было создано дерево фреймов, т.е. модель для решения задач оценки рисков информационной безопасности. Несколько уровней созданного дерева фреймов приведено на рисунке 1.

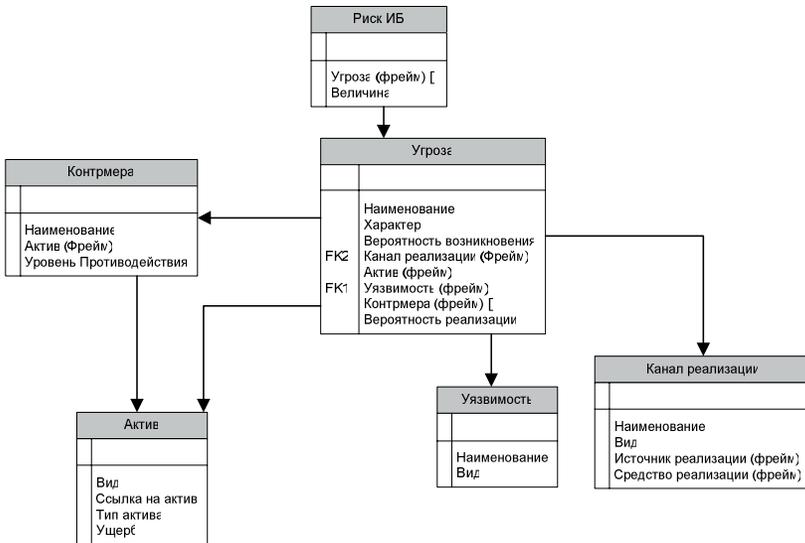


Рисунок 1 – Фрагмент фреймовой модели

Использование системы фреймов обеспечивает наглядность и простоту представление знаний. Правила продукции используются для построения процедур оценки рисков. Они обладают следующими привлекательными свойствами:

- модульность: каждое правило описывает небольшой, относительно независимый фрагмент знаний;
- возможность инкрементного наращивания: добавление новых правил в базу знаний происходит относительно независимо от других правил;
- удобство модификации (как следствие модульности): старые правила можно изменять и заменять на новые относительно независимо от других правил;
- применение правил способствует прозрачности системы.

Производственное правило в общем случае представляется в виде:

$i: S; C; A \rightarrow B; P,$

где i – номер продукции;

S – описание класса ситуаций, в котором эта структура может использоваться;

C – условие, при котором данная продукция активизируется;

$A \rightarrow B$ – ядро продукции (например, «ЕСЛИ A_1, A_2, \dots, A_n , ТО B »);

P – постусловие производственного правила, определяющее действия, которые необходимо произвести после выполнения B .

В случае оценки рисков ИБ, условиями A1, A2, ... An являются условия связанные со значениями слотов каждого фрейма. Например, в правиле:

Если ((Угроза.Характер = «Нарушение конфиденциальности») И (Угроза.Актив.Тип актива = «Информационный Актив»)) Угроза.Актив.Ущерб = Угроза.Информационный Актив.Конфиденциальность

условием A1 является (Угроза.Характер = «Нарушение конфиденциальности»), а условием A2 является (Угроза.Актив.Тип актива = «Информационный Актив»).

Действие В представимо в виде установки значения одного из слотов текущего фрейма либо в установке какого-либо критерия, полезного в дальнейшем рассмотрении. Его можно представить в виде:

В: D = F,

где В – действие;

D – слот фрейма либо критерий, который необходимо установить;

F – фактическое значение которое необходимо установить либо действие, результатом которого является устанавливаемое значение.

Таким образом, была сформирована база знаний для экспертной системой для оценки рисков ИБ на основе продукционно-фреймового представления знаний.

Экспертная система оценки рисков

На основе предложенной модели, была разработана экспертная система оценки рисков ИБ. Экспертная система имеет модульную структуру, что позволяет без лишних усилий модернизировать или заменять модули системы, не затрагивая ядро. Алгоритм функционирования системы представлен на рисунке 2 и состоит из следующих шагов.

- 1) Создание фреймовой модели представления данных. Модель представляет собой дерево фреймов.
- 2) Создание рабочих переменных в рабочей памяти. Выставление их значений.
- 3) Формирование процедур для фреймовой модели. На этом этапе используются: сформированная на первом этапе фреймовая модель и, при необходимости, созданные рабочие переменные.
- 4) В случае, если необходимо получить данные из внешних источников данных, например из отчетов XSpider, вызывается модуль получения данных из внешних источников и указывается файл отчета.
- 5) Заполнение фреймовой модели значениями для конкретной системы.
- 6) Запускается аппарат логического вывода. Если были созданы процедуры, требующие для получения значений слотов – запросы к пользователю или к БД, эксперту необходимо вводить таковые значения или выбирать из хранящихся в БД по мере появления запросов.
- 7) По окончании работы аппарата логического вывода, появляются результаты работы системы. Эксперту необходимо оценить полученные результаты, провести их анализ и сделать соответствующие выводы.

Алгоритм функционирования аппарата логического вывода приведен на рисунке 3. В основе алгоритма лежит обход дерева фреймов в глубину, начиная с нижнего уровня. По каждому фрейму выполняются все заданные процедуры и выставляются значения указанных в них слотов. Предложено 3 варианта получения значения слотов: запрос к пользователю, запрос к БД, вариант вычисления. Соответственно, при получении значения и выполняется одно из вышеперечисленных действий

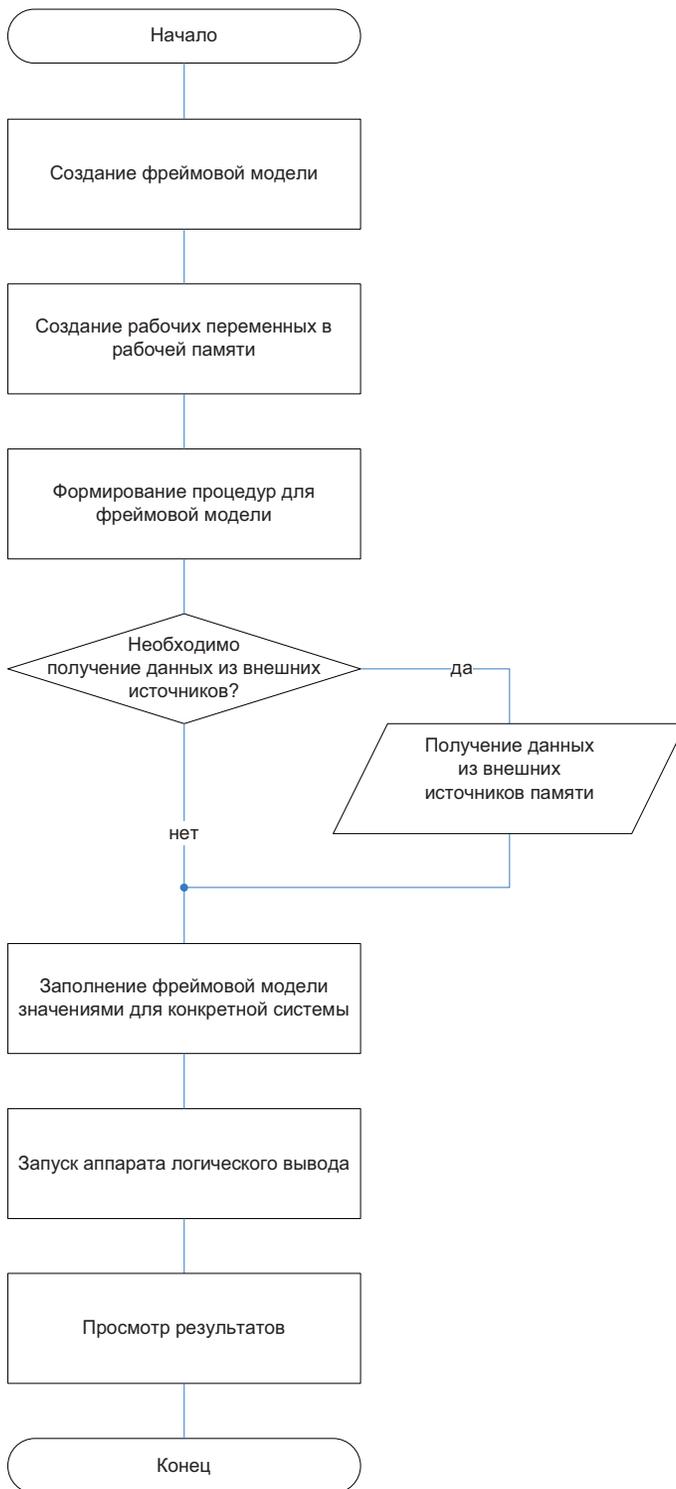


Рисунок 2 – Алгоритм функционирования разработанной ЭС

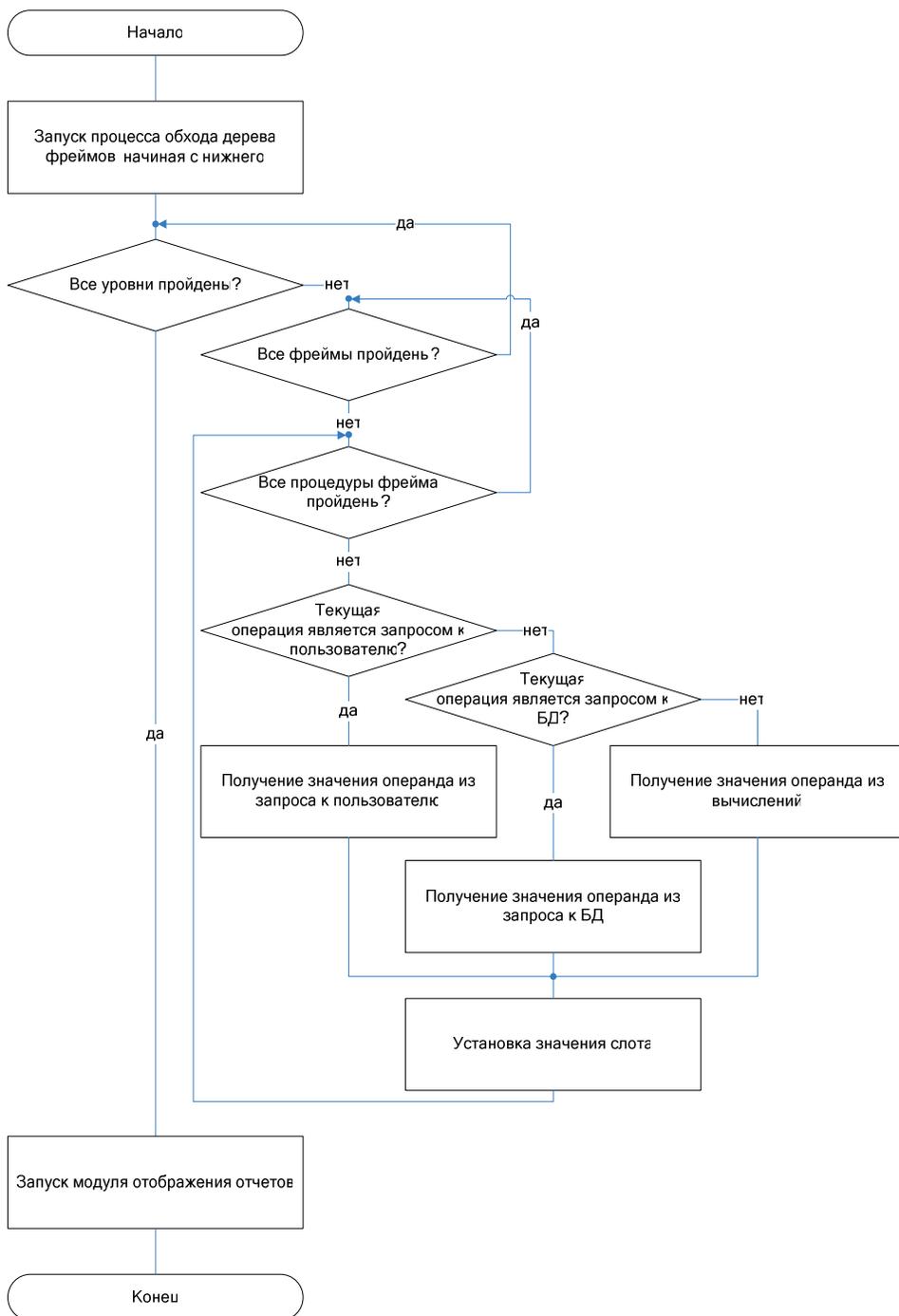


Рисунок 3 – Алгоритм функционирования аппарата логического вывода

Выбор контрмер угроз информационной безопасности

После оценки риска у эксперта возникает вопрос выработки рекомендаций по устранению рисков. От всех рисков избавиться невозможно, так как затраты (как временные, так материальные) на построения контрмер всех выявленных угроз несомненно превзойдут ценность тех ресурсов, которые необходимо защитить. Поэтому необходимо выбрать только оптимальный набор контрмер. На первый взгляд выбраны должны быть контрмерами, которые способны максимально снизить риск информационной безопасности. Однако, эффективность снижения рисков является лишь одним из критериев, по которым следует выбрать контрмеры. Так, например, при выборе из двух разных контрмер, может сложиться ситуация, что контрмера, снижающая вероятность реализации угрозы лишь на несколько процентов, по сравнению с аналогом, будет стоить в несколько раз дороже. Поэтому, первичной задачей является выделить тот минимальный набор критериев, по которому можно эффективно оценить преимущества и недостатки контрмер. Ключевыми из них являются:

- уровень снижения угрозы от внедрения контрмеры;
- экономические затраты, связанные с внедрением контрмеры;
- экономические затраты, связанные с поддержкой контрмеры за определенный период использования;
- затраты человеческих ресурсов, связанные с внедрением контрмеры;
- затраты человеческих ресурсов, с поддержкой контрмеры за определенный период использования.

Однако, даже эти критерии не являются равнозначными, то есть каждый из этих критериев имеет свой вес в задаче управления рисками ИБ. Поэтому необходимо определить вес каждого критерия по отношению к другим.

После этого, в зависимости от критериев необходимо выбрать контрмеру из альтернатив для каждой из угроз.

Для успешной реализации необходима формализация задачи, т.е. для начала определяются угрозы актуальные для данного ресурса. Для определения контрмер из альтернатив для каждой из угроз выделяем минимальный набор критериев, по которому можно эффективно оценить преимущества и недостатки контрмеры.

Пусть n – количество контрмер и m – количество угроз безопасности. Обозначим через i – индекс контрмеры, а через j – индекс угрозы информационной безопасности.

Требуется определить набор оптимальных контрмер для соответствующих угроз. Учитывая, что для каждой угрозы должна существовать хотя бы одна контрмера для снижения значения данной угрозы.

Обозначим через c_{ij} количественную оценку важности контрмеры i в соответствии с критерием для угрозы j (вектор глобальных приоритетов).

Для получения математической записи задачи о выборе эффективного набора контрмер введем $n \cdot m$ переменных следующим образом:

$$x_{ij} = \begin{cases} 1, & \text{если } i\text{-ая контрмера выбрана для } j\text{-ой угрозы} \\ 0, & \text{в противном случае} \end{cases}$$

Тогда задача принимает вид поиска максимума функции $\varphi(x)$.

$$\varphi(x) = \sum_{i=1}^n \sum_{j=1}^m c_{ij} x_{ij} \rightarrow \max$$

При следующих ограничениях на x :

$$\Omega = \begin{cases} x_{ij} \in [0,1], i = \overline{1, n}; j = \overline{1, m} \\ \sum_{i=1}^n x_{ij} = 1 \end{cases}$$

Таким образом, получается задача, которую можно решить методом анализа иерархий. Метод анализа иерархий – методологическая основа для решения задач выбора альтернатив посредством их многокритериального рейтингования. [3]

Иерархическое представление задачи представлено на рисунке 4.

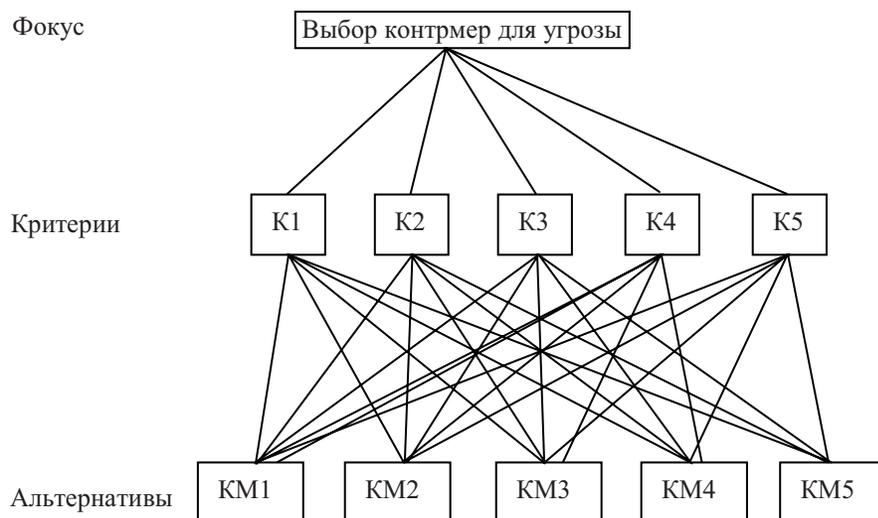


Рисунок 4 – Проблема выбора контрмер для угрозы: иерархическое представление

Алгоритм функционирования метода, основанного на методе анализа иерархий, представлен на рисунке 5 и состоит из следующих шагов.

- 1) Построение иерархии. Начиная с вершины (цель), через промежуточные уровни (критерии) к самому нижнему уровню (который является перечнем альтернатив).
- 2) Выбор текущей угрозы.
- 3) Построение множества матриц парных сравнений для каждого из нижних уровней — по одной матрице для каждого элемента, примыкающего сверху уровня. Этот элемент называют направляемым по отношению к элементу, находящемуся на нижнем уровне, так как элемент нижнего уровня влияет на расположенный выше элемент. Таким образом, получается квадратная матрица суждений. Парные сравнения проводятся в терминах доминирования одного из элементов над другим. Эти суждения затем выражаются в целых числах.
- 4) После проведения всех парных сравнений и ввода данных определяется индекс согласованности.
- 5) Затем, в зависимости от размерности матрицы определяется случайная согласованность.
- 6) Используя отклонение A_{\max} от n , проверяется индекс согласованности.

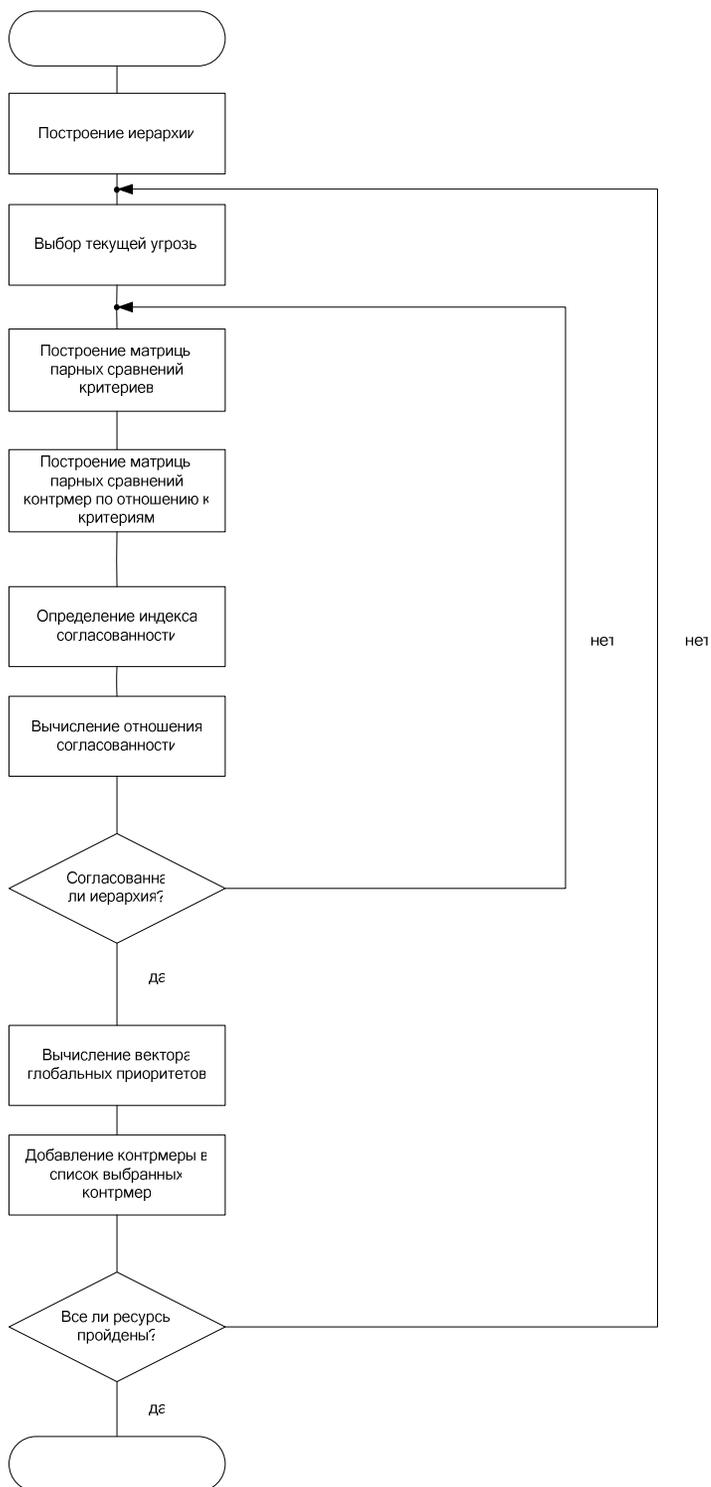


Рисунок 5 – Блок-схема решения задачи выбора контрмер угроз ИБ методом анализа иерархии

- 7) С помощью отношения индекса согласованности и случайной согласованности, необходимо получить отношение согласованности. Данная величина должна быть порядка 10% или менее, что бы быть приемлемой. В некоторых случаях можно допустить 20%, но не более. Если отношение согласованности выходит из этих пределов, то необходимо пересмотреть результаты матриц парных сравнений, т.е. вернуться к 3 пункту.
- 8) Затем необходимо вычислить вектор глобальных приоритетов. Данный вектор содержит значения функции φ , соответствующие каждой контрамере для данной угрозы. Контрмера с индексом, соответствующим индексу максимального значения функции в векторе глобальных приоритетов, считается оптимальной для данной угрозы и ее необходимо добавить в список выбранных контрамер.
- 9) Если пройдены не все угрозы, то необходимо осуществить переход к пункту 2 и повторить пункты 3-8 для следующей угрозы. Если же все угрозы пройдены, то полученный список контрамер будет являться решением задачи.

После решения данной задачи для каждой из угроз будет получен список контрамер, которые являются наилучшими, с точки зрения подобранных критериев и рассмотренных альтернатив.

Заключение

Таким образом, была создана система оценки рисков информационной безопасности на основе продукционно-фреймовой модели представления знаний. Были сформированы критерии выбора контрамер угроз информационной безопасности и предложен метод выбора наиболее эффективного набора контрамер с точки зрения предложенных критериев.

Список литературы

- [1] Гаврилова Т.А., Хорошевский В.Ф Базы знаний интеллектуальных систем. – СПб.: Питер, 2001.
- [2] Minsky M. Minsky's frame system theory. / Theoretical Issues In Natural Language Processing. Representing knowledge: part III. Frames. с: 104 – 116, 1975 г
- [3] Саати Т., Кернс К. - Аналитическое планирование. Организация систем - 1991 г.

НЕЙРОНЕЧЁТКАЯ МОДЕЛЬ ФОРМИРОВАНИЯ БАЗ ЗНАНИЙ ЭКСПЕРТНЫХ СИСТЕМ С ГЕНЕТИЧЕСКИМ АЛГОРИТМОМ ОБУЧЕНИЯ

А.С. Катасёв, Ч.Ф. Ахатова

Казанский государственный технический университет им. А.Н. Туполева
420111, Казань, ул. К.Маркса, 10, Россия, Республика Татарстан
Kat_726@mail.ru
тел: +7 (843) 231-00-56

Ключевые слова: нечеткая нейронная сеть, генетический алгоритм, база знаний

Abstract

This article describes the fuzzy neural network model for expert systems knowledge bases formation. The genetic algorithm of the model tuning is described. The results of its approbation and research on the medical diagnostics data are introduced.

Введение

В настоящее время во многих областях человеческой деятельности для решения важных практических задач используются подходы на основе систем искусственного интеллекта. В частности, при решении неформализованных или плохо формализованных задач, таких как обучение, диагностика, прогнозирование, управление и измерения в сложных технических системах, успешно применяются экспертные системы.

Для данного класса интеллектуальных информационных систем характерно то, что они способны моделировать процесс рассуждения эксперта при принятии им решения и объяснять, почему был получен тот или иной результат. Это достигается путем реализации процедуры логического вывода на формализованных знаниях о предметной области, о тех процессах, которые в ней протекают, о закономерностях, которым подчиняются данные процессы.

Основными проблемами при разработке экспертных систем являются проблемы получения и формализации знаний, составляющих основу ее базы знаний. Первая проблема связана с большими трудностями, возникающими при взаимодействии инженера по знаниям с источником знаний – экспертом. Большинство знаний эксперта основано на его жизненном опыте и интуиции. Часто во время решения конкретных задач, эксперт, не задумываясь, быстро находит правильное решение. Это происходит на бессознательном уровне, на уровне его интуиции. Попытки инженера по знаниям проанализировать и понять процесс рассуждения эксперта могут оказаться безуспешными.

Проблема формализации знаний эксперта связана с необходимостью выбора конкретной модели представления знаний и их «перевода» в рамки выбранной формальной модели. При этом имеющиеся знания проходят через призму тех ограничений, которые присутствуют во всех без исключения моделях, что может привести к потере части полезных знаний, необходимых для качественного решения задачи.

Сформированная таким образом база знаний экспертной системы будет неполной, недостоверной, противоречивой, что может привести к невозможности ее практического применения. Одним из эффективных подходов к решению данных проблем является, с одной стороны, разработка и использование методов автоматического формирования баз знаний экспертных систем на основе интеллектуального анализа данных, а, с другой, выбор гибких моделей представления знаний, адекватных рассматриваемой предметной области. Решению данных задач посвящено настоящее исследование.

1 Постановка задачи исследования

Во введении была обоснована актуальность применения подхода к формированию баз знаний экспертных систем без участия или с минимальным участием человека-эксперта на основе использования методов интеллектуального анализа данных и выбора адекватных моделей представления знаний. Решение данных задач может быть найдено путем разработки и обучения нечеткой нейронной сети.

Как известно, нечеткие нейронные сети представляют собой гибридные модели, сочетающие в себе главные достоинства нейронных сетей (возможность адаптивного самообучения) и нечетких систем (хорошая интерпретируемость получаемого с их помощью результата). При этом любая нечеткая нейронная сеть представляет собой аналог той системы нечеткого логического вывода, в рамках которой она, в результате обучения, формирует систему правил принятия решений. Отличие между ними состоит лишь в том, что система нечеткого логического вывода способна принимать решения на основе готовой, например, сформированной при помощи эксперта, базы знаний, в то время как нечеткая нейронная сеть сама формирует систему нечетких правил в результате обучения на данных из обучающей выборки.

Среди множества моделей представления экспертных знаний в настоящее время большое распространение получили правила продукций из-за их гибкости и простоты описания закономерностей предметной области. К тому же продукции более всего близки человеку, когда он рассуждает, производит умозаключения, принимает решения.

Однако, большинство задач в различных областях человеческой деятельности приходится решать в условиях нечеткости, неполноты и недостоверности исходных данных. В рамках использования обычных продукционных правил подобные задачи решать невозможно. Для этого необходимо применять нечеткие продукции и системы нечеткого логического вывода. Это позволяет, хотя и приближенно, находить решение задачи.

Таким образом, задача настоящего исследования сводится к выбору гибкой нечетко-продукционной модели представления знаний и разработке нечеткой нейронной сети, позволяющей в результате обучения формировать базу знаний экспертной системы правилами, представленными в форме нечетких продукций.

2 Нечетко-продукционная модель представления знаний

В [1] на примере задачи планирования геолого-технических мероприятий на нефтяных месторождениях показано, что зависимость между значениями входных параметров (условиями) и назначаемым результатом может быть эффективно описана экспертом в виде следующей нечетко-продукционной модели представления знаний:

$$(1) \quad R_j : (P^j, \tilde{A}^j, w^j, CF^j, T^j),$$

где $P^j = \{P_i^j\}$ – множество параметров, на которые накладываются ограничения в условиях правила R_j ;

$\tilde{A}^j = \{\tilde{A}_i^j\}$ – множество нечетких ограничений на параметры множества P^j – нечетких условий antecedента правила R_j ;

$w^j = \{w_i^j\}$ – веса нечетких ограничений \tilde{A}^j на параметры P^j ;

CF^j – степень достоверности сформированного правила R_j – степень уверенности эксперта в его универсальности (*CERTAINTY FACTOR*);

$T_j \in T$ – технология, рекомендуемая к применению.

Модель (1) может быть представлена в следующем виде:

		Правило R_j
«ЕСЛИ	P_1^j есть $\tilde{A}_1^j (w_1^j)$	И
	P_2^j есть $\tilde{A}_2^j (w_2^j)$	И
	...	И
	P_n^j есть $\tilde{A}_n^j (w_n^j)$	
ТО	Возможно применение технологии T_j » [CF^j].	

Веса w_i^j нечётких ограничений \tilde{A}_i^j антецедента правила R_j определяют важность ограничений \tilde{A}_i^j в правиле.

В отличие от обычных нечетких продукций, в правилах модели (1) присутствуют такие параметры, как вес нечеткого ограничения w_i^j и достоверность сформированного правила CF^j . Это придает таким правилам большую гибкость при решении реальных практических задач, позволяя учитывать важность отдельных условий и достоверность правила в целом.

Следует отметить [1], что существует система нечеткого логического вывода на правилах, представленных в виде модели (1), принцип работы которой заключается в анализе входных данных, формировании интегральных взвешенных оценок выполнимости каждого правила и выборе оптимального решения, соответствующего правилу с максимальной степенью достоверности принимаемого решения. Как аналог данной системы логического вывода разработана нейронечеткая модель формирования баз знаний экспертных систем.

3 Нейронечеткая модель формирования баз знаний экспертных систем

Разработанная авторами модель формирования баз знаний экспертных система в виде нечеткой нейронной сети реализует алгоритм нечеткого логического вывода на модели (1), что определяет число слоев нейронной сети и их функциональность. Кроме того, структура сети зависит от количества входных, выходных параметров, а также числа градаций входных нейронов. Данные параметры влияют на количество нейронов в слоях нейронной сети.

В качестве примера рассмотрим нечеткую нейронную сеть, имеющую три выхода и два входа с тремя нечеткими градациями (см. рис. 1).

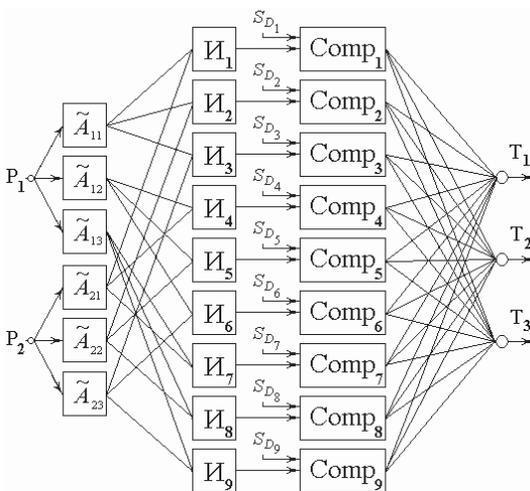


Рисунок 1 – Пример структуры нечеткой нейронной сети

В нулевом слое нейронной сети содержится два Р-нейрона, выполняющих функцию распределения входных сигналов по нейронам первого слоя.

В первом слое сети содержится $2 \cdot 3 = 6$ А-нейронов, которые моделируют нечеткие градации в виде «Р есть \tilde{A}_i ».

Второй слой содержит $3^2 = 9$ И-нейронов, задающих нечеткие продукции в виде: «ЕСЛИ P_1 есть \tilde{A}_1 И P_2 есть \tilde{A}_2 ТО T_j ».

Третий слой нейронной сети состоит из девяти Comp-нейронов, на выходе которых рассчитываются комплексные оценки срабатывания правил. Веса связи нейронов с выходами нейронной сети определяют степень достоверности CF^j сформированного пра-

вила. Произведения выходов нейронов на их веса образуют общий коэффициент достоверности решения.

В четвертом слое содержатся три Т-нейрона, на выходе которых вычисляются взвешенные нормированные оценки достоверности решения.

Для получения системы правил и формирования базы знаний экспертной системы необходимо произвести обучение нечеткой нейронной сети, в результате которого производится настройка параметров \tilde{A}_i ее функций принадлежности.

4 Генетический алгоритм обучения нечеткой нейронной сети

Задача обучения нечеткой нейронной сети относится к классу оптимизационных задач, связанных с нахождением минимума ошибки выхода нейронной сети. В [2] для ее решения был использован алгоритм обратного распространения ошибки, основанный на методе градиентного спуска. Основными недостатками такого подхода являются длительность процесса обучения, а также сложность нахождения глобального минимума.

С другой стороны, для решения задач глобальной оптимизации большое распространение получили генетические алгоритмы, эффективность которых основана на идее случайного направленного поиска путем моделирования биологических процессов естественной эволюции.

В данной работе для повышения эффективности обучения нечеткой нейронной сети, решающей задачу формирования баз знаний экспертных систем, предлагается использование генетического алгоритма. Рассмотрим основные этапы его функционирования.

4.1 Кодирование и декодирование параметров задачи

Как было сказано выше, обучение нечеткой нейронной сети заключается в настройке параметров ее функций принадлежности, которые задаются некоторой параметризованной функцией формы (треугольной, трапециевидальной, гауссовой и др.) с вектором параметров \overline{u}_k .

Так, для треугольной функции принадлежности $\overline{u}_k = (u_{1k}, u_{2k}, u_{3k})$, где u_{1k} – левое основание, u_{2k} – центр (мода), u_{3k} – правое основание (см. рис. 2).

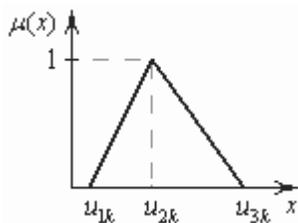


Рисунок 2 – Пример треугольной функции принадлежности

Настройка параметров u_{ik} функций принадлежности осуществляется в окрестности $\Delta = [u_{ik_{\min}}, u_{ik_{\max}}]$, ширина которой определяется индивидуально для каждого параметра функции принадлежности и может составлять до 30 % ширины ее основания.

Разобьем данную окрестность на 2^n интервалов, где n – целое число, например, 10. Тогда параметр u_{ik} сможет принимать одно из 2^{10} значений, которые будут кодироваться хромосомой $\{u_{ik}\}$ длиной 10 бит. При этом хромосома, состоящая из нулей, будет соответствовать значению параметра $u_{ik_{\min}}$, а состоя-

щая из единиц – значению $u_{ik_{\max}}$. Таким образом, общее число хромосом для данного типа функции принадлежности равно $3 \cdot 2^n$. Для нечеткой нейронной сети с одним выходным нейроном общее количество хромосом в популяции равно $3 \cdot N \cdot 2^n$, где N – число А-нейронов сети.

Когда параметры соответствующих функций принадлежности закодированы в виде хромосомного набора, необходимо применить генетические операторы к хромосомам с целью минимизации среднеквадратичной ошибки выхода нечеткой нейронной сети $E = \frac{1}{2}(\Delta y)^2$, где

$\Delta u = y - t$ – абсолютная ошибка выходного нейрона, а t – требуемое значение выхода в соответствии с информацией из обучающей выборки.

Таким образом, ошибка E выступает в роли функции приспособленности для данного генетического алгоритма. При нахождении значения функции ошибки на очередном этапе работы генетического алгоритма необходимо произвести процедуру декодирования параметров соответствующих функций принадлежности, т.е. перейти от хромосомного набора к десятичным значениям данных параметров. Операция декодирования является инверсной по отношению к операции кодирования хромосом.

4.2 Создание начальной популяции хромосом

Введем вспомогательные операции над хромосомами:

- *композиция хромосом* – объединение текущего хромосомного набора $\{u_{11}\}, \{u_{21}\}, \{u_{31}\}, \dots, \{u_{1k}\}, \{u_{2k}\}, \{u_{3k}\}, \dots, \{u_{1N}\}, \{u_{2N}\}, \{u_{3N}\}$ соответствующих функций принадлежности в единую хромосому $\{\bar{U}\} = \bigcup_{k=1}^N \{\bar{u}_k\}$;
- *декомпозиция хромосом* – разбиение хромосомы $\{\bar{U}\}$ на подхромосомы $\{u_{11}\}, \{u_{21}\}, \{u_{31}\}, \dots, \{u_{1k}\}, \{u_{2k}\}, \{u_{3k}\}, \dots, \{u_{1N}\}, \{u_{2N}\}, \{u_{3N}\}$.

Операция декомпозиции хромосом необходима для локальной оптимизации параметров соответствующих функций принадлежности, а операция композиции – для оценки значения функции приспособленности всего хромосомного набора.

Для создания начальной популяции случайным образом выбираются 100 хромосом ch_i из общего хромосомного набора $\{U\}$, мощность которого пропорциональна числу $2^{N \cdot n}$, после чего производится оценка приспособленности данных хромосом.

4.3 Оценка приспособленности хромосом в популяции

Для оценки приспособленности хромосом ch_i в популяции производится расчет значений функции приспособленности E , на вход которой подаются декодированные значения хромосом, полученных в результате операции декомпозиции. При этом каждой хромосоме ch_i из общего хромосомного набора ставится в соответствие величина ее приспособленности $E(ch_i)$.

4.4 Селекция хромосом

Селекция хромосом заключается в выборе по рассчитанным значениям функции приспособленности тех хромосом, которые будут участвовать в создании потомков для следующей популяции. Такой выбор производится согласно принципу естественного отбора, по которому наибольшие шансы на участие в создании новых особей имеют хромосомы с наименьшими значениями функции приспособленности (для минимизации ошибки выхода нейронной сети).

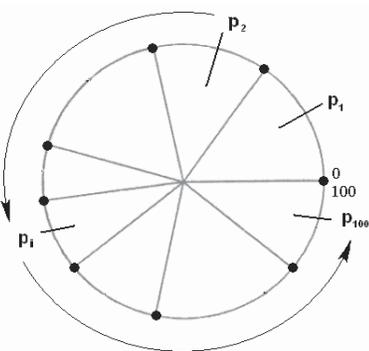


Рисунок 3 – Колесо рулетки для селекции хромосом

Для реализации оператора селекции используется метод рулетки [3]. Для каждой из 100 хромосом текущей популяции формируются секторы колеса рулетки, площадь которых пропорциональна значению инверсных вероятностей селекции данных хромосом

$$p_i = \frac{1 - E(ch_i)}{\sum_i (1 - E(ch_i))} \quad (\text{см. рис. 3}).$$

Розыгрыш с помощью колеса рулетки сводится к случайному выбору числа из интервала $[0, 100]$, указывающего на соответствующий сектор на колесе, которому соответствует конкретная хромосома. Таким способом из текущей популяции выбираются два кандидата

та-родителя, к которому в дальнейшем после операции декомпозиции применяются генетические операторы.

4.5 Применение генетических операторов

Основными операторами генетического алгоритма являются операторы скрещивания и мутации, которые применяются к хромосомному набору параметров конкретных функций принадлежности после выполнения операции декомпозиции. Рассмотрим работу операторов.

Скрещивание родительских хромосом производится в соответствии с вероятностью $p_c \in [0.5, 1]$. В таблице 1 представлен пример работы данного оператора.

Таблица 1 – Пример работы оператора скрещивания

Хромосомы \ Позиции генов	1	...	i	i+1	...	n
X	α_1	...	α_i	α_{i+1}	...	α_n
Y	β_1	...	β_i	β_{i+1}	...	β_n
			↓			
Z	α_1	...	α_i	β_{i+1}	...	β_n
Q	β_1	...	β_i	α_{i+1}	...	α_n

Здесь X и Y – два выбранных решения (родители), Z, Q – новые решения (потомки).

Предположим, что вероятность мутации равна 0,01. Сам процесс мутации заключается в инверсии одного из битов строки дочерних хромосом, выбираемого случайно по равномерному закону. В результате выполнения генетических операторов исходная популяция за счет порожденных потомков увеличилась на 2.

4.6 Формирование новой популяции

Для сокращения популяции до исходного числа особей применяется оператор редукции. Для этого производится операция композиции хромосом и оценивается их приспособленность, после чего из полученной популяции исключаются две хромосомы, чье значение целевой функции максимально.

После останова работы генетического алгоритма получим обученную нечеткую нейронную сеть, аппроксимирующую с заданной точностью данные из обучающей выборки и формирующую базу знаний, состоящую из системы нечетко-продукционных правил.

5 Апробация нечеткой нейронной сети

В качестве примера практического использования нечеткой нейронной сети рассмотрим процесс анализа медицинских данных. Пусть необходимо выявить зависимость между возрастными характеристиками пациента и риском возникновения у него признаков заболевания.

В качестве входов нейронной сети примем «возраст пациента», «стаж заболевания» и «стаж нагрузок», а выходов – «гипертрофия желтой связки», «протрузия диска» и «вакуум-феномен диска». В результате обучения сети получим правила следующего вида:

ЕСЛИ «Возраст пациента»=«средний» ($w=0,48$) И «Стаж заболевания»=«большой» ($w=0,45$) И «Стаж нагрузок»=«большой» ($w=0,52$) ТО «Возможна протрузия диска» ($CF=0,47$).

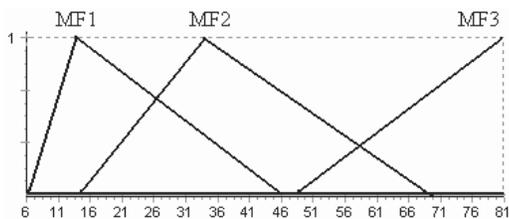


Рисунок 4 – Пример сформированных функций принадлежности

На рисунке 4 показан пример сформированных функций принадлежности.

Здесь по оси абсцисс указан возраст пациента в годах, по оси ординат – степень принад-

лежности значений возраста к нечетким градациям «молодой» (MF1), «средний» (MF2) и «пожилой» (MF3).

Особенности формирования базы знаний с использованием нейронечеткой модели изучались при разработке экспертной системы диагностики клинических проявлений остеохондроза поясничного отдела позвоночника [4]. Для этого были использованы данные обследования 230 женщин в возрасте от 15 до 92 лет и 180 мужчин в возрасте от 16 до 81 года с различными синдромами поясничного остеохондроза на стационарном этапе обострения и в стадии начинающейся ремиссии. Контрольную группу составили 20 женщин в возрасте от 20 до 70 лет и 20 мужчин в возрасте от 17 до 73 лет, никогда не страдавших поясничными болями.

Для получения системы правил и их параметров измерено более 500000 значений признаков течения поясничного остеохондроза по 822 признакам. В результате обучения сети на более чем 100 обучающих выборках получено более 500 правил, значимость которых подтверждена экспертами – специалистами-вертеброневрологами высокой квалификации.

6 Численно-параметрические исследования

Для оценки эффективности генетического алгоритма обучения нечеткой нейронной сети проведен ряд численно-параметрических исследований, основанных на его сравнении с традиционным алгоритмом обратного распространения ошибки по критериям «скорость обучения», «точность аппроксимации» и «устойчивость работы на комплектных и некомплектных обучающих выборках». Кроме того, оценивалась средняя скорость сходимости алгоритмов обучения. Во всех проводимых экспериментах генетический алгоритм показал лучшие результаты при обучении нейронной сети.

Заключение

Практическое значение предложенного подхода заключается в возможности эффективно формирования баз знаний экспертных систем в различных областях человеческой деятельности без участия или с минимальным участием эксперта. Обученная нейронная сеть является системой нечеткого логического вывода. Она может использоваться как инструмент эксперта в составе мягкой экспертной системы и позволяет заменить (или дополнить) эвристический процесс построения базы знаний на процесс автоматизированного ее формирования, извлекая закономерности из статистических выборок данных.

Апробация нечеткой нейронной сети на данных медицинской диагностики показала, что получаемые при этом правила с высокой степенью точности аппроксимируют экспериментальные данные и в большинстве случаев позволяют принимать правильные решения при постановке медицинских диагнозов.

Таким образом, предложенная в данной работе нечеткая нейронная сеть с генетическим алгоритмом обучения вносит значимый вклад в развитие систем интеллектуального анализа данных и формирования баз знаний экспертных систем в сферах диагностики, управления и принятия решений.

Список литературы

- [1] Глова В.И., Аникин И.В., Шагиахметов М.Р. Методы многокритериального принятия решений в условиях неопределённости в задачах нефтедобычи. Препринт 04П2. – Казань: / Изд-во Казан. гос. техн. ун-та, 2004. – 31 с.
- [2] Катасёв А.С. Нейронечеткая модель и программный комплекс формирования баз знаний экспертных систем // Автореф. дисс. на соиск. уч. степ. к-та техн. наук. Казань, 2006. – 20 с.
- [3] Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И.Д. Рудинского. – М.: Горячая линия – Телеком, 2006. – 452 с.: ил.
- [4] Глова В.И., Аникин И.В., Катасёв А.С., Подольская М.А. Формирование базы знаний медицинской диагностической экспертной системы на основе нечеткой нейронной сети // Исследования по информатике. Вып. 12. – Казань: Отечество, 2007. – С. 31-46.

ВЕРИФИКАЦИЯ И СИНТЕЗ ЭФФЕКТИВНЫХ ПРОГРАММ СТАНДАРТНЫХ ФУНКЦИЙ FLOOR, ISQRT И ILOG2 В ТЕХНОЛОГИИ ПРЕДИКАТНОГО ПРОГРАММИРОВАНИЯ

В.И. Шелехов

Институт систем информатики им. А.П.Ершова СО РАН
630090, Новосибирск, пр. акад. Лаврентьева, 6, Россия
vshel@iis.nsk.su
тел: (383) 330-27-21, факс: (383) 332-34-94

Ключевые слова: предикатное программирование, дедуктивная верификация, программный синтез, тотальная корректность программы

Abstract

Methods of deductive verification and program synthesis are illustrated on the examples of development of effective programs for standard functions under predicate programming technology. The proof of program correctness conditions is supported by the PVS Verification System.

Введение

Применение программного синтеза или дедуктивной верификации программ становится настоящим в областях критического программирования с высокой ценой ошибки. Автоматическое доказательство условий корректности программы (или автоматический синтез программы) дает стопроцентную гарантию правильности программы относительно ее спецификации, что недостижимо применением самых совершенных методов тестирования программ.

Методы программного синтеза и дедуктивной верификации иллюстрируются на примерах разработки быстрых программ для стандартных функций: **floor** – целой части плавающего числа, представленного в бинарном формате в соответствии со стандартом IEEE 754-2008, **isqrt** – целочисленного квадратного корня и **ilog2** – целочисленного двоичного логарифма. Программы указанных стандартных функций на языке C++ используются в известной отечественной системе спутниковой навигации “Навител Навигатор” для автомобилей.

Дедуктивная верификация и программный синтез в предикатном программировании определены в разделе 1. Синтез программ целочисленного квадратного корня и целой части плавающего числа описывается в разделах 2 и 3. Обзор смежных работ представлен в разделе 4. В заключении суммируется опыт разработки программ указанных стандартных функций в режимах верификации и синтеза. Условия корректности программы были доказаны с помощью известной системы автоматического доказательства PVS [4].

1 Дедуктивная верификация и программный синтез

Язык предикатного программирования P (Predicate programming language) [1] находится на границе между функциональными и логическими языками. Язык P, его логическая и операционная семантика, построены расширением языков $ССР = P_0 \subset P_1 \subset P_2 \subset P_3 \subset P_4 = P$, где ССР – типизированное исчисление вычислимых предикатов [2]. Базис исчисления содержит: оператор суперпозиции $V(x: z); C(z: y)$; параллельный оператор $V(x: y) \parallel C(x: z)$ и условный оператор **if** (b) $V(x: y)$ **else** $C(x: y)$, где V и C – имена предикатов, x, y и z – различающиеся наборы переменных, причем набор x может быть пустым, b – переменная логического типа. Логическая семантика LS для операторов определяется следующим образом:

$$LS(B(x: z); C(z: y)) \equiv \exists z. (B(x, z) \& C(z, y))$$

$$LS(B(x: y) \parallel C(x: z)) \equiv B(x, y) \& C(x, z)$$

$$LS(\text{if } (b) B(x: y) \text{ else } C(x: y)) \equiv (b \Rightarrow B(x, y)) \& (\neg b \Rightarrow C(x, y))$$

Предикатная программа состоит из набора определений предикатов вида:

$$(1) A(x: y) \equiv \text{pre } P(x) \{ S \} \text{ post } Q(x, y),$$

где A – имя определяемого предиката, x – аргументы, а y – результаты предиката, S – оператор, $P(x)$ – предусловие, $Q(x, y)$ – постусловие. Предусловие должно быть истинно перед исполнением оператора S , а постусловие – после исполнения. Однозначность программы S и тотальность спецификации $[P(x), Q(x, y)]$ определяются, соответственно, формулами:

$$P(x) \& LS(S)(x, y_1) \& LS(S)(x, y_2) \Rightarrow y_1 = y_2,$$

$$P(x) \Rightarrow \exists y. Q(x, y).$$

Тотальная корректность определения предиката (1) со спецификацией $[P(x), Q(x, y)]$ представляется формулой:

$$(2) P(x) \Rightarrow [LS(S)(x, y) \Rightarrow Q(x, y)] \& \exists y. LS(S)(x, y).$$

Здесь $LS(S)$ – предикат, являющийся логическим эквивалентом оператора S в соответствии с логической семантикой. Доказательство тотальной корректности для однозначных программ с тотальной спецификацией можно проводить по более простой формуле:

$$(3) P(x) \& Q(x, y) \Rightarrow LS(S)(x, y).$$

На базе формул (2, 3) построены две серии (R и L) правил доказательства корректности операторов языка P [3].

Задачи верификации и синтеза рассмотрим на примере оператора суперпозиции в следующем определении предиката:

$$(4) A(x: y) \equiv \text{pre } P(x) \{ B(x: z); C(z: y) \} \text{ post } Q(x, y).$$

Предположим, что операторы B и C корректны относительно спецификаций $[P_B(x), Q_B(x, z)]$ и $[P_C(z), Q_C(z, y)]$, соответственно. Пусть спецификация $[P(x), Q(x, y)]$ тотальна. Не трудно доказать, что корректность предиката A гарантируется в случае истинности правил¹:

Правило LS1. $P(x) \vdash P_B(x)$

Правило LS2. $P(x) \& Q(x, y) \& Q_B(x, z) \vdash P_C(z) \& Q_C(z, y)$

Задача верификации. Пусть имеется определение (4). Операторы B и C корректны относительно своих спецификаций. Для доказательства корректности определения (4) требуется доказать тотальность спецификации $[P(x), Q(x, y)]$ и истинность правил LS1 и LS2.

Задача синтеза. Требуется построить программу для предиката A , представленного спецификацией $[P(x), Q(x, y)]$. Допустим, доказана тотальность спецификации. Пусть для некоторых предикатов $P_B(x)$, $Q_B(x, z)$, $P_C(z)$ и $Q_C(z, y)$ удалось доказать истинность правил LS1 и LS2. Тогда для предиката A синтезируется определение (4) с включением в программу двух новых предикатов: B со спецификацией $[P_B(x), Q_B(x, z)]$ и C со спецификацией $[P_C(z), Q_C(z, y)]$. Дальнейшей целью является синтез определений для B и C .

Верификация и синтез зеркальны. Если для предикатной программы сгенерируем условия корректности, а затем, по этим формулам проведем синтез, то получим ту же программу. Наоборот, если на основе спецификации программы и набора формул синтезируем программу, а затем для программы сгенерируем условия корректности, то получим эквивалентную систему формул. Тем не менее, синтезируемая программа отличается от запрограммированной, хотя и эквивалентна ей. Причина в том, что в процессе программирования мы невольно совершаем оптимизацию программы. Оптимизации программы меняют ее логику, обычно в сторону ускорения. Поэтому верификация оказывается более сложной и трудоемкой, чем синтез.

¹ Истинность правил определяется как истинность формул, получающихся заменой \vdash на \Rightarrow .

2 Синтез программ вычисления целочисленного квадратного корня

Целая часть квадратного корня есть функция: $m = \text{isqrt}(x) = \text{floor}(\sqrt{x})$, где $z = \text{floor}(t)$ – целая часть вещественного аргумента t со спецификацией $t \leq z < t+1$. Спецификацией для $t = \sqrt{x}$ является формула $t^2 = x$. Таким образом, функция isqrt имеет спецификацию:

formula $\text{Isqrt}(\text{nat } x, m) = m^2 \leq x \ \& \ x < (m + 1)^2$;
isqrt(**nat** x : **nat** m) **post** $\text{Isqrt}(x, m)$.

Синтез программы isqrt управляется обычным процессом построения программы по технологии предикатного программирования. При этом программа извлекается из доказательства формулы корректности вида (3): $\text{Isqrt}(x, m) \Rightarrow \text{LS}(S(x; m))$ для некоторого оператора $S(x; m)$. Кроме того, необходимо доказать тотальность спецификации, т.е. лемму:

Isqrt_total : **lemma exists nat** m . $\text{Isqrt}(x, m)$.

Сначала синтезируем простейшую программу, определяющую результат m последовательным перебором, начиная с нуля. Для построения программы применяется метод обобщения исходной задачи isqrt . Рассмотрим задачу sq0 , определяемую спецификацией:

formula $P_{\text{sq0}}(\text{nat } x, k) = k^2 \leq x$;
 $\text{sq0}(\text{nat } x, k; m)$ **pre** $P_{\text{sq0}}(x, k)$ **post** $\text{Isqrt}(x, m)$.

Задача $\text{isqrt}(x; m)$ сводится к $\text{sq0}(x, 0; m)$, что определяет программу:

(5) $\text{isqrt}(\text{nat } x : \text{nat } m) \{ \text{sq0}(x, 0; m) \}$ **post** $\text{Isqrt}(x, m)$.

Корректность программы (5) доказывается по правилу FB1:

Правило FB1. $R(z, y) \vdash P(z) \ \& \ Q(z, y)$.

Истинность правила гарантирует истинность формулы $R(z, y) \Rightarrow \text{LS}(A(z; y))$ при условии, что предикат $A(z; y)$ является корректным относительно однозначной спецификации $[P(z), Q(z, y)]$. Таким образом, требуется доказать лемму:

(6) Isqrt_fb1 : **lemma** $\text{Isqrt}(x, m) \Rightarrow P_{\text{sq0}}(x, 0) \ \& \ \text{Isqrt}(x, m)$.

Разработчик программы должен записать эту лемму и доказать ее, а система автоматически, используя правило FB1 в качестве шаблона, должна синтезировать программу (5) по формуле (6). Следующей целью является синтез программы sq0 из доказательства формулы $P_{\text{sq0}}(x, k) \ \& \ \text{Isqrt}(x, m) \Rightarrow \text{LS}(S1(x, k; m))$ для некоторого оператора $S1(x, k; m)$. Отметим, что предусловие $P_{\text{sq0}}(x, k)$ обеспечивает истинность первого конъюнкта в $\text{Isqrt}(x, m)$, а в случае истинности условия $x < (k + 1)^2$ получим решение задачи sq0 : $m = k$. Задача sq0 делится на две: для истинного значения условия $x < (k + 1)^2$ и для ложного, что определяет поиск решения $S1(x, k; m)$ в виде оператора **if** $(x < (k + 1)^2) m = k$ **else** $B(x, k; m)$ для некоторого $B(x, k; m)$. Корректность условного оператора доказывается правилом:

Правило FC. $R(z, y) \vdash (C \Rightarrow \text{LS}(A(z; y))) \ \& \ (\neg C \Rightarrow \text{LS}(B(z; y)))$

Истинность правила гарантирует истинность $R(z, y) \Rightarrow \text{LS}(\text{if } (C) A(z; y) \text{ else } B(z; y))$. Заметим, что ложное условие $x < (k + 1)^2$ эквивалентно $P_{\text{sq0}}(x, k+1)$, а поскольку истинно $\text{Isqrt}(x, m)$, то решение $B(x, k; m)$ следует определить в виде рекурсивного вызова $\text{sq0}(x, k + 1; m)$. Доказательство корректности рекурсивного вызова предиката реализуется правилом FB3, аналогичным FB1, с дополнительным условием: аргументы рекурсивного вызова должны быть по некоторой мере строго меньше аргументов определяемого рекурсивного предиката, где мера – натуральная функция от значений аргументов. В нашем случае мера определяется формулой:

formula $e(\text{nat } x, k; \text{nat}) = (x < (k + 1)^2) ? 0 : x - k^2$;

Таким образом, корректность условного оператора с рекурсивным вызовом определяется леммой:

sq0_rec : **lemma** $P_{\text{sq0}}(x, k) \ \& \ \text{Isqrt}(x, m) \Rightarrow$
 $(x < (k + 1)^2) ? m = k : e(x, k+1) < e(x, k) \ \& \ P_{\text{sq0}}(x, k+1) \ \& \ \text{Isqrt}(x, m)$

По данной формуле, используя в качестве шаблонов правила FC и FB3, система должна автоматически синтезировать рекурсивную программу:

```
sq0(nat x, k : nat m) pre P_sq0(x, k)
{ if (x < (k + 1)^2) m = k else sq0(x, k + 1 : m) } post Isqrt(x, m);
```

Заметным недостатком алгоритма sq0 является вычисление квадрата в выражении $(k + 1)^2$. Поскольку $(k + 1)^2 = k^2 + 2 * k + 1$, то можно использовать значение k^2 , вычисленное на предыдущем шаге введением дополнительного параметра $n = k^2$. Аналогичным образом, введя параметр $d = 2 * k + 1$, можно заменить умножение сложением. Далее, поскольку $x < (k + 1)^2$ эквивалентно $x - k^2 < 2 * k + 1$, то вместо x и n можно перейти к одному параметру $y = x - k^2$. Таким образом, переходим к более общей задаче, определяемой спецификацией предиката sq1:

```
formula P_sq1(nat x, y, k, d) = k^2 <= x & d = 2 * k + 1 & y = x - k^2 ;
sq1(nat x, y, k, d: nat m) pre P_sq1(x, y, k, d) post Isqrt(x, m);
```

Результатом синтеза, проводимого по той же схеме, является программа:

```
isqrt(nat x : nat m) { sq1(x, x, 0, 1: m) } post Isqrt(x, m);
sq1(nat x, y, k, d: nat m) pre P_sq1(x, y, k, d)
{ if (y < d) m = k else sq1(x, y - d, k + 1, d + 2: m) } post Isqrt(x, m);
```

В программе sq1 число шагов равно значению квадратного корня. Рассмотрим алгоритм с логарифмической оценкой для числа шагов. На каждом шаге алгоритма находится очередная цифра в двоичном представлении значения квадратного корня. Предположим, что аргумент n ограничен значением 2^{2^p} . Представим исходную задачу спецификацией:

```
formula P_sqp(nat n, p) = p > 0 & n < 2^(2*p);
isqrt1(nat n, p: nat s) pre P_sqp(n, p) post Isqrt(n, s);
```

Результат s имеет не более p двоичных цифр. Значение старшей цифры равно 0, если $n < 2^{2*(p-1)}$ и 1 в противном случае. Допустим, для результата s найдено очередное приближение q в виде старших двоичных цифр с номерами от p до k . Тогда должно выполняться условие $q^2 <= n < (q + 2^k)^2$. Представим спецификацию обобщенной задачи:

```
formula P_sq2(nat n, p, k, q) = P_sqp(n, p) & k <= p & q^2 <= n & n < (q + 2^k)^2 ;
sq2(nat n, p, k, q : nat s) pre P_sq2(n, p, k, q) post Isqrt(n, s);
```

Справедлива следующая лемма:

```
Isqrt1_fb1: lemma P_sqp(n, p) & Isqrt(n, s) => P_sq2(n, p, p, 0) & Isqrt(n, s)
```

По данной формуле, используя правило FB1 в качестве шаблона, система синтезирует:

```
isqrt1(nat n, p: nat s) pre P_sqp(n, p) { sq2(n, p, p, 0: s) } post Isqrt(n, s);
```

Заметим, что при $k = 0$ решением задачи sq2 будет $s = q$. Если в s цифра с номером $k-1$ есть 1, то следующим приближением для q будет $q + 2^{k-1}$. Если $n < (q + 2^{k-1})^2$, то цифра с номером $k-1$ есть 0, иначе 1. Данные утверждения суммируем в виде леммы:

```
formula e2(k: nat) = k;
```

```
sq2_rec: lemma P_sq2(n, p, k, q) & Isqrt(n, s) =>
(k = 0)? s = q :
```

```
(n < (q + 2^(k-1))^2)? e2(k-1) < e2(k) & P_sq2(n, p, k-1, q) & Isqrt(n, s) :
e2(k-1) < e2(k) & P_sq2(n, p, k-1, q + 2^(k-1)) & Isqrt(n, s)
```

Разработчик должен доказать лемму, а система, используя в качестве шаблонов правила FC и FB3, должна автоматически синтезировать программу:

```
sq2(nat n, p, q, k: nat s) pre P_sq2(n, p, k, q)
{ if (k = 0) s = q
else if ( n < (q + 2^(k-1))^2 ) sq2(n, p, k - 1, q : s)
else sq2(n, p, k - 1, q + 2^(k-1) : s)
} post Isqrt(n, s);
```

В выражении $q + 2^{k-1}$ можно заменить “+” побитовой операцией “or”. Чтобы обеспечить корректность такой замены, необходимо доказать лемму:

$$\text{or_eq_plus: lemma } k > 0 \ \& \ k < p \ \& \ \text{mod}(q, 2^k) = 0 \ \& \ q < 2^p \Rightarrow \\ (q \text{ or } 2^{(k-1)}) = q + 2^{(k-1)}$$

Необходимое условие $\text{mod}(q, 2^k) = 0$ проще обеспечить включением его в предусловие. Условие $n < (q + 2^{(k-1)})^2$ эквивалентно: $n - q^2 < 2^{((k-1)*2)} + q * 2^k$. Как и в sq1 обозначим $y = n - q^2$. Далее мы можем синтезировать программу с данными оптимизациями:

```

formula P_sq3(nat n, p, k, q, y) = P_sq2(n, p, k, q) & y = n - q^2 & mod(q, 2^k) = 0;
isqrt1(nat n, p: nat s) pre P_sq3(n, p, 0, n: s) post Isqrt(n, s);
sq3(nat n, p, k, q, y: nat s) pre P_sq3(n, p, k, q, y)
{ if (k = 0) s = q
  else { nat t = 2^((k-1)*2) + q * 2^k;
        if (y < t) sq3(n, p, k - 1, q, y: s) else sq3(n, p, k - 1, q or 2^(k-1), y - t : s)
      }
} post Isqrt(n, s);

```

Для получения эффективной программы для данной предикатной программы применяется последовательность трансформаций. На первом этапе реализуются склеивания переменных в sq3: $n \leftarrow y$; $s \leftarrow q$. На втором этапе хвостовая рекурсия заменяется циклом. На третьем этапе тело определения sq3 поставляется на место вызова в isqrt1 и реализуются упрощения. В итоге получим программу на императивном расширении языка P:

```

isqrt1(nat n, p: nat s) {
  s = 0;
  for (nat k = p; k = 0; k = k - 1) { nat t = 2^((k-1)*2) + s * 2^k;
                                     if (n >= t) { n = n - t; s = s or 2^(k-1); }
  }
}

```

Далее заменим умножение на степень двойки операций сдвига влево. Для конкретного p (например, p = 16) проведем развертку цикла, представив тело цикла макросом sq(k). Специализируем первый и последний шаг цикла. Получим окончательную программу:

```

isqrt1(nat n, p: nat s) {
  sq(k) = { t = 2^((k-1)*2) + (s << k); if (n >= t) { n = n - t; s = s or (1 << (k-1)) } };
  s = 0;
  if (n >= 2^((p-1)*2)) { n = n - 2^((p-1)*2); s = 2^(p-1) };
  sq(p - 1); ...; sq(2);
  if (n >= s * 2) s = s or 1;
}

```

По сравнению с программой, представленной на сайте:

<http://www.finesse.demon.co.uk/steven/sqrt.html>, параметр k смещен на единицу. Однако поскольку тело sq(k) подставляется открыто на место вызовов sq(p - 1), ..., sq(2), то обе программы идентичны.

3 Генерация программы вычисления целой части плавающего числа

Функцию $y = \text{floor}(x)$, вычисляющую наименьшее целое, не превосходящее вещественного аргумента x, можно определить спецификацией:

```

formula flp(real x, int y) = y <= x & x < y + 1;
floor(real x: int y) post flp(x, y);

```

Требуется построить быстрый алгоритм вычисления floor для чисел в плавающем формате, определяемом стандартом IEEE 754-2008. Плавающее число представляется тройкой (S, E, T), где S = {0, 1} – знак числа, E – сдвинутая экспонента, а T – мантисса в нормализованном представлении без старшего разряда (равного 1). Значение числа определяется по формуле:

$$z = (-1)^S * 2^{E-\text{bias}} * (1 + 2^{1-p} * T).$$

Здесь bias – сдвиг экспоненты, p – число битов в представлении мантиссы, $1 \leq E \leq 2^w - 2$, w – число битов в представлении экспоненты. Значения $E = 0$ и $E = 2^w - 1$ предназначены для кодирования нулей и специальных значений. Значения констант для формата 32 (**float**): $\text{bias} = 127$, $w = 8$ и $p = 24$; для формата 64 (**double**): $\text{bias} = 1023$, $w = 11$ и $p = 53$. Данные константы определим в программе описаниями:

```

const nat bias;
const subtype (nat i: i >= 2) w; // w >= 2
const subtype (nat i: i > 0) p; // p > 0
const nat m = p + w; // число битов в представлении чисел x и y

```

Здесь m – число битов в представлении плавающего числа. Определим типы для S , E и T :

```

type bit = {nat a | a = 0 or a = 1};
type nate = {nat e | 1 <= e & e <= 2^w - 2};
type natp = {nat i | i < 2^(p-1)};
type intm = {int i | -2^(m-1) <= i & i < 2^(m-1)};

```

Результат y должен иметь тип intm той же разрядности, что и аргумент x ; например, типу **float** должен соответствовать **int32**. Определим спецификацию функции **floor**:

```

formula val(bit S, nate E, natp T: real) = (-1)^S * 2^(E - bias) * (1 + 2^(1-p) * T);
formula Floor(bit S, nate E, natp T, intm y) = flp(val(S, E, T), y);
floor(bit S, nate E, natp T: intm y) post Floor(S, E, T, y);

```

Необходимо доказать тотальность спецификации, т.е. лемму:

```

Floor_total: lemma exists intm y. Floor(S, E, T, y)

```

Оказывается, доказать эту лемму невозможно. Причина в том, что для больших E значение y , удовлетворяющее предикату **Floor**, может выйти за пределы типа intm . Например, для формата $m = 32$ результат y может не поместиться даже в формате 64. Значение функции val ограничено по модулю значением $2^{(E - \text{bias} + 1)}$, поскольку $1 + 2^{(1-p)} * T < 2$. Отсюда следует правильное определение типа nate , для которого лемма **Floor_total** доказуема:

```

type nate = {nat e | 1 <= e & e < bias + m - 2};

```

Введем функцию vall :

```

formula vall(bit S, int d, nat z: real) = (-1)^S * 2^d * z;

```

Тогда $\text{val}(S, E, T) = \text{vall}(S, E - \text{bias} - p + 1, 2^{(p-1)} + T)$. Вычисление функции **floor** можно свести к вычислению функции **floor1** со спецификацией:

```

formula P_f1(int d, nat z) = z >= 2^(p-1) & z < 2^p & d < w;
formula Floor1(bit S, int d, nat z, intm y) = flp(vall(S, d, z), y);
floor1(bit S, int d, nat z: intm y) pre P_f1(d, z) post Floor1(S, d, z, y);

```

Сначала следует доказать тотальность спецификации **floor1**, т.е. лемму:

```

Floor1_total: lemma P_f1(d, z) => exists intm y. Floor1(S, d, z, y)

```

Сведение **floor** к **floor1** реализуется доказательством леммы:

```

Floor_fb1: lemma Floor(S, E, T, y) & d = E - bias - p + 1 & z = 2^(p-1) + T =>
  P_f1(d, z) & Floor1(S, d, z, y)

```

По лемме генерируется следующее определение предиката **floor**:

```

floor(bit S, nate E, natp T: intm y)
  { floor1(S, E - bias - p + 1, 2^(p-1) + T: y) } post Floor(S, E, T, y);

```

Алгоритм для **floor1** реализуется разбором случаев на базе леммы

```

Floor1_def: lemma P_f1(d, z) & Floor1(S, d, z, y) =>
  (S = 0)? ((d >= 0)? y = 2^d * z : y = div(z, 2^(-d))) :
  (d >= 0)? y = -2^d * z :
  (mod(z, 2^(-d)) = 0)? y = -div(z, 2^(-d)) : y = -div(z, 2^(-d)) - 1

```

Здесь div – деление нацело, а mod – деление по модулю. Программа для **floor1** генерируется по лемме очевидным образом. Следует отметить, что результат вычисления $y = 2^d * z$ (и

$y = -2^d * z$) может оказаться неточным, если неточным было исходное представление числа в плавающем формате. Погрешность может достигать до 2^d .

Использование битовых операций над целыми позволяет заменить ряд конструкций программы на более эффективные. Применяются следующие замены:

- $2^{(p-1)} + T$ на $2^{(p-1)} \text{ or } T$;
- $z * 2^d$ на $z \ll d$;
- $\text{div}(z, 2^{(-d)})$ на $z \gg (-d)$;
- $\text{mod}(z, 2^{(-d)}) = 0$ на $(z \& (2^{(-d)} - 1)) = 0$.

Значение $2^{(-d)} - 1$ состоит из $(-d)$ единичных битов. Такое значение можно сгенерировать посредством сдвига: $(2^p - 1) \gg (p + d)$. Корректность указанных замен обеспечивается доказательством следующих лемм:

plus_eq: **lemma** $T < 2^{(p-1)} \Rightarrow 2^{(p-1)} + T = 2^{(p-1)} \text{ or } T$;
sh_left: **lemma** $d \geq 0 \ \& \ d < p \Rightarrow z * 2^d = z \ll d$;
sh_right: **lemma** $g > 0 \ \& \ g < p \Rightarrow \text{div}(z, 2^g) = z \gg g$;
mod_eq: **lemma** $g > 0 \ \& \ \text{mod}(z, 2^g) = 0 \Rightarrow (z \& (2^g - 1)) = 0$;
sh_mask: **lemma** $g > 0 \ \& \ g < p \Rightarrow 2^g - 1 = (2^p - 1) \gg (p - g)$;

С учетом указанных замен следует модифицировать лемму `Floor1_def` и получить по ней новую версию программы `floor1`. Далее тело `floor1` подставляется на место вызова в `floor`. Приведем окончательную программу, дополненную вычисление `floor` для нулевого аргумента, но не допускающую других специальных значений в качестве аргумента:

```
floor(bit S, nate E, natp T: intm y)
{ int d = E - bias - p + 1; nat z = 2^(p-1) or T;
  if (E = 0 & T = 0) y = 02
  else if (S = 0) {if (d >= 0) y = z << d else if (-d >= p) y = 0 else y = z >> (-d) }
  else if (d >= 0) y = - (z << d)
  else if (-d >= p) y = - 1
  else if ( (z & ((2^p - 1) >> (p + d))) = 0 ) y = - (z >> (-d))
  else y = - (z >> (-d)) - 1
};
```

4 Смежные работы

Наиболее популярным методом синтеза является извлечение программы из конструктивного (интуиционистского) доказательства (метод `proofs-as-programs`) [6] с применением правил резолюции и унификации. Синтез предикатных программ проводится в классической логике и, в соответствии с обзором [5], реализуется по шаблонам формул (метод `schema-guided`); в качестве шаблонов используются правила доказательства корректности. Метод шаблонов применяется в основном для синтеза логических программ. В *предикативном* программировании Э. Хехнера [7] программа синтезируется как результат последовательных уточнений (`refinements`) спецификации программы применением правил логического вывода.

В работе [9] представлен синтез (методом `proofs-as-programs`) трех программ вычисления целочисленного квадратного корня применением трех разных схем доказательства по индукции теоремы тотальности: $\forall x \exists m. \text{Isqrt}(x, m)$. В отличие от [9] метод синтеза, предлагаемый в настоящей работе, является универсальным и позволяет синтезировать произвольный алгоритм. Построение эффективного алгоритма вычисления целочисленного квадратного корня, модификации алгоритма `sq3`, описано в работе [8]. Исходная программа на языке `Standard ML` верифицирована автоматическим доказательством условий корректности в системе `Nuprl`. Конечная программа – результат серии оптимизирующих трансформаций, переводящие про-

² Из стандарта не очевидно, что `floor(-0)=0`.

грамму на язык HML, а затем на язык интегральных схем. Корректность трансформаций доказывается в Nuprl. Версия алгоритма в разделе 2 сопоставима с реализацией в [8]. Ее можно было бы оттранслировать на язык интегральных схем. Однако процесс построения программы в разделе 2 представлен на одном языковом уровне и значительно проще, чем в [8].

Заключение

Первоначально каждая из трех стандартных функций была запрограммирована на языке P, причем для `isqrt` и `ilog2` реализовано по три версии алгоритмов разной сложности и эффективности. Наиболее эффективные версии алгоритмов преобразованы на язык C++ применением последовательности оптимизирующих трансформаций. Итоговые программы для функций `floor` и `isqrt` оказались почти тождественными с исходными. Программа для функции `ilog2` существенно быстрее исходной программы на C++.

Далее для каждой программы на языке P были сгенерированы условия корректности в соответствии с системой правил серии L. Сгенерированные формулы оттранслированы на язык спецификаций известной системы автоматического доказательства PVS [4]. Доказательство на PVS оказалось нетривиальным и трудоемким процессом, требующем времени на порядок больше по сравнению с программированием. Трудности возникали не столько при доказательстве самих условий корректности, сколько при доказательстве используемых математических свойств, особенно для битовой арифметики.

В процессе верификации обнаружено 5 ошибок. Из них четыре – чисто технические. Принципиальной является ошибка, обнаруженная для функции `floor`: результат функции `floor` для аргумента в формате 32 может не поместиться в формате 32. Это опасная ошибка, приводящая к неверным вычислениям.

Позднее в исследовательских целях был проведен программный синтез всех алгоритмов для трех функций. Оказывается, что в режиме синтеза процесс построения соответствующих теорий и логического вывода в системе PVS реализуется более естественным образом, более вариативно и с большей согласованностью между теориями. Это дает возможность сократить суммарный объем доказательств примерно в два раза по сравнению с режимом верификации. Следовательно, работа в режиме синтеза является более предпочтительной. В результате синтеза получается несколько иная предикатная программа, однако предыдущая программа может быть получена из нее применением нескольких простых эквивалентных преобразований.

Методы верификации и синтеза, предлагаемые в настоящей работе, могут быть применены при разработке бортовых программ самолетов и космических аппаратов, программ управления электростанциями, медицинскими приборами и в других приложениях с повышенными требованиями к надежности программ.

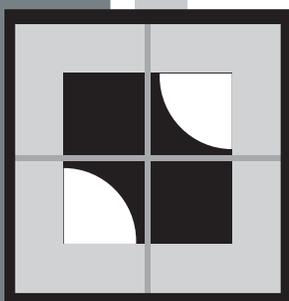
Список литературы

- [1] Карнаухов Н.С., Першин Д.Ю., Шелехов В.И. Язык предикатного программирования P. — Новосибирск, 2009. — 44с. — (Препр. / ИСИ СО РАН; N 153).
- [2] Shelekhov V. The language of calculus of computable predicates as a minimal kernel for functional languages // BULLETIN of the Novosibirsk Computing Center. Series: Computer Science. IIS Special Issue. — 2009. — 29(2009). — P.107-117.
- [3] Шелехов В.И. Модель корректности программ на языке исчисления вычислимых предикатов. — Новосибирск, 2007. — 50с. — (Препр. / ИСИ СО РАН; N 145).
- [4] PVS Specification and Verification System. — SRI International. — <http://pvs.csl.sri.com/>
- [5] Basin D., DeVille Y., Flener P., Hamfelt A., and Nilsson J. Synthesis of programs in computational logic. // LNCS, 3049. — P.30 – 65, 2004.
- [6] Sørensen M.H., Urzyczyn P. Lectures on the Curry-Howard Isomorphism. — 2006. — 457 p.
- [7] E.C.R.Hegner. A Practical Theory of Programming, second edition. — 2004. — <http://www.cs.toronto.edu/~hehner/aPToP/>

- [8] O'Leary J., Leeser M., Hickey J., Aagaard M. Non-Restoring Integer Square Root: A Case Study in Design by Principled Optimization. // Theorem Provers in Circuit Design, 1994 – P.52-71.
- [9] Creitz C. Derivation of a fast integer square root algorithm. – 2003. – <http://www.nuprl.org/Algorithms/03cucs-intsrt.pdf>.

Biography

The author was a developer, systems analyst, and manager in three big compiler projects: Alpha-6 for BESM-6, Algol-Elbrus, and SOKRAT for embedded satellite systems.



Multi-Agent Models and Systems
Мультиагентные модели и системы

USING TOOLS OF COMPLEXITY SCIENCE TO DIAGNOSE THE CURRENT FINANCIAL CRISIS

G. Rzevski

Complexity Science and Design, the Open University, UK
Chairman of the Board, Emergent Intelligence technology, Inc.
1431 Riverplace Blvd. Ste. 3205, Jacksonville FL, USA
Rzevski@gmail.com, www.emergent-intelligence.com
+1 904 638 9125

Key words: *Complexity, Self-Organization, Emergence, Evolution, Adaptation, Agent Based Software, Global Financial Systems*

Abstract

В работе рассматривается возможность применения концепции и средств «науки о сложном» для оценки состояния глобальной экономической системы, а также даются предложения по выявлению причин серьезного спада в экономике, который мы наблюдаем сегодня.

Introduction

The Internet based global economy consists of an exceedingly large and ever increasing number of Agents (suppliers, customers, investors, lenders, borrowers and middlemen) engaged in rich interaction and has all the characteristics of a complex system:

- The system has no central control but it is subject to certain externally imposed rules or norms of behaviour
- Its global behaviour emerges from the interaction of local behaviours of Agents and is unpredictable
- The system is so frequently disturbed that it has no time to return to its equilibrium and therefore is most of the time “far from equilibrium” or “at the edge of chaos”
- Because of system nonlinearity, a small disturbance may be amplified to cause a very large disruption, known as a “black swan” or “butterfly effect”
- The system is capable of autonomously self-organizing and is therefore adaptable and resilient
- The system co-evolves with its environment and resulting changes are irreversible
- In common with all social systems, its constituent Agents have declared and undeclared objectives and propensity to pursue these objectives in competition or co-operation with other Agents

The key problem is unpredictability of its global behaviour. For example, we know that the system exhibits cyclic increases and decreases of economic output but we cannot predict the timing of the next slowdown nor its severity.

Global financial system is a subsystem of the global economic system and it has all the characteristics of its parent system.

The paper applies concepts and tools of complexity science to the global financial system and suggests how to identify reasons for the severity of the current economic downturn.

It examines primarily the effects of structural factors of the global financial network on its global behaviour, taking into account features such as: (a) the speed of Agent interactions, (b) the number of links between Agents, (c) the adequacy of external rules of behaviour imposed upon Agent decision making processes, and (d) the adequacy of social representation in Agent negotiations.

Method

Research results reported in this paper have been achieved by monitoring and analysing behaviour of large-scale complex software systems developed under the leadership of the author and his close collaborators for a number of business clients, which are now in commercial use, and extrapolating observed features to the global financial system. Therefore the results are tentative and it would be prudent to test them by extensive simulation experiments.

Complex Systems

Examples of Complex Systems include: global economy; a national economy; ecology; a swarm of bees; a human being; human society; culture; climate; road traffic; a swarm of software agents and life span of an aircraft or a car.

Examples of complex systems where a capacity to rapidly self-organise provides substantial resilience to strong external attacks include: epidemics, human immune system, terrorist networks and the Internet.

A common feature of all systems listed above is that they do not have a centralised control – their behaviour emerges as a result of the interaction among their components and between components and the environment. The decision making power of different Agents may vary, but as a rule, Agents have certain autonomy limited only by accepted or imposed norms of behaviour; they do not work under precise (algorithmic) instructions. The global behaviour of such a system is not possible to predict.

Complexity versus Determinism

There are two contrasting philosophical theses [1] about the Universe.

The first thesis is that the world is based on the “grand design” and therefore its behaviour is predictable (deterministic). Any uncertainty is due to our inability to understand the world and to reduce uncertainty we have to discover laws of nature. Great philosophers and scientists subscribed to this worldview, including Aristotle, Kant, Newton and Einstein.

The second thesis is that the world is inherently unpredictable (complex) because it evolves with time often due to autocatalytic and self-accelerating properties of some of its elements. Its evolution is irreversible and leads to an increase in complexity. As Prigogine wrote [1], [2], the future is not given; it is under perpetual construction. Supporters of this thesis include Buddha, Darwin, Popper and Prigogine.

Whilst the notion of a deterministic world was very popular at the point in history when we witnessed great discoveries of Newtonian science, at present we seem to be overwhelmed by complexity of environments in which we live and work, and the dynamics of change is such that the maxim “you cannot enter into the same river twice” rings true. Under present circumstances the idea of a perpetually evolving world characterised by rich interaction of its constituent components seems very plausible [3].

Evolution Generates Complexity

The interesting questions are where complexity comes from and why there are currently so many complex issues?

There exists compelling evidence that as the evolution of our Universe takes its course, the ecological, social, political, cultural and economic environments within which we live and work increase in complexity. This process is irreversible and manifests itself in a higher diversity of emergent structures and activities and in an increased uncertainty of outcomes.

Society co-evolves with technology for wealth creation. Industrial Society, where the key resource was Capital and majority of people were employed in industrial production of goods, super-

seeded the Agricultural Society, in which the key resource was Land and majority of people were employed in agriculture. We have now entered a new transition from industrial to Information Society, the society in which the key resource is Knowledge and were majority of people are employed in knowledge-based services (information processing) rather than production of goods.

The current transition towards Information Society is particularly notorious by its very steep increase in complexity caused by globalisation as enabled by (a) the rapid spread of the Internet across the globe, (b) a considerable increase in opportunities (and reduction of costs) for global travel and (c) concerted efforts at reducing tariffs and adopting instruments that facilitate the global exchange of goods and services.

Table 1 – The complexity of markets increases with each transition

	KEY RESOURCE	DISTRIBUTION NETWORKS	SCOPE
Agricultural Society	Land	Local roads	Local
Industrial Society	Capital	Motorways & Railways	National
Information Society	Knowledge	The Internet	Global

Major Paradigm Shifts Created by the Transition to Information Society

Emergence of Global Economy

In the socio-economical sphere there is a very important shift from nation-centred industrial markets to the global economy, enabled by the abolition of tariffs and the rapid growth of the Internet. Global economy consists of a vast number of suppliers and customers rapidly matching supplies to demands and then, as rapidly, changing previously agreed matches as soon as better opportunities present themselves, each agent trying to increase their gain. The global market has all characteristics of a very complex system [4]. The overall distribution of resources to demands emerges from individual transactions. Globalisation increases wealth across the board (there is a saying among economists “globalisation tide lifts all boats”). According to recent statistics during the last 30 years the world as a whole has enjoyed an unprecedented economic growth, with one downside feature - the gap between rich and poor has increased in all countries except Ireland.

The ever increasing numbers of players in the global market and the speed of information flow over the Internet accelerated market dynamics to such an extent that the current generation of data-driven information systems supporting business processes will not be able to cope for long.

Let us illustrate this statement by describing problems experienced by global consumer product supply chains consisting of a large number of suppliers, warehouses, logistics operators, assembly plants and retailers distributed around the world, encompassing together a vast quantity of resources such as machine-tools, robots, conveyers, loading bays, trucks, freight trains, cargo planes and cargo ships, operated by hundreds of crew members, loaders, operators, progress chasers and supervisors. Such supply chains are rendered unmanageable by the current generation of enterprise resource planning systems, schedulers and optimisers because of the frequent and unpredictable occurrence of events such as the arrival of new orders, order cancellations, changes in previously agreed orders and delivery arrangements, delays, failures of logistics resources and human errors. The key problem is that the frequency of market driven changes is typically 1-2 hours whilst the time required to re-schedule a plant is not less than 8-10 hours; in global logistics, once a pallet is assigned to a pipeline its destination cannot be altered until it emerges from the other end of the pipeline, which may take several days.

We have to accept that complexity is a norm and that attempts to simplify complex situations, which was a useful managerial and technological philosophy in industrial society, when complexity of markets was much smaller, is now harmful.

Complexity of markets has to be exploited - it offers rich opportunities for those who master the mindset, skills and tools of adaptation and resilience.

Transition from Manufacturing to Knowledge-Based Services

The shift of manufacturing from the developed to developing countries is a part and parcel of globalisation. The replacement of manufacturing by knowledge-based services as the main wealth creating activity may be achieved only in countries where there exists advanced IT and a large number of high-class knowledge workers: researchers, designers and decision makers in financial services, IT, engineering, consulting, media, construction, architecture, entertainment, etc. The early adoption of knowledge-based services provides an excellent opportunity for economic prosperity at the time when manufacturing-focussed countries face stiff competition.

The limitation is that advanced knowledge-based services require systems capable of knowledge discovery and repository in a format suitable for the easy retrieval and application. Current IT was not designed for such applications.

Emergence of the Digital Leisure Industry and Social Websites

The current shift from paper-based and analogue media to digitally-enabled entertainment as the main leisure activity was made possible by the astonishing progress in designing miniature and powerful processors and memory chips, and appropriate communication and application software, as evidenced by the rapid spread of multi-channel digital TV, digital cameras, cellular phones and digital audio and video devices. This is complemented by a rapid growth of websites for downloading and uploading art, music and multimedia products. Apple iPod and iTunes together with YouTube and My Gallery are the icons of this revolution. The spread of the so called “social websites” where individuals display their creative products to be perused by global audiences is particularly significant and almost certain to continue well into the second half of this century.

The shift from corporate media to the internet-based media has many unexpected implications, not least the “long tail” [5], “the wisdom of crowds” [6] and “economy of attention” [7] phenomena.

Personal Globalization

Trough the specialised websites it is now possible to outsource many personal tasks such as: search for dating partners, personal secretarial tasks, personal website design, bookkeeping for freelancers, private maths coaching for pupils, and graphic design for weddings. The outsourcing provides opportunities to take advantage of better value for money, e.g., services of a college graduate in India for \$15 per hour (versus \$60 per hour in the USA), maths coaching from Bangalore for \$99 per month or graphic design from Argentina for a wedding at \$65 per job; but also, to reduce inequality of knowledge workers around the globe by providing freelance work for those in developing countries.

Supercrunching

As available computer storage space is increasing relentlessly and the Internet enables sharing of data, there is a trend to improve intuitive judgements of experts with extensive data mining aimed at discovering useful patterns in data, which amounts to using computers to discover Knowledge, the key resource of Information Society. The Supercrunching is used extensively in marketing to individual consumers but also in unexpected areas, which include: football coaching, medical diagnosis and in teaching – discovering which teaching methods work with which individual student.

Threats from New Forms of Violence

As globalisation takes its course with the rapid increase in the number of people connected to the Internet and/or travelling across the globe, new global complex systems have emerged, such as, terrorist networks, epidemics, massive hacking attacks, phishing and spamming.

Ineffectiveness of Very Large Systems

Under currently prevailing conditions of perpetual change very large systems appear to be too rigid and unable to adapt to rapid changes in their environments. Examples from the UK are many and include:

- Super-size comprehensive schools with 1,500 to 2,000 pupils have the worst record of bad behaviour in the classroom and limited academic achievements (source: the charity Human Scale Education). In contrast, smaller schools provide a student-friendly pedagogical environment and consequently good academic results.
- Very large private companies as a rule generate less profit than small-to-medium size enterprises.
- Exceedingly large computer systems, including National Health Service System, Ambulance Management System, Air Traffic Control System, Inland Revenue System and Drivers Licence System are too often delivered late, over the budget and, in some cases, never achieved expected performance.
- In contrast to difficulties experienced by large airline companies, a new generation of small and highly efficient jet aircraft, with 3 to 4 seats only, are starting a revolution in aviation by prompting a large number of start-ups promising to offer regional air taxi services.

The Prospect of Semantic Web

The next technological revolution will be the shift from the current data-driven Internet to the new Semantic Web and it is likely to be dramatic, as the new generation of systems will be enabled to “understand” the meaning of data. Semantic Web is an old dream of computer science researchers, which is now firmly on the way to be realised with first programs based on ontology and multi-agent software being released and tested in commercial applications. Further progress is ensured by the concentrated effort of a large number of researchers in EU and USA.

Functionality required to implement Semantic Web is fundamentally different from that offered by the current software technology. Clarifying semantics of sentences expressed in a natural language requires intelligent computational effort rather than brute force of current software. All successful current Semantic Web prototypes are based on ontology and multi-agent technology.

Cloud Computing

Information technology inexorably moves towards the full global interconnection of all digital processing devices and all digitised information content – towards the CLOUD.

There is a trend to store data on servers located on big “server farms”, distributed around the globe and well protected, and to access it from any point in the world where the users may happen to find themselves.

Our information-processing world is growing rapidly, its capacity doubling approximately every two years. In 2008 it was characterised by: 100 billions clicks per day; 4 billion digital devices (such as computers, telephones, RFIDs, etc) connected to the internet (in excess of the number of neurones in a human brain); 55 trillions links between web pages (similar to the number of synapses in a human brain); 2 million emails per second; 8 terabytes of traffic per second; 65 billion phone calls per year; 600 billion RFID tags in use.

It is estimated that by 2030 all computers, personal organisers, telephones, iPods, TVs, DVD players, modems, radios, film projectors, video players, hi-fi decks, including all their content, i.e., alphanumerical data, texts, films, videos, broadcasts, podcasts, newspapers, books, music, emails, blogs, websites and magazines, will be linked forming one Global Network - the Cloud.

Moreover, a great many natural and designed objects on the planet (livestock, aircraft, trains, vehicles, appliances and machines of all kinds) will be linked by means of RFID tags to the, so-called, The Internet of Things, which will be, naturally, a part of the Cloud.

And so will most people through the medium of Smart Cards and, at some stage, through Smart Clothing, Smart Watches, etc.

The idea of Cloud Computing is to have a total interconnectivity of all information creators, processors and consumers.

At the current state of development there are many smaller Clouds, including those operated by Google, Apple, IBM, Microsoft, etc. These Clouds will be gradually linked into the interconnected Global Cloud.

Black Swans

An inherent dangerous characteristic of all complex systems is their high nonlinearity, which may turn small disturbances into catastrophic outcomes, named Black Swans [8]. The most dramatic example of a Black Swan was the recent tsunami. As complexity of the internet-based global economy increases, we can expect from time to time the occurrence of unforeseen cataclysmic events affecting global economy, as exemplified by the current collapse of the global financial system. This is a downside of complexity and we can do very little about Black Swan causes. What we can do is to design information networks in a way, which minimises propagation and acceleration of instability through the networks, and to design our socio-technical systems to be adaptable and resilient. More on this topic later.

Using Complexity Science Concepts and Tools to Resolve Complex Issues

The Thesis

Complex situations can be effectively investigated only by using models that are of similar complexity as situations that are being modelled.

Leaving aside for the moment the definition of “similar”, the following argument supports this thesis - complex situations change during the modelling process and therefore their models must be capable to adapt to these changes; the adaptation must be autonomous (without waiting for instructions from the modeller) which is only possible if models have capabilities of self-organisation and intelligence. In other words,

Models must be able to co-evolve with situations that they model.

The key implication of this thesis is that conventional computer programs cannot be used as models for complex situations because they are not adaptable. They cannot change by themselves - they must be instructed to do so by programmers. At present, only multi-agent software technology can support adaptation.

Two Types of Complex Situations

Complexity of certain unresolved critical issues (such as global warming, poverty and population growth) is so high that we can expect, at best, to build much-simplified models of these issues with a view to gaining some insight into their resilience to our attempts to resolve them. Let us refer to such issues as Exceedingly Complex. An example of a brilliant work on modelling of exceedingly complex systems is a study of global population growth by Kapitsa [9].

In contrast, there are many complex issues where constituent components and their interactions can be identified and modelled by constructing a Virtual World (a model) almost as complex as the Real World that is being modelled. Nearly all complex business situations, as well as many complex social, economic, security and technological problems, as well as the evolution of urban systems, fall into this category [10], [11], [12]. I shall refer to such issues simply as Complex.

The Art of Modelling Complex Situations

The effective modelling consists of several stages:

- 1) *Collecting and organising domain knowledge* - The first step is to collect and organise knowledge on the domain of the real world that is being investigated. The most effective method

of representing knowledge on a complex domain is to construct a network in which nodes are domain Concepts and links are Relations between Concepts. For example, for a supply chain relevant Concepts include: Order, Supplier, Warehouse, Truck, Cargo, Road and Route. Each Concept is characterised by attributes and rules constraining its behaviour. Such domain knowledge representation is called Ontology [13].

- 2) *Constructing a Virtual World* - The next step is to build a Virtual World consisting of instances of Concepts and their Relations from domain Ontology. For a supply chain, a Virtual World will be a network in which nodes are Order O1, Order O2, ... Supplier S1, Supplier S2, ... Cargo C1, Cargo C2, ... Truck T1, etc, and links are "S1 supplies O1", "C1 is loaded on T2", etc. Complex systems, such as supply chains of large international organisations, and Virtual Worlds that represent them, may contain millions of objects, attributes, rules and relations. To construct Virtual Worlds for such complex problems one requires powerful multi-agent software tools [14].
- 3) *Connecting Virtual World to the Real World* - The Real World (i.e., a complex situation that is being modelled) is perpetually changing. The changes are represented by the occurrence of unpredictable events exemplified, in a supply chain, by: Arrival of an Order, Order Modification, Order Cancellation, Road Closure, Track Failure, etc. The occurrence of every Real Event must be communicated instantly to the Virtual World where an equivalent Virtual Event is created, causing the affected part of the Virtual World to adapt to changes originated in the Real World. Every change (adaptation) of the Virtual World must be communicated to the Real World before the occurrence of the next Real Event [15].
- 4) *Empowering Virtual World to manage the Real World in real time* - An Agent (a small computer program) is assigned to every node of the Virtual World with responsibility to maintain its integrity. For example, if a loaded Virtual Truck breaks down, the Truck Agent sends messages to Agents of all affected nodes letting them know that this Virtual Truck does not exist any more. The message provokes a flurry of activities among affected Agents who try to accommodate the failure by searching for a replacement. As soon as a solution is found, it is conveyed to the Real World for implementation, ensuring that the two worlds co-evolve (change in unison).
- 5) *Investigating behaviours of the Real World by experimenting in the Virtual World* - Once a suitable Virtual World is constructed in software, it could be used to simulate behaviours of the Real World under different states of its environment, e.g., studying behaviour of a supply chain under varying market conditions.

Examples of Solved Practical Complex Problems

The author and his team have designed a large number of agent-based software systems for investigating or managing large-scale complex situations, many of which have been implemented and are in regular use. Examples are described in detail in a series of papers and they include:

- Real-time scheduling of car rentals across Europe for one of the largest rent-a-car companies in the world [15]
- Real-time scheduling of over 2,000 taxis in London [16]
- Real-time scheduling of a fleet of very large tankers (10% of the world total capacity) transporting crude oil around the globe [17]
- Real-time production scheduling [18]
- Dynamic pattern discovery for a leading insurance company [19]
- Semantic analysis and search for a genetic research organization [20]
- Simulation of virtual enterprises [21].

The scale of the solved problems and the commercial success of developed systems have attracted attention of European Complexity Science Community, which singled out the author's work as a success case in applied Complexity Science.

Can we Learn from Complexity Science how to Avoid Financial Crises?

The answer is a tentative YES.

Financial networks are, after all, complex systems and therefore lessons learned during designing artificial complex systems, such as large-scale real-time multi-agent schedulers, should be applicable.

The key lesson learned from ten years of experience in constructing solutions to large-scale complex problems is that however carefully you design a complex system, it will not immediately work as expected. This is a rule, rather than an exception, because the overall behaviour of such systems emerges from the interaction of constituent Agents and is, in principle, not predictable.

However it is always possible to tune the behaviour of the designed system by adjusting critical system parameters. The issue we want to avoid is the propagation of oscillations through the system, which causes instability and may result in system disintegration. The result of successful tuning of the system is that its behaviour, although unpredictable in detail, never leaves the specified region of the behaviour space, and recovers speedily after each disturbance.

Main parameters that need tuning are:

- 1) The degree of autonomy given to Agents
- 2) The weight given to the Enterprise Agent
- 3) The speed of transactions between Agents
- 4) The degree of interconnectivity among Agents

Let us consider the importance of each above parameter in turn.

The degree of autonomy of Agents is perhaps the most important factor that determines the behaviour of a complex system. If Agents are severely constrained, the complexity is lost and with it the adaptability and resilience of the system. The software behaves exactly as a conventional computer program - it is deterministic - and if a situation is encountered that was not foreseen by the programmer, the program stalls. If, on the other hand, Agents are given too much freedom, the complexity increases and the system behaves erratically. To the best of author's knowledge there is no theory that is capable of prescribing the correct degree of freedom. Even more importantly, it appears that rules governing Agents need to be tighten up or relaxed depending on the dynamics of the situation. In the presence of extreme events it seems to be advantageous to give to Agent a greater freedom to rapidly find an effective way to adapt, whilst when the environment is less turbulent, it may be more appropriate to limit the risk of uncontrolled innovation.

Complexity is, by definition, present in situations where autonomous component interact in unpredictable manner. When agent-based software is used for modelling of complex situations, software Agents are assigned to each autonomous component with a view to exchanging messages among themselves and thus simulating component interactions. The author's experience shows that it is always advisable to introduce an additional software Agent to represent the situation as a whole. The author has named this Agent as the Enterprise Agent. Its role is to ensure that the no Component Agent gains an advantage for itself on expense of the system as a whole. It is important to underlie that Enterprise Agent does not issue instructions to Component Agents; it is rather an equal partner in negotiations.

Whenever a new event occurs that disturbs the current state of the system, it takes a certain period of time for the system to adapt to the new situation and for Agent negotiations aimed at achieving adaptation to settle. The excessive speed of transactions among Agents may lead to instability because it may cause the initiation of a new wave of Agent negotiations before the previous one had settled. On the other hand, a too slow speed may prevent the system to adapt to a disturbance in time before the next disturbance occurs. A skilful tuning of the transaction speed is required.

In a large network in which all nodes are connected to each other the interactions among nodes are completely unpredictable and often lead to extreme events - Black Swans. Reducing severely the interconnections limits the adaptability of the network. The best practical solution is to partition net-

works into regions with densely interlinked nodes within each region, and to interconnect regions sparsely. A proper partitioning limits the frequency of occurrence and intensity of extreme events.

As experience in tuning complex systems accumulated, it became possible to compile rules for managing complexity, which are articulated below. Consistent application of these rules during early stages of design reduces a need for tuning.

Rules for Managing Complexity

Let me reiterate that any attempt to simplify a complex problem, which is not fully under our control (such as global warming, or dynamics of the global market), is counter productive. The solution we obtain is a solution to a simplified, rather than to the original, complex problem.

However, if we are in a position to design from scratch, or to modify key parameters of a complex system (such as a multi-agent software or a global digital network linking financial institutions), there is a lot we can do to ensure its stability.

Key design rules, derived from discussions above, are quite straightforward:

- Construct network ontology incorporating rules that regulate, but do not curtail autonomy of Agents; continuously review the effectiveness of rules and modify them when required.
- Appoint Enterprise Agents to represent global community interests in Agent negotiations.
- Partition the network into regions (which may be logical rather than geographical) and limit the number of links and speed of information transfer between regions.
- Limit the speed of transactions among Agents.

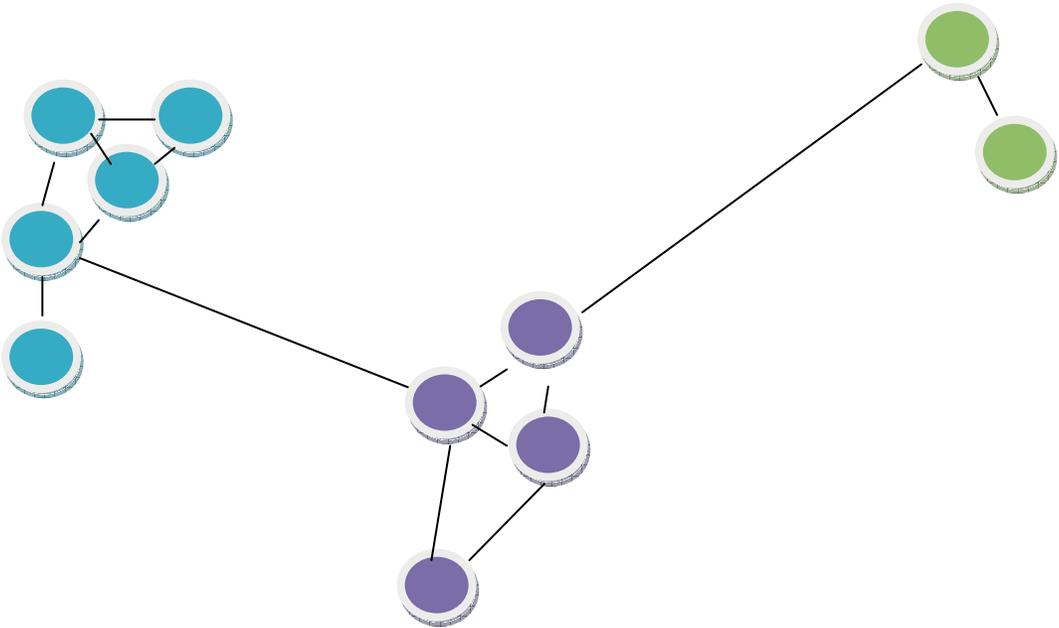


Figure 1 – Partitioning a network into regions

Let us apply these rules to an analysis of current highly interconnected digital network, which links financial institutions around the glob.

Conclusion: Designing Stable Global Financial Systems

According to analysis presented in this paper the current global financial network has the following *systemic* faults:

Regulations imposed upon financial institutions are not flexible - the degree of freedom given to bankers to innovate was excessive under conditions of stability and insufficient under conditions of the crisis.

The excessive speed of financial transactions and excessive interconnectivity of financial institutions increased the risk of instability and may have contributed to the occurrence of extreme events.

The solution may include:

- Partitioning of the sector in a number of logical or geographical regions to limit the risk of the propagation of extreme events throughout the whole global system.
- Developing Financial Services Ontology for each region, incorporating the best available expertise on banking, investments, financial trading and regional peculiarities.
- Constructing a set of multi-agent simulators, perpetually evaluating the performance of regional and total financial systems.
- Instituting intelligent regulations, possibly different for different regions, based on a continuous evaluation of system performance.

As stated at the beginning of this paper, these conclusions are derived by analogy rather than by direct investigation of the financial system and are therefore tentative. However, a study of the global financial system from the standpoint of Complexity Science would most probably yield similar results.

Acknowledgement

This research was only possible because of the most valuable contribution of Dr Petr Skobelev, who was in charge of a talented Russian software team that has developed large-scale intelligent systems used in research experiments.

Literature

- [1] Prigogine, Ilya, "The End of Certainty: Time, Chaos and the new Laws of Nature". Free Press, 1997.
- [2] Prigogine, Ilya, "Is Future Given?" World Scientific Publishing Co., 2003.
- [3] Holland, John, "Emergence: from Chaos to Order". Oxford University Press, 1998.
- [4] Beinhocker, Eric, "The Origin of Wealth: Evolution, Complexity and the Radical Remaking of Economics". Random House Business Books, 2007.
- [5] Anderson, Chris, "The Long Tail: How Endless Choice Is Creating Unlimited Demand", 2006.
- [6] Surowiecki, James, "The Wisdom of Crowds". Abacus, 2004
- [7] Lanham, Richard, "The Economic of Attention". The University of Chicago Press, 2007.
- [8] Taleb, Nassim Nicholas, "The Black Swan: The Impact of the Highly Improbable". Penguin Books, 2007.
- [9] Kapitsa, Sergey, "Global Population Blow-Up and After: The Demographic Revolution and Information Society". Report to the Club of Rome, 2006.
- [10] Rzevski G., "Investigating Current Social, Economic and Educational Issues using Framework and Tools of Complexity Science". Journal of the World University Forum, Volume 1, Number 2, 2008.
- [11] Rzevski G., Skobelev, P., "Emergent Intelligence in Large Scale Multi-Agent Systems". International Journal of Education and Information Technology, Issue 2, Volume 1, 2007, pp 64-71.
- [12] Rzevski G. "A New Direction of Research into Artificial Intelligence", Invited Paper, The Annual Conference of Sri Lankan Association for Artificial Intelligence, Colombo, 2008.
- [13] Rzevski G. "Ontology and Relativity". Volga Conference on Complex Adaptive Systems, Samara, Russia, 2005.

- [14] Rzevski G., Skobelev P., Andreev V. "MagentaToolkit: A Set of Multi-Agent Tools for Developing Adaptive Real-Time Applications". In Marik, V., Vyatkin, V., Colombo, A. W. (eds.) *Holonic and Multi-Agent Systems for Manufacturing*. Third International Conference on Industrial Applications of Holonic and Multi-Agent Systems, HoloMAS 2007, Regensburg, Germany, September 2007, pp 303-314. Springer LNAI 4659.
- [15] S. Andreev, G. Rzevski, P. Shveikin, P. Skobelev, I. Yankov, "Multi-Agent Scheduler for Rent-a-car companies", Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009), Decker, Sichman, Sierra, and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary.
- [16] Glaschenko A., Ivaschenko A., Rzevski G., Skobelev P. "Multi-Agent Real Time Scheduling System for Taxi Companies". Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009), Decker, Sichman, Sierra, and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary. 14. Glaschenko, A., Ivaschenko, A., Rzevski, G., Skobelev, P. "Multi-Agent Real Time Scheduling System for Taxi Companies". Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009), Decker, Sichman, Sierra, and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary.
- [17] Rzevski G., Himoff J., Skobelev P. "Magenta Technology: A Family of Multi-Agent Intelligent Schedulers". Workshop on Software Agents in Information Systems and Industrial Applications (SAISIA). Fraunhofer IITB. February 2006.
- [18] Andreev M., Rzevski G., Skobelev P., Shveykin P., Tsarev A., Tugashev A. "Adaptive Planning for Supply Chain Networks". In Marik, V., Vyatkin, V., Colombo, A. W. (eds.) *Holonic and Multi-Agent Systems for Manufacturing*. Third International Conference on Industrial Applications of Holonic and Multi-Agent Systems, HoloMAS 2007, Regensburg, Germany, September 2007, pp 215-225. Springer LNAI 4659.
- [19] Rzevski G., Skobelev P., Minakov Ig., Volman S., "Dynamic Pattern Discovery using Multi-Agent Technology". Proceedings of the MATH, SIP, TELE-INFO Conferences in Dallas, March 2007.
- [20] Rzevski G., Skobelev P. "Agent Based Semantic Web", Semantic Universe Website, 2009.
- [21] Rzevski G., Skobelev P., Batishchev S., Orlov A.: "A Framework for Multi-Agent Modelling of Virtual Organisations". In Camarinha-Matos, L M and Afsarmanesh, H (eds), *Processes and foundations for Virtual Organisations*, Kluwer Academic Publishers, 2003, pp. 253-260.

ИСПОЛЬЗОВАНИЕ АГЕНТНО-ОРИЕНТИРОВАННОГО ПОДХОДА К РЕАЛИЗАЦИИ МЕТОДОЛОГИИ СИСТЕМНОГО АНАЛИЗА

А.И. Павленко

Московский авиационный институт
Москва, Волоколамское шоссе, 4, Россия
kaf 302@mai.ru
тел: +7 (499) 158-44-33

Ключевые слова: *системный анализ, проблемная ситуация, этапы, интеллектуальные агенты, определение целей*

Abstract

Consideration is given to the applicability of the agent-oriented approach for realizing the system supporting the systems analyses methodology at a variety of decision-making phases. The system analyses steps are analyzed in terms of the operation of the decision support system. Its architecture and tasks for intellectual agents in estimating the ultimate goals are investigated.

Введение

В организационных системах управления обычно всегда предусматривается существование лица, принимающего решение (ЛПР), или группы лиц, ответственных за принятие решения. В зависимости от степени важности решаемого вопроса к анализу существующей ситуации и прогнозированию последствий тех или иных решений привлекается, как правило, группа специалистов: системных аналитиков, консультантов, советников, экспертов.

Необходимость привлечения ЛПР при подготовке решений дополнительных специалистов основана на том, что глобальные, стратегические, принимаемые на перспективу решения, обычно требуют сложных аналитических исследований, трудоемких вычислений, интерпретации получаемых на различных этапах исследований результатов и т.д.

Широкое внедрение в практику компьютерных систем позволяет аналитикам при подготовке решения активно использовать результаты математического и имитационного моделирования ситуаций, анализировать имеющуюся неопределенность и т.д.

Таким образом, при принятии управленческих решений кроме ЛПР имеется еще одно звено, участвующее в подготовке решения: аналитики, эксперты.

Необходимость обеспечивать высокую эффективность решений в данных условиях привела к развитию достаточно универсального подхода к решению проблем, который получил название – системный анализ.

Современные концепции организационного управления и принятия решений в условиях высокой неопределенности и слабой структурированности проблем основаны на использовании методологии системного анализа, основные этапы которой в настоящее время достаточно детально определены.

Вместе с тем, отсутствие у специалистов звена подготовки решений опыта в использовании методологии системного анализа часто приводит к тому, что решения готовятся в спешке, «поверхностно», часто по не до конца продуманным основаниям.

В связи с этим, становится актуальной задача компьютерной поддержки действий ЛПР, позволяющей ему более эффективно реализовывать свои знания и опыт не только для анализа подготовленных решений, но также и на этапах подготовки решений.

Эту задачу возможно решить лишь при использовании методов искусственного интеллекта – создании *интеллектуальных агентов*, способных реализовывать часть функции системных аналитиков и взаимодействовать с другими агентами и ЛПР.

Под *интеллектуальным агентом* будем понимать вычислительный процесс, постоянно и автономно функционирующий на компьютере, способный к восприятию состояния среды и к выработке решений в соответствии со сформулированными перед агентом целями.

1 Системный анализ и агенты

Сложность современных систем, их многофункциональность и многокритериальность, высокие степени риска при принятии решений требуют всестороннего учета сложившихся условий и оценки последствий принимаемых решений, что привело к широкому использованию методов системного анализа.

В настоящий момент, методология системного анализа в самом общем виде рассматривается как совокупность определенного набора действий системного аналитика, которые можно представить в виде последовательности этапов действий с обратными связями, позволяющими в случае необходимости вернуться к более ранним этапам исследования. В ряде работ по системному анализу [1-6] приводится последовательность этапов действий аналитиков в различной степени детализации. Можно выделить следующие этапы:

а) уяснение проблемы, б) структуризация проблемы, в) определение конечных целей, г) разработка вариантов решения, д) анализ ограничений, е) анализ взаимовлияния целей, альтернатив и ресурсов, е) принятие решения, ж) реализация решения.

Каждый этап характеризуется перечнем конкретных задач, выполнение которых можно рассматривать в качестве целей этапа. Кроме того, исполнение каждого этапа требует применения методов и экспертных процедур, свойственных именно данному этапу.

Несмотря на то, что каждый этап, предполагает применения творческих нестандартных действий, он также несет в себе и рутинные операции, основанных на знании нужных процедур и методов и исполнении формальных действий.

Рассмотрим решение слабоструктурированной проблемы требующей для своего решения применения методологии системного анализа, как формальную задачу *<Дано, Требуется>*: $\langle \Sigma, x^* \rangle$, где Σ - проблемная ситуация, x^* - рациональные решения, обеспечивающие достижения поставленных целей.

Под проблемной ситуацией Σ будем понимать следующую систему [7]:

$$\Sigma = \{T, Z, S, H, V, W, X, Q, M, \xi, P, \Delta F_{\Delta T}\}$$

где T – момент времени возникновения проблемной ситуации; Z – тип задачи; S – ситуация, соответствующая состояниям предметной области внешней среды; H – характеристика неопределенности, присущая решаемой задаче; V – множество целей, характеризующих принимаемое решение; W – множество критериев; X – множество управляющих воздействий x ; Q – множество ограничений; M – множество альтернативных вариантов принимаемых решений, ведущих к достижению целей; $\xi : X \rightarrow M$ – функция, отображающая набор управляющих воздействий в пространство решений; P – предпочтения ЛПР; $\Delta F_{\Delta T}$ – «облако» последствий решений в момент $T + \Delta T$.

Тип задачи – Z определяется в основном кругом вопросов, решаемых ЛПР. Так, в системах организационного управления можно укрупнено рассматривать следующие типы задач: оптимизацию отдельных показателей, выбор, упорядочение, планирование, распределение ресурсов, компенсацию возникающих отклонений от плановых показателей, обслуживание, составление расписаний, управление запасами и др.

Последовательное выполнение этапов системного анализа предполагает постепенное оп-ределение и уточнение компонентов проблемной ситуации Σ .

Для уяснения возможностей применения формальных методов, которые можно возложить на интеллектуальных агентов, рассмотрим первые три этапа методологии более детально.

Так первый этап – уяснение проблемы может быть представлен в виде последовательности следующих действий системного аналитика:

- обнаружение симптомов проблемы;
- концептуальный анализ проблемы;
- концептуальное определение проблемы;
- анализ развития проблемной ситуации;
- анализ типа возможного вмешательства в проблемную ситуацию.

Поскольку данный этап предполагает в основном построение концептуальной модели проблемной ситуации, то инструментарий системного аналитика содержит набор процедур и методов, обеспечивающих диалог ЛПР с компьютером на естественном языке (в рамках предметной области проблемной ситуации) и исполнение сценариев диалога, позволяющих зафиксировать в компьютере концептуальное (вербальное) описание ситуации.

Полученное на первом этапе описание служит исходным материалом для построения формализованной модели S , характеристики которой определяются на втором этапе анализа: структуризация проблемы.

Данный этап системного анализа ориентирован на исследование проблематики и определение структуры самой проблемной ситуации и предполагает выполнение следующих действий:

- исследование потребностей (или определение позиции наблюдателя [1] или определение миропонимания ЛПР [2]);
- выявление проблематики;
- выделение проблемы из среды;
- разработка концептуального сценария проблемной ситуации;
- классификации элементов;
- выявление структуры;
- определение функций;
- построение модели текущего состояния системы – «как есть» (*as is*);
- исследование ресурсных возможностей.

В формализованном виде проблемная ситуация Σ может быть задана посредством создания справочников всех понятий (объектов) предметной области проблемы $S = \{S_1, S_2, \dots, S_n\}$, свойств (атрибутов) объектов – A , качественных или количественных параметров атрибутов и выявления связей различного типа, определяющих структуры проблемы.

Инструментарий аналитика на данном этапе должен обеспечить переход от концептуальной модели ситуации к модели ситуации S , включающей справочники понятий, атрибутов A , диапазонов параметров D и матриц смежностей, отражающих связи между понятиями. Методы классификации и кластеризации, графовые модели и методы анализа структур составляют необходимый набор инструментов данного этапа.

Следующий этап системного анализа – определение целей, позволяет на основе сформулированной проблемы и установленной проблематики сформулировать *цель* – направление, в котором надо двигаться, чтобы решить проблему.

Понятие цели является центральным в проблеме принятия решений. Под целью понимается желаемое состояние результатов деятельности. Именно достижением тех или иных целей руководствуется ЛПР при принятии решений.

Выбор конечных целей зависит от характера решаемой проблемы. Так как система предназначена для сопровождения процессов принятия решений в организационных системах, то и

конечные цели должны быть конкретными. Предполагается, что ЛПР может сформулировать эти цели хотя бы в виде декларации.

Поскольку проблема представляет собой не что иное, как несоответствие между желаемым и существующими состояниями субъекта, и поскольку у нас имеется уже определенная модель существующего состояния, то возникает естественное желание определить модель желаемого состояния: «как должно быть» (*to be*). Желаемое состояние определяется на основании интервьюирования ЛПР с целью выявления его представлений о том, что следует изменить в текущей ситуации. Поскольку уже была построена модель «как есть», то предъявление ЛПР этой модели облегчает ему формулирование цели.

Формализованное представление желаемого состояния позволяет более четко сформулировать то, что существенно для нашей дальнейшей работы – определить *цель*.

Несмотря на наличие модели ситуации «как должно быть», определение целей все равно представляет собой достаточно сложную задачу из-за субъективного характера целеполагания.

Фиксация множества всех целей, вытекающих из сравнения состояний «как есть» и «как должно быть» необходима для дальнейшей работы аналитика. После того как цели сформулированы, появляется возможность выбора связанных с ними критериев.

Выбор критериев также представляет собой один из наиболее сложных вопросов системного анализа. Каждая цель является многомерной и требует введения нескольких критериев. Поэтому в общем случае задачи системного анализа - многокритериальные.

С учетом развития подходов, позволяющих ввести в процесс решения интеллектуальных агентов, следует рассмотреть возможность передачи части функций системного аналитика компьютерным программам, являющихся виртуальными интеллектуальными агентами.

В связи с этим, рассмотрим задачу поддержки действий ЛПР как *мультиагентную систему* [8,9], выполняющую по существу роль системного аналитика, и отслеживающую при принятии решений ЛПР выполнение всех этапов системного анализа.

Архитектура мультиагентной системы поддержки решений включает в себя двухуровневую иерархию агентов (см. рисунок 1).

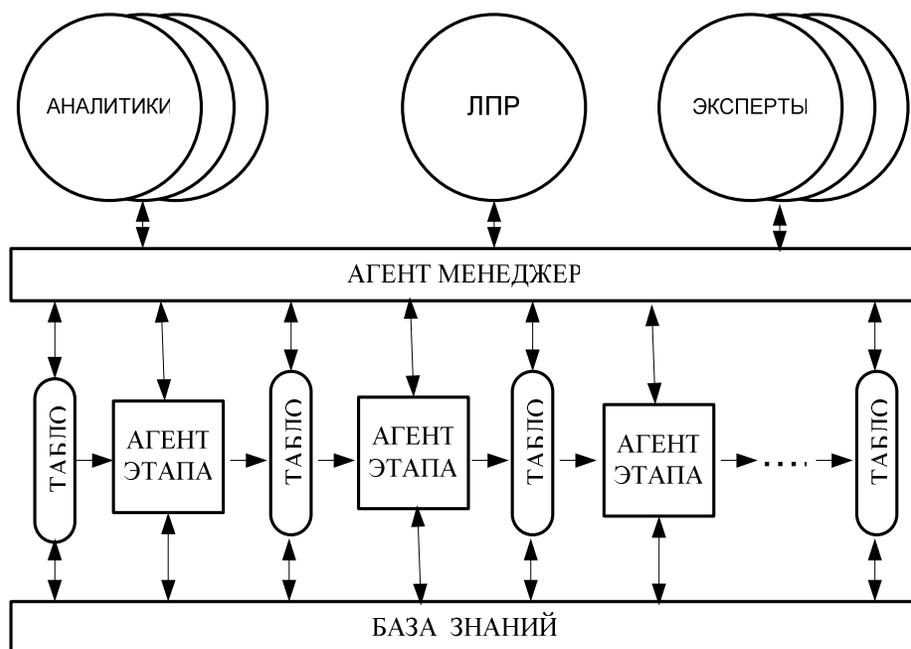


Рисунок 1 – Архитектура агентно-ориентированной системы поддержки решений

Верхний уровень представляет собой *агента менеджера*, обеспечивающего контроль работы всех агентов более низкого уровня иерархии – *агентов этапов*. Каждый агент этапа выполняет только ему свойственные функции аналитика ситуации подчиненного ему этапа. Агенты этапов могут общаться между собой посредством «табло взаимодействия», на котором отражается информация, полученная агентом в процессе его функционирования.

Агент этапа имеет собственную цель – максимально полное заполнение «табло взаимодействия», находящегося в его сфере ответственности. В том случае, когда он не может собственными силами достичь поставленной цели, агент формирует запрос к агенту-менеджеру (и к агенту предыдущего этапа), снижая тем самым последнему достигнутый им максимум заполнения «табло» и понуждающему его к действиям направленным на достижение максимума.

Для полноценного взаимодействия агентов и «понимания» ими запросов других агентов требуется соответствующая онтология предметной области проблемы (храняемая в базе знаний системы), позволяющая в едином информационном пространстве строить и понимать запросы агентов друг к другу.

Структуру данных «табло взаимодействия» агентов этапов готовит агент менеджер, на основе запросов к ЛПР. Агент менеджер обладает наиболее интеллектуальными способностями из всех виртуальных агентов и только он один может общаться напрямую с ЛПР, или транслировать последнему запросы агентов этапов. Такая архитектура системы упрощает организацию диалогового взаимодействия компьютерной системы поддержки решений с ЛПР.

Рассмотрим более подробно работу агентов этапов на примере *агента этапа определения целей*.

Поскольку конечную цель обычно не удается сразу связать со средствами ее достижения, требуется осуществить разбиение глобальной цели на более мелкие локальные цели, позволяющие выявить средства их достижения.

При этом идет выявление составляющих предметной области S , которые могут служить в качестве управляющих, т.е. определяются параметры, находящиеся в распоряжении ЛПР и оказывающих влияние на принимаемое решение.

Выбор целей решаемой задачи V определяется целями, задаваемыми ЛПР в текущий момент времени T . Как правило, эти цели формулируются на вербальном уровне, и для количественной оценки уровня достижения цели вводятся критериальные функции $w(x)$.

Таким образом, агент данного этапа получает через «табло взаимодействия» предыдущего этапа информацию о параметрах проблемной ситуации S , включающую список понятий (объектов) предметной области, список атрибутов этих понятий A , значения (или интервалы возможных значений) параметров D и множество отношений между понятиями R .

С другой стороны, эффективность работы данного интеллектуального агента будет определяться после заполнения им своего табло взаимодействия с агентами последующих этапов системного анализа. Это табло должно содержать список целей V , и отношений между отдельными целями (дерево целей) $-R_V$, перечень критериев, характеризующих степень достижения поставленных целей и коэффициенты относительной важности целей K_V .

Задачу агента данного можно представить как задачу

$$\langle (S, W_{\text{эц}}, (V, R_V, w(x), K_V) \rangle,$$

где S – состояние среды, $W_{\text{эц}}$ – критерий эффективности агента этапа целей, перечень целей – V , дерево целей – R_V , K_V – коэффициент относительной важности критериев $w(x)$.

Решение данной задачи основывается на владении агентом методами построения деревьев целей и коэффициентов относительной важности.

При этом можно выделить три основные ситуации при построении дерева целей:

- 1) анализируя модель «как есть» ЛПР, может указать, что в текущей ситуации его не устраивают только значения некоторых параметров атрибутов объектов;

- 2) более сложная ситуация возникает когда ЛПР на множестве имеющихся объектов (понятий) указывает на потребность изменения отношений между объектами;
- 3) потребности ЛПР заключаются в необходимости включения в модель ситуации новых понятий и объектов.

Связь «ситуация – дерево целей» позволяет определить классы деревьев целей, используемых на данном этапе.

В настоящее время, используется несколько подходов к построению дерева целей [10]:

- структурный,
- контент-анализа,
- репертуарных решеток.

Поскольку метод построения дерева целей не единственный, то агент должен обладать интеллектом, чтобы в зависимости от сложившейся ситуации выбрать возможность применения того или иного метода построения дерева целей или даже использовать совокупность методов, обеспечивающих максимальную эффективность его деятельности.

Представляет интерес подход, основанный на привлечении ЛПР к построению им когнитивных карт (карт «причина – следствие»), на основе предъявления ему информации о ситуации «как есть» и выявления его представлений о желаемом состоянии. Поскольку ЛПР вынужден формулировать свои цели в терминах, заложенных в описание ситуации S , то такой подход к определению целей облегчает построение дерева целей и выбор критериев эффективности.

Заключение

Таким образом, предложенный в данной работе подход к разработке систем компьютерной поддержки решений, основанный на реализации методологии системного анализа интеллектуальными агентами основан на объективной возможности возложить часть рутинных операций системного аналитика на компьютер, что позволит ускорить процессы подготовки решений.

Список литературы

- [1] Черняк Ю.И. Системный анализ в управлении экономикой. М.: Экономика, 1975.191 с.
- [2] Дж. ван Гиг. Прикладная общая теория систем. Т. 1, 2.М.: Мир,1981, 730 с.
- [3] Сурмин Ю. П. Теория систем и системный анализ. Киев : МАУП, 2003. 368 с.
- [4] Голубков Е.П. Системный анализ как методологическая основа принятия решений. М : Менеджмент в России и за рубежом, №3 / 2003.
- [5] Павленко А.И., Росин М.Ф. Математическое обеспечение систем принятия решений. М.: МАИ, 1990. 42 с.
- [6] Павленко А.И. Введение в системный анализ. М.: МАИ-ПРИНТ, 2008, 96 с.
- [7] Павленко А.И. Формализация задач принятия решений и выбора. . М.: МАИ-ПРИНТ, 2009, 86 с.
- [8] Городецкий В.И., Грушинский М.С., Хабалов А.В. Многоагентные системы (обзор) // Новости искусственного интеллекта, 1998, №2, с.64-116.
- [9] Садыков С.С., Симаков Р.А. Агентно-ориентированный подход к автоматизации работы пользователей. //Международный сигнал «Программные продукты и системы», 2005, №1, 6 с.
- [10] Абрамов А.Л., Свиридова А.Ю. Методика построения дерева целей стратегического плана с применением репертуарных решеток. //Современный экономический анализ на Дальнем востоке России: позиция молодых исследователей. М. 2007, с. 86-126.

СИСТЕМА ФИЗИЧЕСКОЙ ЗАЩИТЫ НА ОСНОВЕ АГЕНТНО-ОРИЕНТИРОВАННОГО ПОДХОДА И НЕЧЕТКОЙ ЛОГИКИ

А.Д. Тарасов

Оренбургский государственный аграрный университет
460795, Оренбург, ул. Челюскинцев, 18, Россия
тел./факс: +8 (3532) 77-52-30

***Ключевые слова:** система физической защиты, система контроля доступа, агентно-ориентированный подход, нечеткая логика, фазификация, дефазификация, нечеткие базы знаний Мамдани*

Abstract

It is described mathematic modeling of objects and processes with systems of physical defense by using agent-oriented methods and fuzzy logic (on example of standard package Fuzzy Logic in MATLAB system).

Введение

Развитие современных программно-аппаратных средств позволяет существенно повысить защищенность объекта от несанкционированного доступа посторонних лиц. Широкое распространение получает применение автоматизированных систем физической защиты (СФЗ) в частности систем контроля доступа (СКД).

Разработка программного обеспечения сложных интегрированных комплексов и компьютерных сетей имеет большое значение для расширения возможностей компьютерных систем поддержки принятия решений и повышения эффективности процессов обработки данных в компьютерных системах. Один из способов проектирования программного обеспечения для СФЗ это моделирование в виде распределенной системы состоящей из автономных модулей, т. е. применение агентно-ориентированного подхода.

Во время решения задачи построения интегрированной системы физической защиты появляется проблема использования неточной, неполной и часто неоднозначной информацией [1]. Для работы в такой информационной среде можно применять неклассические виды логик, например, так называемую нечеткую логику.

1 Многоагентные системы

Многоагентная система (МАС) – это система, образованная несколькими взаимодействующими интеллектуальными агентами. Многоагентные системы могут быть использованы для решения таких проблем, которые сложно или невозможно решить с помощью одного агента или монолитной системы.

Интеллектуальные агенты – программные объекты (особый род компьютерных программ), способные к взаимодействию друг с другом и анализу информации, полученной через их сообщения друг другу. В оправдание определения «интеллектуальные», Агенты должны быть способны к принятию решений в условиях неопределенности ситуации, действовать при отсутствии полной информации, хотя бы и в какой-либо узкой области. Интеллектуальный агент владеет определенными знаниями о себе и об окружающей среде, и на основе этих знаний он способен определять свое поведение. Как правило, агенты скорее обучены, чем запрограммированы для выполнения конкретной работы [2].

В случае применения многоагентного подхода при построении СФЗ, система представляет собой сеть взаимодействующих между собой автономных программно или аппаратно реализованных агентов, каждый из которых выполняет определенные функции в системе.

2 Нечеткая логика

Математическая теория нечетких множеств (fuzzy sets) и нечеткая логика (fuzzy logic) являются обобщениями классической теории множеств и классической формальной логики. Данные понятия были впервые предложены американским ученым Лотфи Заде (Lotfi Zadeh) в 1965 г.

Понятие лингвистической переменной играет важную роль в нечетком логическом выводе и в принятии решений на основе приближенных рассуждений. Лингвистической называется переменная, принимающая значения из множества слов или словосочетаний некоторого естественного или искусственного языка. Множество допустимых значений лингвистической переменной называется терм-множеством.

Задаваемые экспертами исходные данные в нечеткой форме с помощью операций нечеткого логического вывода позволяют получить значения выходных переменных. Выходные переменные в свою очередь могут быть входными для следующего уровня обработки нечеткой информации. Образуется иерархическая система, в которой информация от экспертов, в конечном счете, определяет, например, категорию объекта защиты.

Основой для проведения операции нечеткого логического вывода является база правил, содержащая нечеткие высказывания в форме 'Если-то' и функции принадлежности для соответствующих лингвистических термов.

Например имеется m правил вида:

R1: ЕСЛИ x_1 это A_{11} ... И ... x_n это A_{1n} , ТО y это B_1

...

Ri: ЕСЛИ x_1 это A_{i1} ... И ... x_n это A_{in} , ТО y это B_i

...

Rm: ЕСЛИ x_1 это A_{m1} ... И ... x_n это A_{mn} , ТО y это B_m ,

где x_k , $k=1..n$ – входные переменные; y – выходная переменная; A_{ik} – заданные нечеткие множества с функциями принадлежности.

Результатом нечеткого вывода является значение переменной y на основе заданных значений x_k , $k=1..n$. Такую нечеткую базу знаний можно представить в виде таблицы.

В общем случае механизм логического вывода включает четыре этапа: введение нечеткости (фазификация), нечеткий вывод, композиция и приведение к четкости, или дефазификация.

Алгоритмы нечеткого вывода различаются главным образом видом используемых правил, логических операций и разновидностью метода дефазификации. Разработаны модели нечеткого вывода Мамдани, Сугено.

Лингвистические правила генерируются экспертом либо получаются в результате обработки нечетких знаний из экспериментальных данных.

Исходная модель должна быть настроена (обучена) чтобы выдавать верный конечный результат. Правильность конечного результата определяется экспертами, т. е. эксперты должны указать ответ для конкретных исходных данных. Несколько наборов данных вида x_k , $k=1..n$ – входные переменные и соответствующая y – выходная переменная используются для настройки модели. В процессе обучения нечеткая база знаний изменяется таким образом, чтобы модель выдавала результаты как можно более близкие к мнению экспертов.

Обучение происходит автоматически с помощью существующих программных средств путем нахождения таких параметров нечеткой базы знаний, которые минимизируют отклоне-

ние модельных и экспериментальных результатов. После настройки модель можно использовать для любых исходных данных [3].

3 Модель формирования мнений агента на основе нечеткой логики

Предлагается модель СКД на основе многоагентного подхода (рисунок 1). Стрелками показаны направления передачи информации (сообщений) от одних агентов другим. Рассмотрим подробно действия одного агента. Агент-вахтер – модуль, отвечающий за пропуск посетителей на территорию объекта. Данный агент должен принимать решение о пропуске или задержке посетителя (об открывании или закрывании турникета), а также о подаче сигнала тревоги в случае нарушения пропускного режима или сбое в работе системы, и решение о простом сигнале автоматического открытия турникета для информирования охранника. Эти действия зависят от следующих сообщений других агентов:

- 1) Состояние карточки посетителя от агента сканера;
- 2) Информация из базы данных об уровне допуска посетителя;
- 3) Информация из базы данных о наличии посетителя на территории;
- 4) Действия агента охранника (желание открыть турникет вручную).

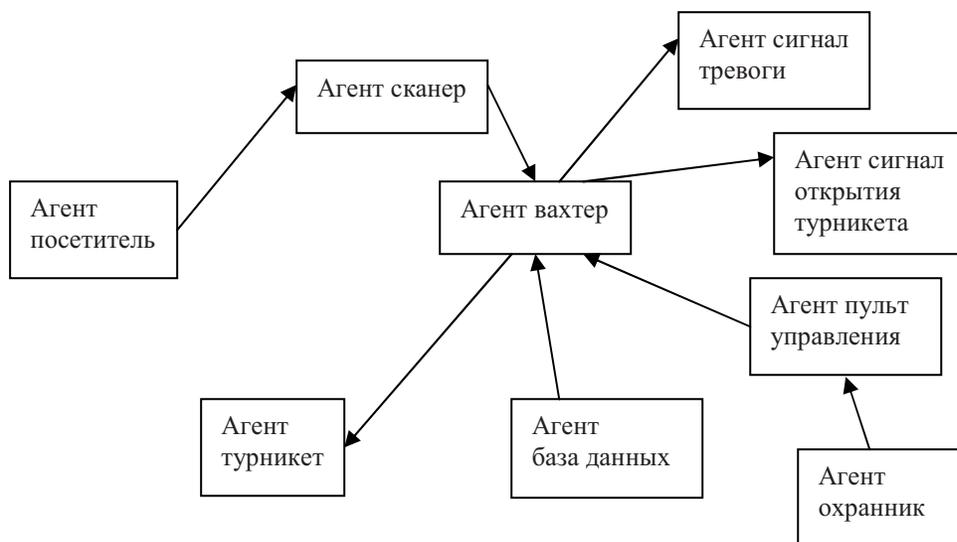


Рисунок 1 – Модель MAC системы контроля доступа

Соответственно для агента-вахтера имеется набор сообщений – исходных данных, на основе которых агент принимает решения, формирует свои мнения, которые будут представлены в виде сообщений для других агентов – выходных данных. Процесс преобразования входных данных в выходные представим в виде модели, в которой предлагается использовать методы нечеткой логики с помощью пакета Fuzzy Logic в системе MATLAB.

Модель принятия решения можно представить функциональным отображением: $X = (x_1, x_2, \dots, x_n) \rightarrow Y = (y_1, y_2, \dots, y_m)$, где X – вектор влияющих факторов, а Y – вектор решений

Влияние факторов отобразим в виде иерархического дерева логического вывода (рисунок 2). Элементами дерева являются:

- вершины - выходные данные – мнения агента (y_1, y_2, y_3);
- вершины - входные данные – влияющие факторы (x_1, x_2, x_3, x_4);

y_1 также является входным фактором для y_3 . Такая переменная называется укрупненным входным фактором.

Описание факторов приводится в таблице 1. Расчеты y_1, y_2, y_3 будут осуществляться посредством логического вывода по нечетким базам знаний.

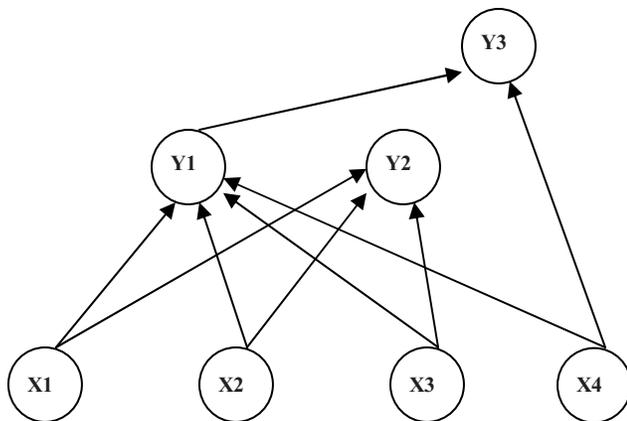


Рисунок 2 – Иерархическое дерево модели принятия решения агентом-вахтером

Таблица 1 – Входные и выходные данные в процессе формирования мнений агента-вахтера

Название	Описание
X1	Состояние карточки в устройстве считывания
X2	Информация об уровне доступа посетителя
X3	Информация о нахождении посетителя на территории
X4	Сигнал от пульта управления
Y1	Решение об открытии турникета
Y2	Решение о включении сигнала тревоги
Y3	Решение о подаче сигнала об автоматическом открытии турникета

3.1 Базы знаний для модели формирования мнений агента

Значения факторов будем рассматривать как лингвистические переменные. Для входных факторов значения определяются из следующих терм множеств:

- x1- “Присутствует, Отсутствует, Нераспознана”,
- x2- “Нет доступа, Есть доступ, Нет информации”,
- x3- “Отсутствует, Присутствует, Неизвестно”,
- x4- “Есть, Нет”,

для выходных факторов:

- y1- “Открыть, Закрыть”,
- y2- “Включать, Не включать”,
- y3- “Включать, Не включать”,

Для моделирования выходных факторов используются нечеткие базы знаний типа Мамдани. Переменные нечетких правил связаны логической операцией И. Правила нечеткой базы знаний заполняются с соблюдением условия: для любого набора значений входных переменных должно существовать хотя бы одно правило.

Откуда следует, что минимальное количество правил равняется числу всех возможных вариантов наборов значений входных переменных. В случае, когда количество термов у входных переменных одинаковое, количество наборов будет равно A в степени B , где A – количество термов входных переменных, B – количество входных переменных [4, 5].

В результате, например, для трех входных переменных с тремя термами в каждой необходимо сформулировать 27 правил базы знаний. Но в случае, когда при некоторых значениях одних входных переменных значения остальных не влияют на решение, количество правил сокращается. Например, при открывании турникета вручную (x_4 равняется “Есть сигнал от пульта”) на выходную переменную y_1 не влияют значения остальных переменных. База знаний с сокращенным набором правил показана в таблице 2.

Пример графика функций принадлежности нечетких термов для лингвистической переменной “Информация о нахождении посетителя на территории” в Fuzzy Logic приведен на рисунке 3. Используется треугольная функция принадлежности.

Таблица 2 – Нечеткая база знаний для формирования мнения агента-вахтера

x_1	x_2	x_3	x_4	y_1	y_2	y_3
Присутствует	Нет доступа	Отсутствует	Нет	Закрыть	Не включать	Не включать
Присутствует	Нет доступа	Присутствует	Нет	Закрыть	Включать	Не включать
Присутствует	Нет доступа	Неизвестно	Нет	Закрыть	Не включать	Не включать
Присутствует	Есть доступ	Отсутствует	Нет	Открыть	Не включать	Включать
Присутствует	Есть доступ	Присутствует	Нет	Закрыть	Включать	Не включать
Присутствует	Есть доступ	Неизвестно	Нет	Открыть	Не включать	Включать
Присутствует	Нет информации	Отсутствует	Нет	Закрыть	Включать	Не включать
Присутствует	Нет информации	Присутствует	Нет	Закрыть	Включать	Не включать
Присутствует	Нет информации	Неизвестно	Нет	Закрыть	Включать	Не включать
Отсутствует	-	-	Нет	Закрыть	Не включать	Не включать
Нераспознана	-	-	Нет	Закрыть	Включать	Не включать
-	-	-	Есть	Открыть	Не включать	Не включать

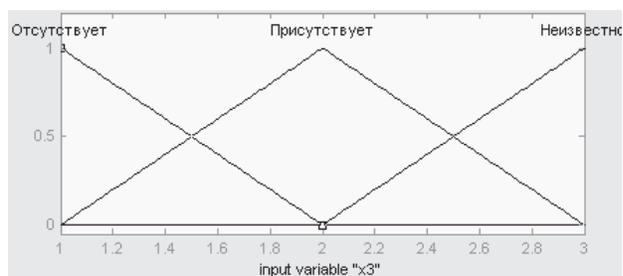


Рисунок 3 – Функции принадлежности нечетких термов для лингвистической переменной “Информация о нахождении посетителя на территории”

3.2 Нечеткий вывод для модели формирования мнений агента

При нечетком моделировании необходимо определять степени принадлежности входов к термам из базы знаний. При нечетких исходных данных необходимо определить степень принадлежности одного нечеткого множества значения входной переменной, к другому нечеткому множеству терму из базы знаний. Степень принадлежности равна высоте пересечения этих нечетких множеств.

На рисунке 4 отображен пример работы системы нечеткого логического вывода в Fuzzy Logic. Решение о подаче сигнала тревоги (y_2) выходная переменная, зависящая от трех входных – состояние карточки (x_1), уровень доступа (x_2) и нахождение на территории (x_3). Пронумерованные строки показывают работу правил из базы знаний: степень принадлежности переменных x_1 , x_2 и x_3 к термам задает принадлежность переменной y_2 . В правом столбце все результаты правил объединяются в итоговое значение выходной переменной, которое дефаззифицируется по методу центра тяжести (фигура справа внизу). Используются алгоритмы нечеткого вывода Мамдани. В качестве треугольной нормы выбрано умножение. Четкое значение $y_2=0,33$ приравнивается к нечеткому терму “Есть сигнал тревоги”.

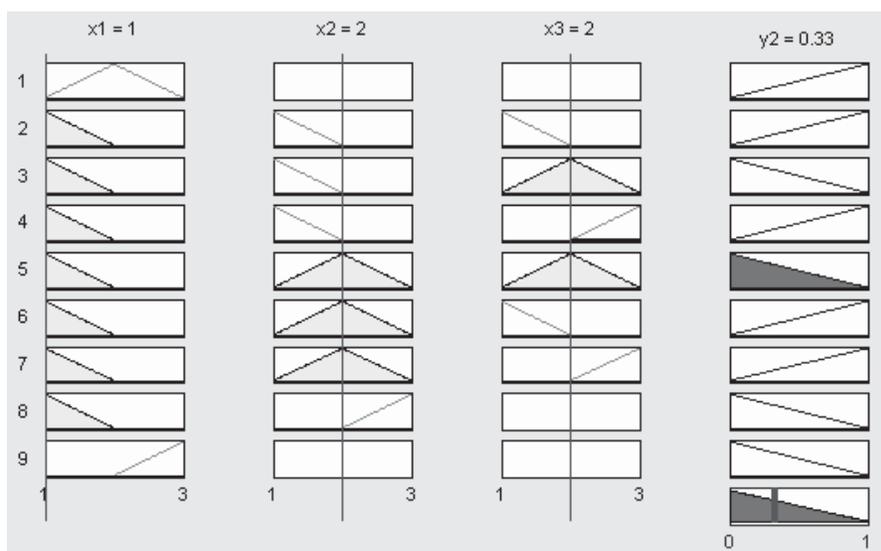


Рисунок 4 – работа системы нечеткого логического вывода

Заключение

При использовании средств Fuzzy Logic Toolbox нечеткая модель принятия решения агентом-вахтером реализована тремя системами нечеткого вывода:

V-y1.fis - нечеткая система принятия решения об открытии турникета,

V-y2.fis- нечеткая система принятия решения о включении сигнала тревоги,

V-y3.fis- нечеткая система принятия решения о подаче сигнала об автоматическом открытии турникета.

Иерархический нечеткий вывод по дереву (рисунок 1) осуществляется функцией Vahter.m. Текст программы набран на языке программирования MATLAB. Функция возвращает три выходных аргумента: значения Y1, Y2, Y3. Функция вызывается с четырьмя входными аргументами, которые задают значения факторов X1, X2, X3, X4. Логический вывод происходит через функцию evalfis.

Моделирование с использованием методов нечеткой логики в области СФЗ позволяет применить программные средства в процессе принятия решений агентов в многоагентном подходе с неточной и неполной информацией. Система поддержки принятия решений на этой основе повысит эффективность создания и работы программных средств для многоагентных систем в СФЗ.

Список литературы

- [1] Бояринцев А.В., Бражник А.Н., Зуев А. Г. Проблемы антитероризма: Категорирование и анализ уязвимости объектов. – СПб.: ЗАО «НПП «ИСТА-Системс», 2006. – 252 с.
- [2] Смирнов А.В., Шереметов Л.Б. Многоагентная технология проектирования сложных систем - Автоматизация проектирования. – 1999.- №1.
- [3] Заде Л. Понятие лингвистической переменной и ее применение к принятию приближенных решений. – М.: Мир, 1976. – 167 с.
- [4] Орловский С.А. Проблемы принятия решений при нечеткой исходной информации. – М.: Радио и связь, 1981. – 286 с.
- [5] Ротштейн А.П. Интеллектуальные технологии идентификации: нечеткая логика, генетические алгоритмы, нейронные сети. - Винница: УНИВЕРСУМ-Винница, 1999. – 320 с.

МЕТОД АДАПТИВНОГО ПЛАНИРОВАНИЯ В МУЛЬТИАГЕНТНОЙ СИСТЕМЕ МОДЕЛИРОВАНИЯ РАБОТЫ НАЗЕМНЫХ СЛУЖБ АЭРОПОРТА

А.В. Иващенко¹, У. Инден², П.О. Скобелев³, И.А. Сюсин³, А.В. Царев¹

¹Группа компаний «Генезис знаний»
443080, Самара, ул. Санфириковой 95, этаж 4, офис 416, Россия
ivashenko@kg.ru, at@kg.ru
тел: +7 (846) 222-91-72

²Cologne University of Applied Sciences
53913, Swisttal, Kuchenheimer Weg 18, Germany
udo.inden@fh-koeln.de
тел: +49(0)2633 473656

³Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
petr.skobelev@gmail.com, ilya.syusin@gmail.com
тел: +7 (846) 332-39-27

Ключевые слова: авиационная логистика, катеринг, мультиагентные технологии, онтологии, адаптивное планирование, ресурс, наземная служба

Abstract

The paper describes multi-agent approach to managing and scheduling of airport ground services' operations. The review of existing systems and researches in aviation logistic domain is given, advantages and lacks are found. The data structure, scheduling logic and system architecture are shown. The paper gives examples of algorithms of agents' communications, decisions making. The experience of applications of the system and future works are described.

Введение

Один из наиболее динамично развивающихся видов транспорта на сегодняшний день – это пассажирская авиация. Современные авиалайнеры – очень сложные высокотехнологичные комплексы, требующие продолжительного технического обслуживания (проверки, плановые осмотры, ремонт и т.д.). Но в связи с крупным пассажиропотоком и стремлением получать максимальную прибыль авиакомпания стараются уменьшить время пребывания самолета на земле: пока самолет не в воздухе, он не приносит прибыль. В современных динамичных условиях ведения бизнеса авиакомпания стремятся к оптимизации процессов управления всеми службами обеспечения и подготовки лайнера к полету, сокращения времени выполнения соответствующих операций и минимизации затрат на их выполнение.

В работе служб аэропорта ярко проявляются свойства сложной системы; аэропорт представляет собой открытую систему, которая подвержена постоянным и разнообразным воздействиям извне. Прежде всего, любой аэропорт является частью сети, в которой осуществляются рейсы. События, происходящие в одном аэропорте, могут непосредственно отразиться на деятельности другого, задержка вылета влечет за собой задержку прилета, погодные явления и особенности работы службы снабжения и события на борту влияют на работу служб в аэропорте прилета и т.д. В связи с этим, задача управления службами аэропорта связана с необходимостью адекватной и быстрой реакции на различные события, с чем не всегда способен эффективно справиться человек.

Таким образом, в современных условиях становится крайне актуальным применение новейших методов автоматизации динамического планирования для совершенствования системы

управления службами аэропорта, на основе широкого использования формализованных знаний о бизнес-процессах. При этом роль диспетчера изменится: вместо полного контроля над всеми процессами, что крайне сложно для человека и ведет к ошибкам и финансовым потерям, он перейдет на управление автоматизированной системой и обеспечение стабильности ее работы [1].

На сегодняшний день разработаны и внедрены системы поддержки принятия решений, способные работать, в том числе, и в сфере авиационной логистики. С одной стороны, многие авиакомпании (например, Lufthansa), имеют собственные подразделения по разработке программных комплексов. Однако при этом обычно не учитываются интересы аэропортов, а также основных подрядчиков (таких, как катеринговые компании). С другой стороны, некоторые компании фокусируют свое внимание на решениях проблем автоматизации работы наземных служб аэропортов и производят универсальные системы. Такие системы в целом решают поставленные перед ними задачи, однако не обладают такими возможностями, как адаптивное планирование, реакция на внешние события, контроль выполнения плана, оптимизация в реальном времени, интерактивное взаимодействие с участниками процессов, индивидуальный подход к каждому рейсу, сервису, ресурсу, реализовать которые позволяет использование принципов самоорганизации и интеллектуальных агентов.

В настоящей работе для решения задач планирования и оптимизации работы наземных служб аэропорта в реальном времени предлагается использование мультиагентных технологий (МАТ), поскольку:

- интеллектуальные агенты способны создавать временные связи для поиска решений, что позволяет оперативно адаптироваться к изменениям в реальном мире;
- процесс переговоров между агентами приводит к быстрому поиску подходящего решения, которое при последующих итерациях может улучшиться;
- использование общего информационного пространства – мира агентов – позволяет беспрепятственно переговариваться агентам, отвечающим за работу разных наземных служб, ресурсов, рейсов, аэропортов. Внедрение новых агентов в такое пространство происходит быстро, прозрачно и не оказывает негативного влияния на остальных участников переговоров;
- наличие у каждого агента собственных моделей поведения и целей позволяет инкапсулировать основные его свойства и скрыть приватную информацию от других агентов, действующих в том же мире агентов. Такой подход позволяет индивидуально описывать свойства и поведение каждой службы, ресурса, а также создавать модель взаимодействия агентов с конфликтующими целями (например, при переговорах подрядчика, стремящегося повысить прибыль, и авиакомпании, стремящейся снизить затраты).

Разработанная архитектура системы и реализованный набор компонент образуют собой современную программную платформу для разработки приложений для аэропортов, которая может иметь разнообразные применения. В частности, система разрабатывалась для решения следующих задач:

- поддержка принятия решений с возможностью планирования ресурсов и бизнес-процессов для аэропорта;
- моделирование виртуальной сети аэропортов, обладающей большим набором настраиваемых параметров моделирования, с возможностью контролировать положение и процесс обслуживания для авиакомпании;
- симулятор для апробации, прогона и оценки экономической эффективности бизнес-процессов в цехах катеринговой компании;
- создание модели виртуального аэропорта или сети аэропортов, способной в режиме реального времени обрабатывать поступающие внешние события и адаптироваться к ним таким образом, чтобы всегда максимально точно соответствовать реальному миру.

В настоящее время система используется для обработки RFID технологии для поставки продуктов питания на борт самолета в рамках проекта, реализуемого при поддержке Министерства экономики Германии в консорциуме с производителями авиационного оборудования, информационных технологий, логистическими и консалтинговыми компаниями (Airbus Germany, EADS, AutoID-Systems GmbH, Modular Galley Systems GmbH, B&W-Engineering GmbH). В ближайшем будущем планируется проводить испытания в тестовом режиме, а затем и в режиме полной нагрузки в аэропортах среднего размера (таких как Кельн-Бонн, около 100 рейсов в день), тесно взаимодействуя с уже установленными и работающими автоматизированными информационными системами.

1 Анализ проблемы

На сегодняшний день существует ряд программных комплексов, позволяющих решать определенный круг задач в сфере авиационной логистики, в том числе управления работой наземных служб аэропорта. Одна из таких систем – GroundStar [2], разработанная немецкой компанией Inform, функционирует более чем в 200 аэропортах по всему миру и предназначена для планирования, оптимизации и контроля выполнения планов в сфере пассажирских авиаперевозок. Основная функциональность системы включает расчет потребностей и построение графиков работы для ресурсов и служб (долгосрочное планирование) и построение суточных планов для ресурсов и служб (краткосрочное планирование).

Система GroundStar позволяет эффективно решать задачи управления наземными службами аэропорта, планирования их операций, производить определенную оптимизацию плана. Кроме того, программный комплекс задействует такие проблемные сферы, как поддержка диспетчера в плане управления ресурсами, управление взаимодействия с клиентами (CRM), наблюдение за процессами, происходящими в различных службах. Основное достоинство системы состоит в том, что при ее создании был использован большой опыт автоматизации действующих аэропортов.

Эта система предоставляет базовые функциональные возможности по планированию ресурсов, а также множество дополнительных модулей, сервисов и инструментов, находящихся в тесном взаимодействии и призванных решать широкий спектр задач. Однако и сам модуль планирования, и компоненты взаимодействия обладают рядом недостатков:

- централизованный подход к построению расписаний и оптимизации ограничивает возможности системы по масштабируемости и адаптации к новым условиям реального мира;
- отсутствует полноценный механизм адаптации плана к внешним событиям в автоматическом режиме, что не позволяет системе максимально учитывать реальную обстановку;
- отсутствуют средства контроля выполнения плана, интегрированные с подсистемой адаптации и перепланирования (имеется лишь модуль поддержки диспетчера).

В связи с этим можно сделать вывод, что система GroundStar не позволяет строить адаптивные планы, которые были бы адекватны быстро меняющейся реальной обстановке, обеспечивать возможность отслеживания выполнения расписаний в реальном времени и генерировать приемлемые (в соответствии с заданными критериями) решения в ограниченные промежутки времени.

Другой подход к планированию операций и управлению бизнес-процессами наземных служб в сфере авиационной логистики был предложен Сальденом и Мао [3]. Авторы одними из первых дали описание мультиагентной среды с агентами ресурсов и проектов, представили математическую модель для расчета целевых функций и стоимостей ресурсов, что в результате является критерием оценки качества созданных планов.

Для эффективного решения проблемы обеспечения снабжения рейса наземными службами в данной работе предлагается создание мультиагентной гетерогенной платформы для планирования, в которой основные действия совершаются агентами проектов и ресурсов. Для оп-

тимизации процесса построения плана и сокращения числа перепланирований наряду с предварительным планированием используется подход «непосредственного» (online) планирования, представляющий собой пошаговую корректировку предварительного плана в соответствии с поступающими извне данными. Такой подход позволит системе адаптироваться к изменениям в реальном мире.

Временные буферы между операциями используются для предотвращения и минимизации распространения незначительных изменений по всему расписанию. Длину буферов должны устанавливать агенты, используя метод совместного постепенного обучения (co-evolutionary learning), а также расчеты по предоставленным формулам и алгоритмам. Предлагаемая математическая модель для проблемы планирования множества заказов (multi-project scheduling problem (MPSP)) описывает проекты (заказы) как совокупность активностей (операций), ограничения на использование проектов и активностей (ограничения временных интервалов для активностей, порядок выполнения активностей, невозможность прервать начатую активность и т.д.) и ресурсы, их отношения с активностями. Для оптимизации планирования используются алгоритмы для вычисления функций целей агентов, стоимостей утилизации ресурсов и других показателей эффективности.

Основными достоинствами данной работы [3] является постановка и иллюстрация проблемы на простом примере – обслуживание рейса одной наземной службой, в котором показаны алгоритм и схема разрешения конфликтной ситуации за счет взаимодействия двух конкурирующих активностей, стремящихся использовать один и тот же интервал времени. Весьма высокую ценность представляет предлагаемый подход для решения поставленной задачи – использование неоднородной мультиагентной платформы для планирования. Авторами была изучена предметная область, выявлены основные проблемы и предложены методы их решения с помощью мультиагентных технологий: использование буферных времен, применение «непосредственного» планирования и другие. Важная роль в статье отводится формулам и алгоритмам для расчета показателей эффективности предлагаемой платформы и их применению.

Однако предложенный подход реализован на модельном примере и представляется слишком узким для построения согласованных планов различных служб, а приведенные примеры довольно сильно упрощены. Реальная система управления и планирования в сфере авиационной логистики должна быть намного более интеллектуальной, продвинутой, технологичной, для чего необходимо рассматривать следующие аспекты:

- виртуальный мир аэропорта (мир агентов) будет иметь множество типов, представляющих собой развитую сеть, что требует более глубокой проработки;
- взаимодействие между различными агентами (как разного типа, так и внутри одного типа) должно осуществляться с помощью системы эффективного обмена сообщениями;
- планирование десятков и сотен заказов в одном аэропорте наделит систему свойством эмерджентности, что должно быть учтено при ее проектировании;
- для возможности гибкого редактирования структуры ресурсов, бизнес-процессов, а также кастомизации системы необходимо использование онтологии;
- необходима реализация мультиагентной платформы для планирования, обеспечивающая высокую производительность и масштабируемость, а также промышленную реализацию создаваемых систем.

Все это делает актуальной и значимой задачу создания мультиагентной системы управления наземными службами аэропорта.

2 Постановка задачи и описание предлагаемого подхода к планированию работы наземных служб аэропорта

Аэропорт – сложная система, объединяющая множество самых различных служб, выполняющих разные функции, имеющих пересекающиеся сферы деятельности, порой конкури-

рующих между собой, но вынужденных тесно взаимодействовать при обслуживании одних и тех же ресурсов. Большинство наземных служб активизируются по приходу внешнего события (каким является, например, начало обслуживания самолета) и вынуждены выполнять определенный объем работ в строго регламентированные интервалы времени. Общая схема информационного взаимодействия между прибывающим самолетом и наземными службами аэропорта показана на рисунке 6. Когда самолет заходит на посадку, он сообщает о себе диспетчеру – в систему поступает событие «Прибытие рейса». Все наземные службы, предоставляющие услуги по обслуживанию прибывающего рейса, активизируются и занимают исходные позиции.

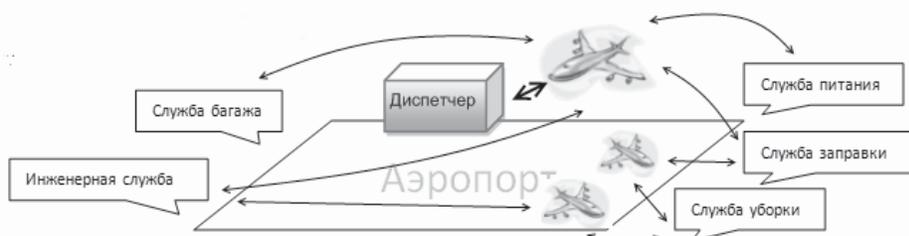


Рисунок 1 – Схема информационного взаимодействия между прибывающим самолетом и наземными службами аэропорта

Каждая наземная служба имеет собственную структуру, бизнес-процессы, ресурсную базу, поэтому к каждой из них необходим индивидуальный подход, подразумевающий как инкапсуляцию некоторых механизмов (таких как внутренняя логика, правила выполнения работ), так и организацию прозрачного взаимодействия между различными службами (в частности, передачи управления и данных).

В качестве примера можно привести организацию взаимодействия между пассажирской, заправочной службами и службой уборки авиалайнера. Каждая из них имеет собственную сферу ответственности при обслуживании общего ресурса – самолета. Однако сервисы не знают подробности работы друг друга – уборщики не имеют сведений о составе топлива, которым заправляется самолет, а заправщики не имеют данных о том, сколько пассажиров прибыло. При этом существуют жесткие правила взаимодействия этих трех служб: 1) по прибытию рейса все пассажиры должны покинуть борт самолета, 2) после высадки пассажиров самолет заправляют топливом, 3) после заправки самолет подвергается чистке и уборке, 4) после уборки пассажиры следующего рейса доставляются на борт. В соответствии с техникой безопасности порядок выполнения операций не может быть изменен, поэтому службы должны передавать друг другу информацию о завершении работ.

Будем называть планировщиком относительно автономный программный модуль, имеющий для окружающих его подсистем вид «черного ящика» и выполняющий функции создания адаптивного расписания для ресурса с учетом как внутренних (особенности бизнес-процессов, ограничения, условия), так и внешних (различные непредсказуемые события) факторов. Для своей работы планировщику необходимо получать данные о ресурсах, логике работы службы в качестве входных параметров. Выходными данными является созданное расписание. Кроме того, в процессе работы модуль должен незамедлительно получать все самые свежие порции информации, чтобы адаптировать к ним создаваемый план и всегда быть актуальным и адекватным внешнему миру. Для этого планировщик всегда имеет доступ к созданному им расписанию с возможностью вносить в него необходимые изменения.

Основная цель взаимодействия различных планировщиков между собой – разрешать конфликты, возникающие в ходе построения расписаний, находить компромиссы в сложных ситуациях. Критериями приемлемых решений являются качество и время обслуживания, а также стоимости рисков и возможных альтернатив.

В связи с описанной организации обслуживания наиболее эффективной является организация информационной сети, в которой узлами являются наземные службы, а связями – возможные направления взаимодействия между сервисами. При этом с каждой из служб должен быть соотнесен собственный планировщик, который будет настроен на наиболее эффективное решение задач именно для соответствующего сервиса. Планировщик будет учитывать все особенности бизнес-процессов связанной с ним службы и стремиться максимально повысить эффективность ее работы. Планировщики различных служб должны взаимодействовать посредством общей подсистемы обмена данными – общей шины предприятия (ESB), которая позволит оперативно передавать необходимые объемы данных между подключенными к ней модулями. Схема организации такого взаимодействия представлена на рисунке 2.



Рисунок 2 – Аэропорт как сеть взаимодействующих планировщиков: схема организации взаимодействия между отдельными модулями планирования посредством общей шины предприятия (ESB)

Такая организация представляет собой сервис-ориентированную архитектуру (SOA), позволяющую создавать слабозависимые, легкозаменяемые модули планирования для служб (например, в случае прекращения работы одних служб и/или появления новых). Непосредственное информационное взаимодействие между отдельными модулями, минуя общую шину, при такой организации практически отсутствует, что делает архитектуру прозрачной и позволяет эффективно управлять потоками информации в системе.

Поскольку задачи, выполняемые отдельными модулями планирования в составе сети планировщиков аэропорта, имеют общие черты, можно создать базовый модуль, обладающий всеми необходимыми функциональными возможностями, и на его базе создавать индивидуальные планировщики для каждой службы. Такой подход позволит создать систему, состоящую из схожих модулей, что значительно упростит ее реализацию и поддержку. Кроме того, разработка одного общего базового модуля значительно сократит затраты на внедрение в систему новых служб и упростит масштабирование системы.

Основные направления в разработке базового модуля включают обеспечение возможности настройки на различные типы мобильных и стационарных ресурсов и реализации различных операций бизнес логики, используя базовые типы операций, такие как «транспортировка», «загрузка», «выгрузка», «ремонт», «чистка» и т.д. Кроме того, базовый модуль планирования должен реализовывать следующие функции:

- анализировать состояния ресурса на возможность его использования;
- создавать расписание для ресурсов в соответствии с правилами бизнес логики;
- оперативно реагировать на различные внешние события;
- вносить изменения в части расписания, не нарушая целостности плана;
- учитывать различные показатели эффективности при планировании и перепланировании.

При этом базовый планировщик должен предоставлять возможность переопределять эти и другие функции в дальнейших реализациях, а также выполнять стандартные операции по обмену и сохранению информации.

3 Метод и реализация мультиагентного планирования ресурсов наземного сервиса

Для реализации мультиагентного подхода к решению задач планирования наземных служб аэропорта необходимо представить объекты предметной области в качестве сети с четко определенными сущностями, их ролями и целями, а также с правилами взаимодействия между этими сущностями. Систему, в которой имеются какие-либо ресурсы (производственные, персонал), а также использующие их услуги (операции) или заказы, которые необходимо оказывать или выполнять, можно представить как совокупность множества потребностей и множества возможностей (ПВ-сетей) [4, 5].

Все заказы и ресурсы системы взаимодействуют, вступая в связи между собой. В этом контексте можно считать, что каждый ресурс или заказ раздваивается на две противоположности: возможности и потребности. Например, грузовик для перевозки багажа для рейса имеет возможность перевозить определенное количество багажа, но при этом имеет потребность в наличии водителя. «Характеры» заказов и ресурсов также оказываются прямо противоположными: ресурсы более консервативны и должны стремиться быть максимально использованными и просуществовать как можно больший срок (для чего необходимо осуществлять специальный регламент, такой как отдых для человека или профилактика для техники), а заказы, наоборот, более активные и стремятся как можно быстрее реализоваться с наилучшими характеристиками, в частности, как можно меньше использовать ресурсы, чтобы меньше потратить денег на расчеты с ними.

Данную концепцию можно применить к модели виртуального аэропорта с наземными службами: потребностью в этом случае выступает заявка (заказ) на обслуживание рейса, причем каждый рейс создает множество заказов и рассылает их различным службам в сети. Службы, в свою очередь, стремятся выполнить полученные заказы с максимальной выгодой (минимальными издержками), привлекая для этого ресурсы различного типа. Ресурс играет роль возможности, при этом в системе могут быть различные типы ресурсов (мобильные, стационарные, персонал и др.), в ПВ-сети аэропорта они будут единообразно взаимодействовать с потребностями-заказами.

В качестве примера взаимодействия потребностей и возможностей рассмотрим процесс обслуживания рейса катеринговой компанией. Практически все авиакомпании на сегодняшний день предлагают пассажирам возможность перекусить на борту самолета во время полета. При этом меню рейса может содержать сотни элементов продуктов питания и сопутствующих товаров. Кроме того, каждый рейс обязательно должен обеспечиваться такими аксессуарами, как питьевая вода в баллонах, туалетная бумага, свежая пресса и т.д. Доставкой всей этой продукции занимается катеринговая компания. Задачи ее – обеспечить рейс всем необходимым (не больше и не меньше) на определенном уровне качества, при этом затратив минимальное количество средств. Задачи аэропорта – отправить рейс с доставленной вовремя всей необходимой продукцией на борту, также затратив на это минимальное количество средств. Как видно, задачи эти некоторым образом противоречат друг другу и могут быть решены с помощью проведения переговоров агентов. Приемлемым решением будет считаться компромисс, удовлетворяющий условиям как катеринговой компании, так и аэропорта.

Одним из преимуществ использования концепции ПВ-сети является возможность нахождения компромиссных решений, которые порой не являются оптимальными в глобальном смысле, но позволяют получить вполне приемлемое решение в условиях ограничений на время поиска и недостатка информации. Эта особенность является важным преимуществом при использовании в сфере авиационной логистики, где часто возникают непредвиденные ситуации, требующие принятия новых решений в режиме реального времени.

Мультиагентная платформа использует виртуальный мир агентов: множество интеллектуальных агентов различных типов, переговаривающихся между собой посредством обмена сообщениями и взаимодействующих с другими подсистемами. Основные типы агентов и их описание представлены в таблице 1.

Одним из основных достоинств предлагаемого подхода является возможность создавать динамические расписания и возможность адаптировать их к меняющейся обстановке в реальном мире. Это возможно за счет внедрения функций сдвига для операций, цепочки которых в конечном итоге обеспечивают заказ. Агенты операций при возникновении конфликтов ведут переговоры друг с другом. Цели этих переговоров, с одной стороны, устранить конфликты (возникающие в основном из-за захвата разными операциями одних ресурсов), а с другой – оптимизировать процесс выполнения заказа путем поиска более эффективного решения.

Таблица 1 – Основные типы агентов

Тип агента	Назначение	Модель поведения	Цель
Агент ресурса	Хранит информацию об использовании ресурса в прошлом, расписание на будущее. Предоставляет информацию о доступности ресурса в заданные интервалы времени. Является основным типом агента возможности.	Стремится оптимизировать использование ресурса, в частности за счет повышения его утилизации. При одновременном поступлении заказов может выбирать более выгодный (например, заказ с большей стоимостью).	Повышение эффективности использования ресурса на заданном интервале времени.
Агент заказа	Регистрирует заказ в системе, ищет все необходимые ресурсы для его выполнения, договаривается с другими агентами заказов для выявления и предотвращения конфликтных ситуаций (например, при использовании одного и того же ресурса). Является основным типом агента потребности.	Создает цепочку необходимых операций, контролирует их выполнение. Стремится выполнить заказ максимально эффективно (с низкими издержками).	Обеспечить успешное и своевременное выполнение заказа с минимальными затратами.
Агент операции	Каждый заказ состоит из одной или последовательности операций, т.о. агент операции реализует выполнение заказа. Является одним из типов агента потребности и основным типом агента подпотребности.	Ищет ресурсы, находящиеся в определенных состояниях. При этом старается минимизировать стоимость выполнения операции и всего заказа в целом.	Обеспечить выполнение заказа, поиск подходящего решения с минимальными затратами.
Агент события	Регистрирует внешнее событие в системе, обеспечивает передачу информации, пришедшей в сообщении, остальным агентам. Отвечает за своевременную реакцию системы на событие, а также за однократное его выполнение (без повторений).	Стремится как можно быстрее распространить информацию о событии среди агентов системы. Вносит изменения в сцену в соответствии с типом и содержанием события.	Обеспечить корректное и своевременное выполнение события.

Основные функции, позволяющие операциям изменять расписания и реализующие механизм адаптивного планирования, включают:

- поиск и проверку набора ресурсов на доступность в заданном временном слоте (временным слотом будем называть непрерывный конечный интервал на числовой оси, который может быть занят операцией для своего выполнения);

- сдвиг операции на Δt или на новый временной слот. При этом отдается предпочтение вариантам с использованием тех же ресурсов. Если это невозможно, производится поиск новых ресурсов;
- сдвиг операций на ближайший доступный временной слот левее или правее;
- отмена операции с освобождением всех захваченных ресурсов. Используется при невозможности выполнения заказа в данной обстановке. Однако операция при этом не удаляется и может быть вновь запланирована при наступлении благоприятных обстоятельств. Заказ при этом считается активным, но невыполненным;
- удаление операции. При этом освобождаются все захваченные ресурсы и стирается вся информация об операции. Такая операция в дальнейшем не может быть перепланирована. Используется при отмене рейса. Заказ при этом считается неактивным и невыполненным.

Основная идея адаптивного планирования – создавать расписание как систему неустойчивого равновесия, которая при отсутствии внешних воздействий остается неизменной, однако при возникновении каких-либо возмущений легко разрывает связи между связанными с возмущением объектами и пытается установить новые связи в соответствии с изменившимися данными. При этом благодаря отсутствию жесткой структуры и постоянных связей перепланирование для системы является естественным процессом и может происходить постоянно, что необходимо для использования в таких подверженных постоянным изменениям сферах как авиационная логистика.

Гармоничное использование механизмов планирования и перепланирования позволяет реализовать такую важную функциональную возможность, как контроль за выполнением плана. Эта функция позволяет:

- в режиме реального времени отслеживать текущее состояние расписания, распознавать моменты старта операций;
- ожидать в соответствии с заданными правилами и протоколами сигналы, подтверждающие начало выполнения необходимых операций;
- подавать предупредительные сигналы или выполнять иные действия, если сигнал подтверждения выполнения операции не получен;
- перепланировать фрагменты расписания в соответствии с полученными сигналами подтверждения начала выполнения операции или отсутствием этих сигналов (сдвинуть операции, если фактическое выполнение началось не вовремя; перепланировать или отменить операции, если подтверждение не пришло);
- пометить операции, для которых пришел сигнал подтверждения успешного начала и окончания как фактически выполненные – такие операции в дальнейшем не могут быть изменены или удалены;
- перепланировать расписание (частично или полностью) в соответствии с полученными сигналами об окончании выполнения операций. Данная возможность позволяет оперативно распознавать и исправлять ошибки, допущенные при выполнении операций.

Описанная выше функциональность была реализована в мультиагентной системе адаптивного планирования ресурсов наземного сервиса. Для удобства работы с разными типами данных применен многооконный интерфейс (рисунок 3). В правой части экрана расположен динамический список событий, зарегистрированных в системе. В правой верхней части экрана показан список обслуживаемых рейсов с указанием их статусов. В правой нижней части экрана находится окно с ошибками, предупреждениями и рекомендациями, выдаваемыми системой в процессе планирования.

Также при реализации системы особое внимание было уделено вопросам интеграции, в частности была разработана подсистема генерации событий, позволяющая унифицировать процесс обмена данными с внешними системами и выделять логику и интерфейсы взаимодействия с каждым отдельным комплексом в отдельный модуль. Модуль взаимодействия с внеш-

ними системами на платформе Sonic был испытан и протестирован на выставке Demonstrator, в которой, помимо описываемой системы планирования также участвовали диспетчерская служба, предоставляющая информацию о рейсах и заказах, несколько катеринговых компаний, обеспечивающих рейс провизией. Информационные системы всех партнеров были объединены посредством Sonic и взаимодействовали по протоколу OpenTRANS. В результате была продемонстрирована высокая слаженность работы всех информационных систем, а система планирования и управления продемонстрировала реальную возможность не только составлять эффективные расписания, но и контролировать их выполнение посредством обмена сообщениями, а также в режиме реального времени корректировать расписание в связи с изменениями в работе систем партнеров.

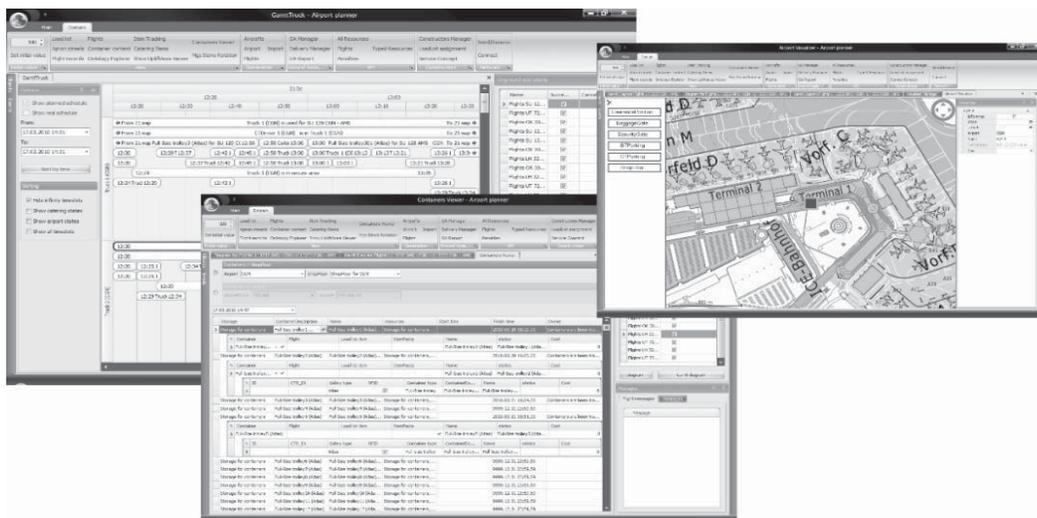


Рисунок 3 – Экраны, показывающие работу системы в различных режимах

4 Пример решения проблемы конфликтов планирования наземных служб

Продемонстрируем пример решения конфликтов планирования на примере доставки провизии для рейса к самолету. Онтология, описывающая этот бизнес-процесс, показана на рисунке 4. Рассмотрим процесс формирования расписания:

- 1) Получено событие о начале подготовки самолета к отправлению. Активизируются различные службы, в том числе катеринговая. Грузовик в это время находится в гараже пустой, контейнеры также пустые ожидают в специальном хранилище, самолет стоит в зоне обслуживания.
- 2) Агент заказа извлекает из онтологии данные, необходимые для планирования бизнес-процесса. По полученным данным агент заказа определяет, что необходимо создать цепочку операций «Сборка-Загрузка в грузовик-Транспортировка-Загрузка в самолет».
- 3) Планирование операций начинается от последней к первой. Для выполнения операции «Загрузка в самолет» создается агент этой операции, который обращается к БД и пытается найти необходимых ему агентов потребностей – самолета, грузовика, заполненных контейнеров, которые доступны в течение всей операции и подходят по заданным критериям. Если агенты найдены, операция начинает вести переговоры с ресурсами с целью установления связи, если не найдены, то агент операции может попытаться адаптироваться к обстоятельствам согласно описанию в таблице 1. Если запланировать операцию все же не удастся, то заказ считается активным, но невыполненным. Если все ресурсы найдены, операция пытается найти возможность для своего успешного выполнения – контейнеры

должны быть транспортированы к месту стоянки самолета. После выполнения этой операции грузовик должен находиться в зоне обслуживания самолета, контейнеры с провизией должны быть загружены в самолет.

- 4) Если все ресурсы для операции «Загрузка в самолет» найдены, она требует планирования предыдущей операции в цепочке – «Транспортировка». Для этого также создается агент для этой операции, который ищет в БД доступные и подходящие ресурсы – грузовик, заполненные контейнеры. При этом предпочтение будет отдано ресурсам, связанным со следующей операцией, поскольку и грузовик, и контейнеры должны быть использованы одни и те же во всех операциях цепочки. Если ресурсы из следующей операции оказываются недоступны, может быть найден и другой набор ресурсов, однако в этом случае придется перепланировать следующую операцию. Если операции не удастся установить связи с ресурсами, заказ считается активным, но невыполненным. Если все ресурсы найдены и с ними могут быть установлены связи, операция пытается найти возможность для своего успешного выполнения – контейнеры должны быть загружены в грузовик. После выполнения этой операции грузовик должен находиться в зоне обслуживания самолета, контейнеры с провизией должны быть загружены в грузовик.

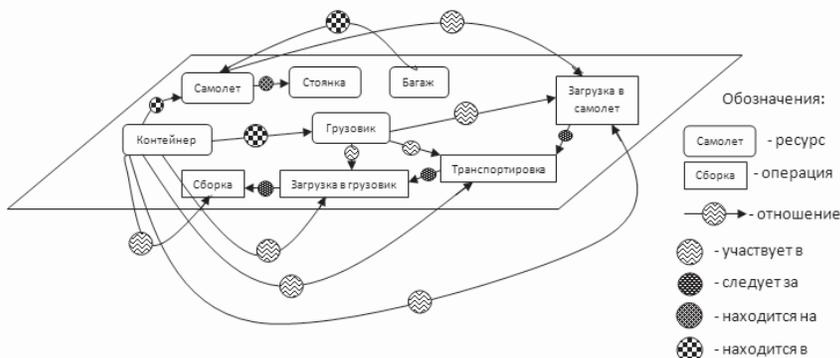


Рисунок 4 – Онтология бизнес-процесса доставки провизии для рейса к самолету

- 5) Если все ресурсы для операции «Транспортировка» найдены, она требует планирования предыдущей операции в цепочке – «Загрузка в грузовик». Для нее создается агент, который пытается найти в БД необходимые ресурсы – грузовик, заполненные контейнеры, зону погрузки. При этом зона для загрузки может быть выбрана любая (в соответствии с заданными критериями), а грузовик и контейнеры должны быть неизменными на протяжении всех операций. Если ресурсы не найдены, заказ считается активным, но невыполненным. Если все ресурсы найдены и задействованы в операции, она пытается найти возможность для своего успешного выполнения – контейнеры должны быть загружены необходимой провизией, грузовик должен стоять пустой у зоны погрузки. После выполнения этой операции Грузовик должен находиться в зоне погрузки, контейнеры с провизией должны находиться в грузовике.
- 6) Если все ресурсы для операции «Загрузка в грузовик» найдены, она требует планирования первой операции в цепочке – «Сборка контейнера». Для этой операции создается агент, который ищет необходимые ресурсы в БД – контейнер, провизию для загрузки, рабочую станцию. При этом грузовик уже не задействован, для него планируется параллельная операция транспортировки из гаража. Таким образом, операция «Загрузка в грузовик» требует выполнения двух параллельных операций – «Сборка контейнера» и «Транспортировка грузовика из гаража». После выполнения этой операции пустой грузовик и собранные контейнеры с провизией должны находиться в зоне погрузки.

7) При успешном планировании всех операций цепочки заказ считается активным и выполненным. Однако какие-либо изменения в сцене могут привести как к отмене, так и к перепланированию этого заказа, возможно, с использованием других ресурсов.

Данный пример показывает, как реализован процесс планирования поступившего заказа. Концепция ПВ-сети реализуется не только в отношении агентов ресурсов и заказов (операций), но и между операциями – одна из них может иметь потребность в выполнении другой.

Рассмотрим случай, описанный выше, при возникновении конфликта – недоступность ресурса типа грузовик. Пусть в системе зарегистрированы 2 грузовика (см. рисунок 5), которые уже участвуют в выполнении трех заказов. Допустим, в систему поступает Заказ 4 на доставку провизии к самолету, у которого отправление в 14:00. Провизия должна быть загружена на борт не позднее, чем за полчаса до отправления, т.е. к 13:30. Планируя «Загрузку в самолет» длительностью 20 минут, агент операции ищет доступные ресурсы. Таким образом, появляется задача найти грузовик, доступный на интервале времени 13:10-13:30.

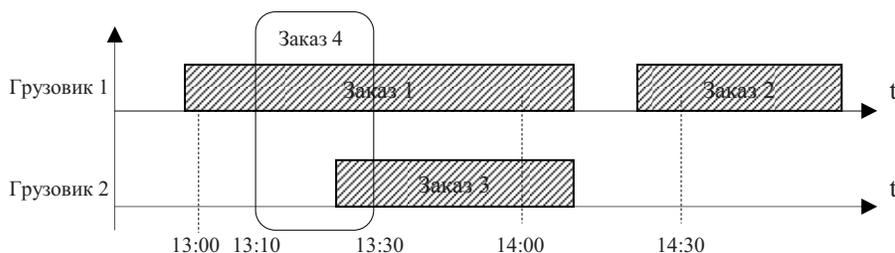


Рисунок 5 – Диаграмма Гантта для грузовиков до прихода нового заказа

Агент операции загрузки Заказа 4 не находит ни одного доступного на всем заданном интервале времени грузовика (грузовик 1 недоступен не всем интервале, грузовик 2 доступен лишь на части интервала времени). Возникает конфликтная ситуация, когда несколько заказов (1, 3 и 4) пытаются использовать одновременно один и тот же ресурс. Для разрешения этого конфликта агенты конкурирующих заказов начинают переговариваться между собой для поиска компромиссного решения. Агент Заказа 4 предлагает агентам Заказов 1 и 3 внести изменения в свои расписания и оценить стоимость такого изменения. При этом Заказ 1 оценивает стоимость такого изменения в 100 у.е., а Заказ 3 – в 60 у.е., поскольку для Заказа 3 изменение расписания не ведет к серьезным штрафам и т.д. Агент Заказа 4 сравнивает стоимости изменения расписаний и определяет, что выгоднее договориться с агентом Заказа 3 с целью внесения изменений в его расписание и использовать Грузовик 2. При этом агент Заказа 4 также сравнивает стоимость изменения графика Заказа 3 со стоимостью внесения изменения в собственное расписание. Если дешевле сдвинуть собственный Заказ (т.е. $\text{Стоимость}(\text{Заказ 3}) > \text{Стоимость}(\text{Заказ 4})$), то конфликт разрешается путем сдвига Заказа 4 влево по оси времени (при этом основное требование к заказу выполняется – он будет выполнен не позже чем за 30 минут до отправления рейса, см. рисунок 6. Если же дешевле внести изменения в расписание для Заказа 3 (т.е. $\text{Стоимость}(\text{Заказ 4}) > \text{Стоимость}(\text{Заказ 3})$), то Заказ 3 сдвигается вправо по числовой оси, освобождая Грузовик 2 (см. рисунок 7).

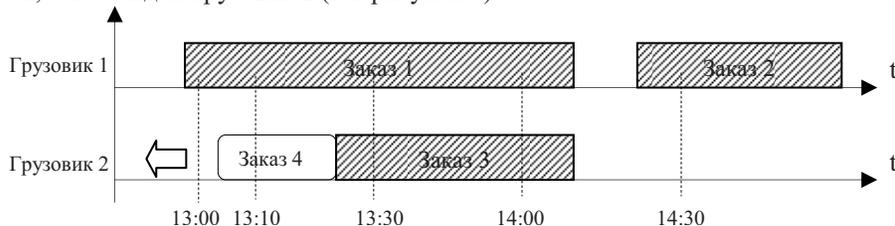


Рисунок 6 – Расписание после разрешения конфликта, случай $\text{Стоимость}(\text{Заказ 3}) > \text{Стоимость}(\text{Заказ 4})$

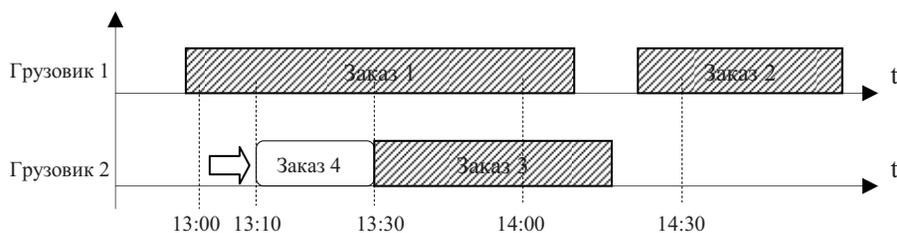


Рисунок 7– Расписание после разрешения конфликта, случай $\text{Стоимость(Заказ 4)} > \text{Стоимость(Заказ 3)}$

Описанная схема может быть применена итерационно, когда сдвиг одного заказа или операции приводит к волновому распространению возмущения (агенты многих заказов вынуждены перепланировать свои операции, искать новые ресурсы) или к «эффекту бабочки», когда незначительный сдвиг в одной части расписания приводит к непредсказуемым изменениям в другой его части. Все эти проблемы могут быть решены за конечное число шагов с использованием переговоров между агентами потребностей и возможностей в процессе поиска ими компромиссного решения. При этом основными критериями для поиска приемлемого решения являются показатели эффективности альтернативных вариантов, что заведомо приводит к решению, близкому к оптимальному.

Заключение

В статье предложен подход для динамического планирования деятельности наземных служб аэропорта, обеспечивающий новые возможности: реакция на внешние воздействия в режиме реального времени, событийно-ориентированная архитектура, позволяющая гармонично интегрироваться с различными сторонними программно-аппаратными комплексами, реализация концепции ПВ-сети, позволяющая получать приемлемые решения в сжатые временные сроки, выбирать между альтернативными решениями.

Результаты испытаний созданного прототипа системы на реальных данных аэропортов малых и средних размеров (Курумоч, Россия; Кельн-Бонн, Германия), а также аналитические расчеты подтвердили эффективность предлагаемого метода планирования как работоспособного, гибкого и адаптируемого механизма работы в сложных конфликтных ситуациях.

Таким образом, можно сделать вывод о перспективах разработанного подхода к динамическому распределению и планированию ресурсов при автоматизации сферы логистики пассажирской авиации.

Список литературы

- [1] Peter Jones. Flight Catering. 2004. ISBN-13: 978-0-7506-6216-1. ISBN-10: 0-7506-6216-6.
- [2] <http://www.groundstar.aero/>
- [3] Xiaoyu Mao, Nico Roos, Alfons Salden. Stable Multi-project Scheduling of Airport Ground Handling Services by Heterogeneous Agents. Материалы конференции «AAMAS 09» (Венгрия, Будапешт, 10-15 мая 2009 года) – с. 537 – 544.
- [4] Витих В. А., Скобелев П.О. Мультиагентные модели взаимодействий для построения сетей потребностей и возможностей в открытых системах // Автоматика и телемеханика, 2003. – №1. – с. 177-185.
- [5] Сюсин И.А., Царев А.В. Построение мультиагентной ПВ-сети для решения задач в сфере авиационной логистики. / Труды XI-ой международной конференции «Проблемы управления и моделирования в сложных системах». Самара: СНЦ РАН, 2009 – с. 624 – 633.

ПРИМЕНЕНИЕ КОНСТРУКТОРА ОНТОЛОГИИ В МУЛЬТИАГЕНТНОЙ СИСТЕМЕ ДИНАМИЧЕСКОГО ПЛАНИРОВАНИЯ ПЕРСОНАЛЬНЫХ ЗАДАЧ ДЛЯ МОБИЛЬНЫХ ПОЛЬЗОВАТЕЛЕЙ

М.В. Андреев¹, А.В. Иващенко¹, Д.М. Мартышкин², П.О. Скобелев², Л.В. Уланова¹, А.В. Царев¹

¹ГК «Генезис знаний»
443080, Самара, ул. Санфириковой 95, Россия
ivashenko@kg.ru, <http://www.kg.ru>
тел/факс: +7 (846) 332-21-01

²Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
petr.skobelev@gmail.com
тел: +7 (846) 332-39-27, факс: +7 (846) 333-27-70

Ключевые слова: онтология, семантическая сеть, мультиагентные системы, динамическое планирование, планирование персональных задач

Abstract

This paper describes a solution to develop knowledge management toolset in multi-agent system for dynamic scheduling of personal tasks.

Введение

В настоящее время весьма актуальной является разработка новых методов и средств динамического планирования персональных задач пользователей, позволяющих автоматически создавать и изменять формирующееся расписание мобильного пользователя по событиям с учётом структуры и взаимосвязей задач и возможности их исполнения в зависимости от местонахождения пользователя. Применение мультиагентных технологий для решения этой проблемы [1, 2] позволяют сделать работу существующих календарей интеллектуальной для обеспечения пользователей более удобными и полезными, гибкими и эффективными инструментами планирования.

Однако для того, чтобы добиться такого эффекта в мультиагентной системе, необходимо предусмотреть эффективные средства накопления знаний, которые бы позволили адекватно и точно описывать различные специфические типы задач в предлагаемой системе для различных предметных областей, таких как бизнес, спорт, отдых, культура и другие. Пользователь должен иметь возможность самостоятельно осуществлять настройку, изменение и добавление типов задач в соответствии с личными требованиями. При этом будут задаваться свойства и атрибуты задач, связи с другими задачами, совместимость/несовместимость друг с другом. Агенты разных типов операций смогут настраиваться на эти описания и из этих описаний знать точно, какое время требуется для операции, в какой цепочке она исполняется, когда она может быть начата и т.д.

Для эффективной реализации такой базы знаний в мультиагентной системе можно предложить онтологию, которая содержит базовые понятия и отношения между ними, представляемые в форме семантической сети. Редактор и средства хранения онтологий и правил [3, 4] могут использоваться для ведения и изменения общей онтологии различных задач, которая затем применяется в процессе планирования для описания связей и отношений между задачами. Редактор онтологий может позволить всем категориям пользователей самостоятельно опи-

сывать шаблоны задач и обеспечивать обмен опытом решения проблем путем передачи друг другу таких шаблонов.

В данной статье описываются результаты применения онтологий в мультиагентной системе планирования персональных задач, разработанной при выполнении работ по теме «Разработка мультиагентной системы динамического планирования персональных задач для мобильных пользователей» (шифр «2009-04-1.4-00-07-012»), выполняемой в рамках ФЦНТП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2012 годы».

1 Задача представления расписания пользователя в системе динамического планирования персональных задач

Расписание пользователя мультиагентной системы планирования персональных задач представляет собой одновременно его план и отчет об исполнении, т.е. является динамической, непрерывно «катящейся» (continuously rolling) структурой данных, поскольку построение и корректировка плана всегда сочетается с его исполнением в реальном времени. В этой связи, в ходе работы пользователя, к расписанию постоянно добавляются новые дни с новыми задачами и делами, а старые, отработанные, удаляются из расписания.

На рисунке 1 представлена структура расписания и для примера показано, что расписание пользователя в текущий момент представлено списком из трех дней, причем расписание на 17 мая самое заполненное и имеет 5 задач (операций). Здесь отношения показаны стрелками, причем, несмотря на однонаправленность показанных здесь стрелок, списки предполагаются дву- и более связные, соответствующие многоарным отношениям между задачами.

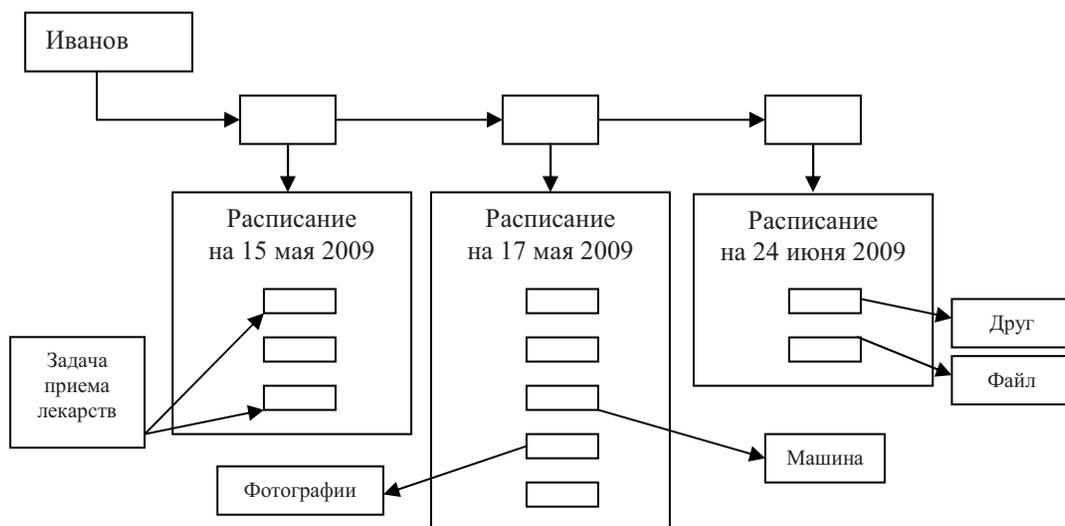


Рисунок 1 – Пример структуры расписания пользователя в отношении с другими объектами мира пользователя

Кроме того, часть операций может быть составной в расписании, представляя собой параллельно идущие нити работ пользователя, и раскрываться операциями следующего уровня.

В базе знаний расписание целесообразно хранить в виде семантической сети связей задач, операций, ресурсов и результатов пользователя, связывающей «все-со-всем», что позволяет видеть зависимость между операциями и разрешать конфликты в расписании.

Физически, сцена мира и расписание хранятся в памяти как многосвязные списки с индексами по объектам, отношениям и процессам (см. рисунок 2) – здесь, например, отношение

1 связывает объект 1 и объект N, а процесс N реализуется над объектом 2. В этой структуре данных каждый концепт, в свою очередь, описывается набором свойств и атрибутов с присущими им значениями, и его можно быстро найти как по индексу, так и последовательно перемещаясь от объекта – к отношению, от отношения к объекту, от объекта к процессу и т.д.

В качестве инструментария для работы со сценами в ходе проекта может быть предложен специальный отладчик, позволяющий в ручном режиме «пройти» сцену, красиво выполняя визуализацию сети и пошаговую навигацию от одного объекта к другому, а также осуществлять поиск объектов – независимо от приложения к планированию. Дополнительно в процессе работы может быть использована компонента, отображающая часть соответствующую сцены в виде расписания, в табличном виде, в виде списка с взаимосвязанными работами или в виде диаграммы Ганта.

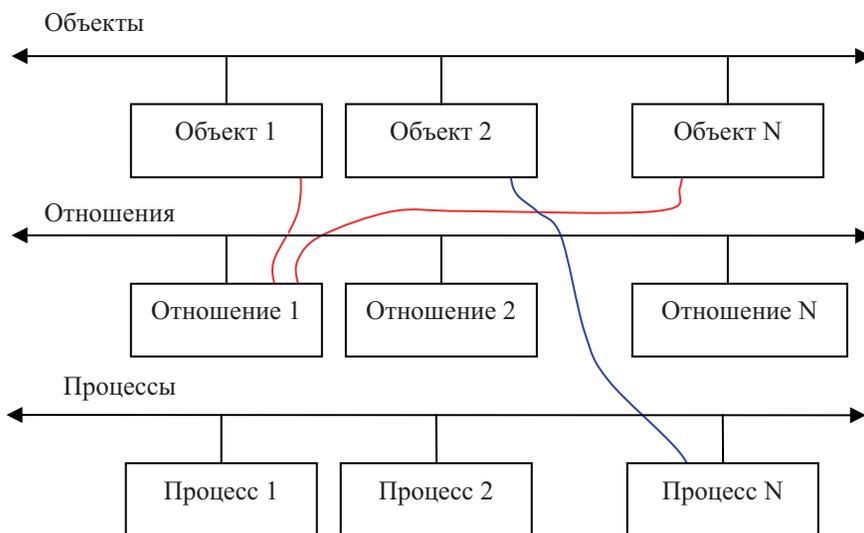


Рисунок 2 – Физическое представление сцены мира в памяти

2 Планирование персональных задач с использованием онтологических шаблонов

Одной из особенностей реализации системы планирования персональных задач является тот факт, что задачи в персональном расписании имеют сложную структуру и могут быть связаны различными отношениями друг с другом. При этом можно выделить некоторые типовые последовательности задач (которые в некоторых случаях могут быть достаточно сложными), которые разные пользователи планируют одинаково. Это могут быть хорошие решения сложных проблем (best practice), либо рекомендованные последовательности действий в определенных ситуациях. Для описания такой сложной последовательности задач в онтологии можно предложить такую формализацию, как онтологический шаблон.

Онтологические шаблоны могут создаваться пользователями, как для себя, так и для других пользователей в случае, когда они хотят поделиться опытом. После создания и редактирования шаблона, пользователь может его сохранить. Если предполагается что использование шаблона будет полезно и другим пользователям, данный шаблон может быть опубликован, то есть, открыт для свободного или платного доступа. У шаблона задается положительный и отрицательный рейтинг, с помощью которого обеспечивается ранжирование наиболее удачных шаблонов.

Работа с шаблонами может производиться на трех уровнях, соответствующих применению основных компонентов системы планирования персональных задач:

- рабочее место (компонент) для редактирования онтологии используется для создания, поиска, редактирования онтологических шаблонов, а также редактирования сцены, экземпляры которой связаны с пользователем;
- мобильный клиент для работы с личным расписанием используется пользователем для загрузки онтологических шаблонов и, при необходимости, их инстанциации, после чего задачи, описанные в шаблонах, отправляются на планирование в мультиагентную систему динамического планирования;
- мультиагентное ядро системы динамического планирования персональных задач применяется для составления расписания пользователя, при этом ограничения и взаимозависимости между задачами, описанные в шаблонах учитываются агентами в ходе планирования.

Для реализации данной концепции жизненный цикл шаблона может быть представлен в виде последовательности онтология – инстанциация – сцена – генерализация (OISG). Созданные пользователями онтологические шаблоны загружаются на мобильные устройства, на которых производится их инстанциация в сцену и последующее планирование задач, в результате которого задачи встают в расписание пользователя. В свою очередь, как на этапе планирования, так и в процессе выполнения существующих задач, пользователь может создавать и планировать дополнительные задачи, связывая их отношениями с теми, которые уже существуют.

При добавлении в расписание новых задач, пользователь может принять решение о необходимости обновления какого-либо существующего шаблона или создания нового. В результате анализа истории выполнения задач пользователем система планирования может заметить наиболее часто повторяемые паттерны в расписании и проактивно предложить пользователю выделить группу связанных задач в новый шаблон или дополнить/изменить существующий.

Преобразование шаблона в сцену (инстанциация шаблона) может осуществляться разными способами. Наиболее простым представляется вариант, когда для каждого объекта и процесса в шаблоне создается экземпляр в сцене. Если какие-либо концепты в онтологическом шаблоне были связаны друг с другом отношениями, то и в сцене соответствующие им экземпляры будут связаны друг с другом. Поясним это на примере. Допустим, есть некоторая последовательность связанных задач в личном расписании пользователя: после завтрака нужно выпить первое лекарство, перед обедом за 30 минут – второе, пообедать, а после обеда – опять первое. В таком шаблоне задачи «Выпить лекарство после завтрака» и «Выпить лекарство после обеда» должны ссылаться на одно и то же лекарство. В этом случае, на мобильном устройстве при инстанциации система попросит пользователя ввести параметры лекарства только один раз.

Инстанциацию можно рассматривать не только как преобразование части онтологии в группу экземпляров сцены, но и как встраивание шаблона в текущий контекст ситуации. При этом контекст можно рассматривать в двух аспектах:

- персональный профайл (семантический дескриптор);
- сложившаяся ситуация во всех деталях.

Персональный профайл используется для описания текущего состояния пользователя учета особенностей этого состояния в ходе динамического планирования. Например, контекст некоторого пользователя может содержать свойство «GPS координаты», отношение «Взял утром из гаража автомобиль» и процесс «Совершать переезд из А в В», что позволит планировщику понять, что по умолчанию при организации переездов данного пользователя следует планировать перемещения на машине и смотреть за пробками на дорогах. При этом планировщик может действовать по умолчанию или, наоборот, всякий раз при необходимости спрашивать подтверждения у пользователя по этому поводу (что регулируется настройками

пользователя). И, наоборот, если пользователь хочет бросить машину и поехать куда-то на такси или пойти пешком, он может изменить контекст, для чего будет достаточно запланировать событие «Пойти пешком» или «Поехать на такси». Таким образом, пользователь сообщит системе, что теперь выбирается другой способ перемещения, который перекрывает значение по умолчанию.

В качестве контекста «по умолчанию» для пользователя может задаваться следующая информация, уточняющая исходную сцену в любой рабочий день:

- утром пользователь находится у себя дома;
- состояние здоровья – хорошее;
- машина находится на стоянке (в гараже пользователя);
- выезд на работу обычно в 9.30;
- спортивные принадлежности находятся в машине.

Наличие машины изменяет алгоритм расчета времени переезда из одного места в другое, а плохое состояние здоровья может означать для планировщика пожелание пользователя сократить длительность рабочего дня.

Учет особенностей сложившейся ситуации (второй аспект рассмотрения контекста пользователя) позволяет адаптировать применение шаблона для конкретного расписания пользователя. Для этого при инстанциации задач шаблона производится анализ уже запланированных задач с целью поиска нужных корректировок. Например, в случае, если необходимо согласно одному шаблону принять лекарства до обеда, а согласно другому – после, при их совместном применении будет инстанцирован только один экземпляр задачи «Обед». В случае если в исходном шаблоне используется более общая задача «Прием пищи», при инстанциации пользователю будут предложены в качестве вариантов существующие в его расписании завтрак, деловой обед, торжественный ужин и т.п., а также создание новой задачи.

Использование онтологии при решении данной задачи позволяет осуществлять контекстно-зависимое планирование и выявлять конфликты. Например, пусть в шаблоне задано, что после приема пищи необходимо сразу принять лекарство. Однако при инстанциации пользователь указал в качестве приема пищи торжественный ужин в ресторане, посвященный годовщине свадьбы. Данный ужин был запланирован уже месяц назад и был заказан столик в ресторане. В редакторе онтологии было указано, что торжественный ужин связан отношением «Участвовать в процессе» с объектом «Красное вино. Бордо. 1986 г.». Точно такое же вино было открыто на первом свидании и осталось в памяти пары как символ начала их отношений.

В тоже время, «Красное вино. Бордо. 1986 г.» является разновидностью «Красное вино», что в свою очередь является «Алкогольным напитком». Необходимо учесть, что между объектом «Лекарство» и «Алкогольный напиток» задано отношение «Несовместимы». В этом случае при попытке размещения задач, входящих в шаблон по приему лекарств, в расписание, система выведет сообщение, что обнаружен конфликт с проактивными предложениями по его разрешению:

- перенести торжественный ужин на другое время/день, создав вместо ужина новый прием пищи или его вариант;
- перенести прием лекарства;
- отказаться от горячительных напитков;
- отказаться от приема лекарства, если срок приема подошел к концу.

3 Типизация отношений для описания сложных шаблонов последовательности задач

Для создания сложных сценариев действий и разрешения конфликтов между действиями в ходе динамического планирования в онтологиях задаются следующие временные отношения между действиями:

- Hard start – ЗАПУСК (D_1, T_1) – жесткий запуск действия D_1 в момент времени T_1 . Если на данный момент времени идет другое действие, то оно должно быть перепланировано (сдвинуто или перемещено) на другой срок;
- Soft start – МЗАПУСК (D_1, T_1) – мягкий запуск действия D_1 в момент времени T_1 . Если в данный момент времени запланировано другое действие, то необходимо решить, какое действие, имеющееся или новое, будет перепланировано (сдвинуто или перемещено) на другой срок;
- After – ПСЛЕД (D_1, D_T, D_2) - простое следование (ПС) действия 2 за действием 1, означающее, что как только D_1 завершится, следует через интервал времени D_T запустить D_2 (при этом T может быть функцией или константой);
- Synchronous – СИНХ ($D_1, D_2, \dots, D_n; D_T, D_{n+1}$) – синхронное следование (СС) действия D_{n+1} после завершения всех действий D_1, D_2, \dots, D_n через интервал D_T ;
- Asynchronous – АСИНХ ($D_1, D_2, \dots, D_n; D_T, D_{n+1}$) – асинхронное следование (СС) действия D_{n+1} после завершения хотя бы одного из действий D_1, D_2, \dots, D_n через интервал D_T ;
- Before – ДО (D_1, T, D_2) - запланировать действие D_1 так, чтобы оно успело выполниться не позже чем за время T до момента начала D_2 (планировать от D_2 к D_1);
- Concurrent – ПАР (D_1) – параллельное исполнение действия D_1 по отношению к любым другим действиям в расписании пользователя;
- Open slot – СВОБ (D_1) – запускать действие D_1 только в свободное время.
- Alternative before – АЛЬТДО ($D_1, D_2, \dots, D_n; D_T, D_{\text{опт}}; D_{n+1}$) запланировать действия D_1, D_2, \dots, D_n так, чтобы они успели выполниться не позже чем за время T до момента начала D_{n+1} (планировать от D_{n+1} к D_1, D_2, \dots, D_n). $D_{\text{опт}}$ – означает необязательность выполнения указанных задач D_1, D_2, \dots, D_n .

Знание временных отношений между действиями позволяет системе согласованно перепланировать время выполнения этих действий в случае возникновения различного рода событий. Действия могут запускаться по событиям:

- наступления заданного абсолютного момента времени или истечения относительного интервала времени;
- начала или завершения выбранного действия (например, завтрака);
- поступления приглашения на встречу или Conf Call, электронного письма или СМС сообщения от выбранного отправителя на заданную тему или по заданному имени вложенного файла;
- завершения звонка от выбранного абонента.

4 Описание действий в онтологиях пользователей

Для динамического планировщика наиболее важны описания действий в онтологиях, между которыми могут быть определены специальные отношения следования. При выборе и загрузке онтологии планировщику становится доступен весь список действий, определенных для объектов данной онтологии. Такой список действий соответственно открывается и пользователю, который может выбирать задачи для планирования, причем параметры этих задач настраиваются по ситуации на основе контекста пользователя. Рассмотрим несколько примеров таких действий, как элементарных процессов, так и составных, которые могут осуществляться над объектами онтологии.

Первый пример посвящен онтологии автомобиля – здесь предполагаются наличие определенных свойств объекта машины, например, габаритов, расхода топлива на 100 км пути, общего объема и стоимости сожженного топлива и т.д. Для простоты в данном примере не будем рассматривать отношения временного следования между задачами (например, отдать машину в ремонт и потом забрать из ремонта, забрать из гаража и поставить в гараж и т.п.).

Рассмотрим возможные описания данных задач в расчете на использование в мультиагентной системе планирования.

- «Взять машину из гаража» – задача, которая может задаваться пользователем вручную или инстанцироваться из типового сценария действий пользователя в течение дня. При этом система знает, где находится пользователь и где находится гараж (стоянка), и может примерно оценить требуемое время, нужное, чтобы добраться до гаража пешком. Параметры задачи: место нахождения гаража, как добраться (пешком или на транспорте), среднее время в пути или расчет по ситуации, затраченное время на доставку, время выезда из гаража.
- «Поставить машину в гараж» – задача, которая может задаваться пользователем вручную или инстанцироваться из типового сценария действий пользователя в течение дня. При этом система знает, где находится пользователь и где находится гараж (стоянка), и может примерно оценить требуемое время доезда до гаража. Параметры задачи: место нахождения гаража, среднее время в пути или расчет по ситуации, затраченное время на доставку, время приезда в гараж.
- «Заправить машину» – задача, которая может задаваться вручную или возникать периодически в расписании пользователя в тех случаях, когда он часто использует машину. Если приложение планирует поездки с использованием электронной карты, то довольно точно будет известен пройденный километраж и время предыдущей заправки. В этом случае можно разумно предупреждать пользователя о необходимости заправиться и предлагать спланировать такое действие, в особенности, если пользователь в текущее время находится на машине, причем в его расписании обнаруживается свободное время, а рядом есть заправка, которую он постоянно использует. Также может быть обеспечена возможность активации данной задачи по сигналу от датчика расхода или уровня топлива в баке автомобиля. Параметрами данной задачи могут быть: домашняя АЗС (по умолчанию), условие активации (время от времени по заказу пользователя или при критическом уменьшении запаса топлива в баке), время и место последней заправки, стоимость бензина при последней заправке, текущий и ожидаемый остаток топлива в баке, реальный остаток (для сравнения и корректировки расхода), объем и цена новой заправки, время заправки, включая ожидание в очереди.
- «Помыть машину» – задача, которая может задаваться вручную или возникать периодически в расписании пользователя в тех случаях, когда он часто использует машину или по событию от сервера погоды, указывающего на дождь – в случае поездок или нахождения машины на открытой стоянке во время дождя. Параметрами этой задачи могут являться следующие: домашняя мойка (по умолчанию), условие активации – время от времени по заказу пользователя или по наличию поездок во время дождя, время и место последней мойки, стоимость последней мойки, объем и цена новой мойки, время новой мойки, включая ожидание в очереди.
- «Отдать машину в ремонт» – задача, которая может задаваться вручную или возникать из списка задач пользователя (ToDoList). Параметры задачи: домашний сервис (по умолчанию), предмет, срок и стоимость последнего ремонта, затраченное время, чтобы сдать машину в ремонт. Эта задача инициирует появление в расписании последующей задачи забрать машину из ремонта – по назначенному пользователю сроку.
- «Забрать машину из ремонта» – задача, которая порождается задачей сдачи машины в ремонт. Параметрами задачи являются предмет, срок и стоимость нового ремонта, а также время, затраченное чтобы забрать машину в сервисе.
- «Пройти техобслуживание» – задача, которая должна выполняться раз в три года, после чего заново планирует себя на следующий период времени.

- «Купить машину» – устанавливает отношение наличия движимого имущества между пользователем и машиной, задает все параметры машины, позволяет далее активизировать все сценарии по работе с машиной.
- «Продать машину» – разрывает отношение наличия движимого имущества между пользователем и машиной, уничтожает все параметры машины и не позволяет активировать сценарии работы с машиной, для которых необходимо наличие имущественных отношений с машиной.

Аналогично расписаны действия, типичные для процесса приема лекарств, что позволит планировщику помогать пожилым людям (см. рисунок 3). Рассмотрим в качестве примера возможное применение задачи по приему лекарства. Пользователь может вводить такую задачу вручную, или она может автоматически планироваться до или после приема пищи, например, через 1 час после еды. Если вдруг в расписание пользователя никакая еда не задана (концепт «Прием пищи» или его подклассы вида «Завтрак»), то система попросит указать время, когда назначить прием лекарства или сама создаст задачу «Прием пищи», от которой (или до которой) будет отсчитываться время динамически с запросом пользователю на подтверждение завершения приема пищи.

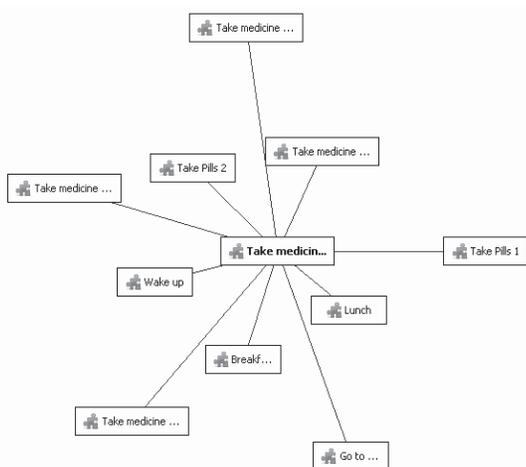


Рисунок 3 – Фрагмент онтологии приема лекарства

Предположим, что прием лекарств состоит из нескольких связанных задач. Например, одно лекарство необходимо принять за полчаса до еды, а два других – спустя час после еды. Для создания такой цепочки связанных задач целесообразно создать шаблон, который можно будет неоднократно использовать, в частности открыть для скачивания другим пользователям. Прием лекарств может быть и более сложным: принять лекарство после сна, позавтракать, принять лекарство после завтрака, принять лекарство перед обедом, пообедать, принять два лекарства после обеда, принять лекарство перед сном.

Список литературы

- [1] Иващенко А.В., Мартышкин Д.М., Скобелев П.О., Уланова Л.В., Царев А.В. Построение мультиагентной системы динамического планирования персональных задач для мобильных пользователей / Вестник Самарского государственного технического университета. Серия «Технические науки», 2009, № 3 – с. 32 – 37.
- [2] Виттих В.А., Скобелев П.О. Метод сопряженных взаимодействий для управления распределением ресурсов в реальном масштабе времени / Автометрия. – 2009. – № 2. – С. 78 – 87.
- [3] Андреев В.В., и др. Конструктор онтологий мультиагентных систем // Труды III Международной конференции по проблемам управления и моделирования сложных систем, Самара, 4-9 сентября 2001 – Самара: СНЦ РАН. 2001. – с. 480 – 488.
- [4] Абрамов Д.В. и др. Разработка средств построения и использования онтологий для поддержки принятия решений. // Труды VII Международной конференции по проблемам управления и моделирования сложных систем – Самара: СНЦ РАН. 2005. – с. 420 – 427.

ОРГАНИЗАЦИЯ ИНТЕРАКТИВНОГО ВЗАИМОДЕЙСТВИЯ В МУЛЬТИАГЕНТНОЙ СИСТЕМЕ УПРАВЛЕНИЯ ТРАНСПОРТНО-ЭКСПЕДИЦИОННОЙ КОМПАНИЕЙ

А.В. Вайсблат¹, А.Р. Диязитдинова¹, А.В. Ивашенко², П.О. Скобелев¹, А.В. Царев¹

¹ГК «Генезис знаний»

443080, Самара, ул. Санфириковой 95 «А», Россия

at@kg.ru, <http://www.kg.ru>

тел./факс: +7 (846) 222-91-72, 222-91-73

²Институт проблем управления сложными системами РАН

443020, Самара, ул. Садовая, 61, Россия

ivashenko@kg.ru

тел./факс: +7 (846) 333-27-70

***Ключевые слова:** распределение ресурсов, планирование грузоперевозок, динамическое планирование, мультиагентные системы, распределенное принятие решений, онтология*

Abstract

This paper describes the result of implementation and first probation of full truck load (FTL) scheduling system based on multi-agent platform in transportation logistics. Basic attention is given to providing interactive cooperation of all system users which are operators, dispatchers, drivers (via mobile handheld devices) and top management to involve them into collaborative decision making process.

Введение

Продолжение работ по созданию новой системы автоматизированного управления мобильными ресурсами на основе применения мультиагентных технологий, обозначенных в работе [1], привело к необходимости решения ряда актуальных задач в области применения современных принципов управления транспортно-экспедиционной компанией на основе мультиагентных решений. В частности, весьма важно решение проблем внедрения созданных компонентов мультиагентной платформы, обеспечения бесперебойной работы мультиагентной системы управления и адекватности результатов планирования, а также объяснения их конечным пользователям, обучения пользователей, перехода от ручного планирования к автоматизированному и проработки пользовательского интерфейса.

Однако наиболее сложной и актуальной является задача организации интерактивного взаимодействия операторов, диспетчеров, водителей и руководства компании по согласованному планированию ресурсов. Новые информационные технологии, обеспечивающие постоянную связь с помощью мобильных устройств (так называемых, «наладонников»), возможность задания индивидуальной логики принятия решений средствами онтологий, а также современные методы организации сопряженного взаимодействия [2] представляют собой мощный инструмент организации совместной деятельности пользователей по управлению мобильными ресурсами. Вместе с тем, решение этой задачи на практике требует дополнительных усилий по доработке пользовательского интерфейса и бизнес-процесса планирования, для того, чтобы все возможности новых технологий стали доступны и удобны в использовании.

Кроме этого, весьма важно настроить логику агентов, правильно выбрав цели и ограничения каждого агента, а также задав критерии оптимизации индивидуальных планов. Эти критерии могут быть разными даже в случае внедрения мультиагентной системы управления мобильными ресурсами в одной области: например, в двух разных транспортно-экспедиционных компаниях могут использовать разные критерии, а существующие могут быть определены с

разными приоритетами. Эта разница, в основном, определяется спецификой постоянных заказчиков и особенностями используемых ресурсов, составляющих флот компании. Например, в случае преобладания заказов из столицы в регионы, бывает выгоднее поступиться выручкой от каждой поездки, но быстрее вернуть грузовик в столицу, что позволит выполнить большее количество заказов.

В данной статье основное внимание уделено именно таким нюансам: описаны особенности организации процесса обработки заявок в транспортно-экспедиционной компании, сформулированы возможные критерии автоматизированного планирования заявок и приведены примеры экранных форм, иллюстрирующие базовые сценарии планирования поступающих заявок, описана организация согласованного взаимодействия по планированию транспортных ресурсов.

1 Особенности планирования ресурсов в транспортно-экспедиционной компании

Задачи транспортной логистики заключаются в сокращении сроков транспортировки грузов и оптимизации затрат компании на всем пути движения грузов при максимально высоком уровне обслуживания посредников и конечных потребителей. В коммерческих транспортных организациях, где имеется распределенная структура с развитой филиальной сетью и используется большое количество подвижного грузового транспорта, задача по автоматизации планирования поездок, управлению подвижными ресурсами, оптимизации материальных и временных затрат на перевозку, контролю и учету операций бизнес-процесса приема и исполнения заявок, а также повышению качества оказываемых услуг в целом становится особенно актуальной с ростом числа клиентов и заказов. Анализ технического оснащения современных транспортно-экспедиционных компаний показывает, что такие компании представляют собой сложные объекты управления, которые характеризуются значительным числом взаимозависимых ограничений и факторов, обусловленных технологическим регламентом и определяющих эффективность работы объекта управления, что вызывает определенные трудности при решении задачи планирования выполнения заявки на транспортировку.

В процессе оказания услуги по транспортировке грузов задействовано три стороны: заказчик (грузоотправитель), транспортно-экспедиционная компания (ТЭК) и грузополучатель. Каждая из указанных сторон может иметь свои собственные, иногда противоречащие друг другу, цели. На рисунке 1 показана упрощенная схема осуществления транспортировки груза, на которой представлены основные аспекты взаимодействия сторон, между которыми необходимо обеспечить информационный обмен для эффективного управления процессом транспортировки. Наличие некоторого центра, выполняющего функции диспетчера сквозного перевозочного процесса, осуществляющего единую функцию управления, дает возможность эффективно проектировать движение материальных и информационных потоков. Кроме того, качественное управление возможно только при наличии системы поддержки принятия решений, которая будет учитывать многокритериальность и неопределенность исходной информации, а также будет обеспечивать автоматизацию процесса поиска наилучших вариантов. При автоматизации процесса управления транспортными ресурсами необходимо учитывать такие факторы, как неравномерность поступления заявок на транспортировку и транспортных средств, возможные форс-мажорные обстоятельства при выполнении заявки (например, поломка транспортного средства), неравномерность загрузки транспортных средств, переменный уровень эксплуатационной надежности и т.п.

Можно выделить следующие задачи управления, решение которых имеет определяющий характер для компаний, занимающихся оказанием транспортно-логистических услуг:

- организация процесса планирования ресурсов на выполнение заказа с учетом особенностей деятельности грузоотправителя/грузополучателя (например, режима работы складов и выходных);

- обеспечение обратной связи с заказчиком (грузоотправителем) и двусторонней связи с водителем для получения актуальной информации о состоянии выполнения заказа, а также возможных проблемных ситуациях.

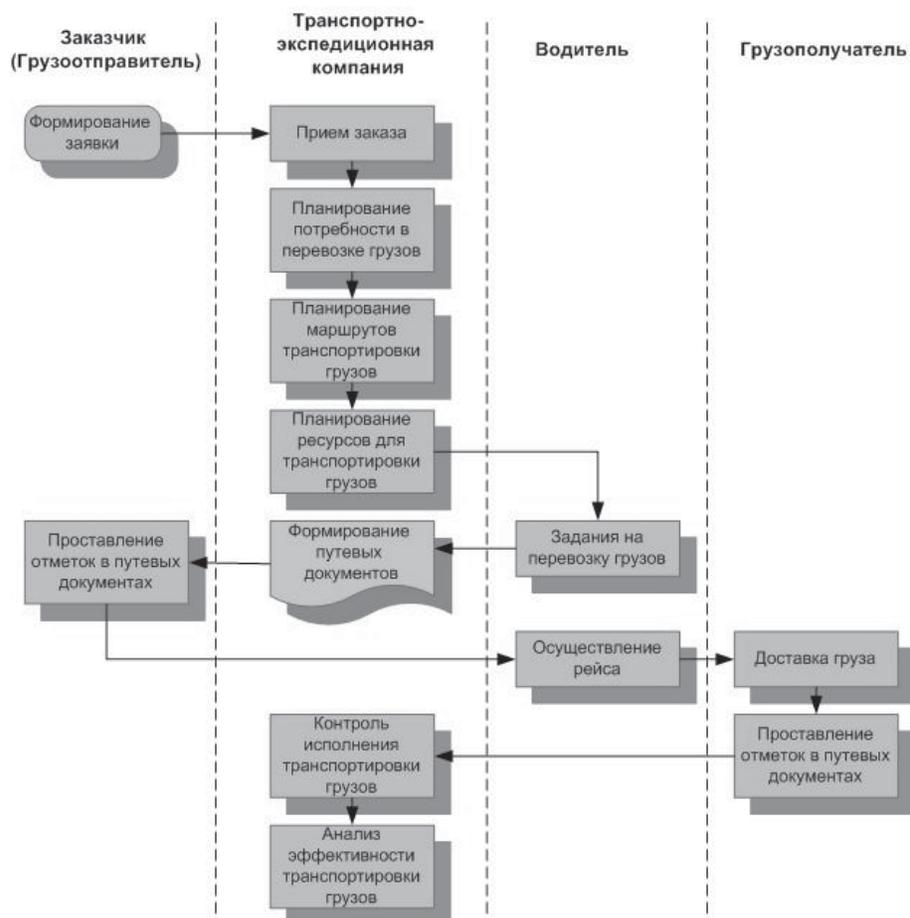


Рисунок 1 – Схема взаимодействия по обработке заявки

Собственно при планировании средствами автоматизированной интеллектуальной системы необходимо учитывать следующие критерии:

- Утилизация – отношение времени, потраченного на перемещение до места загрузки и выполнение заявки, к общему рабочему времени в месяц. Необходимо, чтобы грузовики не простаивали и постоянно ездили по заявкам (утилизацию надо постоянно повышать);
- Абсолютная выручка – выручка, получаемая при выполнении заявки. Необходимо, чтобы выручка была максимальной (автоматизированная система должна отдавать предпочтение заявкам с максимальной абсолютной выручкой).
- Удельная выручка – выручка на километр. Система должна оценивать заявки по удельной выручке и сравнивать значение удельной выручки с некоторой граничной величиной. Среди заявок, с удельной выручкой, больше заданной границы, будет выбрана заявка с большей абсолютной выручкой.

Среди заявок, приблизительно равных по абсолютной выручке необходимо выбирать те, у которых удельная выручка больше. Из двух заявок, удельная выручка которых больше заданной границы, следует выбирать ту, которая выгоднее по абсолютной выручке. Если, со-

гласно данным, указанным в заявке, один грузовик вынужден будет совершить более протяженный переезд до точки погрузки, чем другой, следует выбрать грузовик, находящийся ближе с целью уменьшения количества «порожних» переездов.

Кроме этого, надо учитывать специальные правила: так как заявки из крупного транспортно-портного узла в другие города найти легче, необходимо планировать возврат в транспортный центр как можно быстрее, пусть и с меньшей выручкой. При этом не стоит ждать, пока появятся более выгодные заявки или выполнять заявки, требующие перемещения в другой город. Например, если одна из двух заявок из Самары в Москву является более выгодной по выручке, чем другая, но при этом погрузка более выгодной заявки должна начаться на день позже, то следует выбрать менее выгодную заявку, позволив грузовику вернуться в Москву раньше.

2 Построение системы управления транспортно-экспедиционной компанией на базе мультиагентной платформы планирования мобильных ресурсов

Ключевым этапом в общем процессе обработки заявок, поступающих в транспортно-экспедиционную компанию, является поиск ресурсов для их выполнения. С учетом описанных особенностей рассматриваемого бизнеса при автоматизации этого этапа целесообразно использование компонентов мультиагентной платформы [1], которые позволяют добиться следующих преимуществ:

- обеспечение индивидуального подхода к выполнению каждого заказа;
- автоматизация основных операций операторов по планированию ресурсов;
- накопление знаний о специфике бизнеса, позволяющих уточнять и совершенствовать правила принятия решений при планировании;
- повышение оперативности принятия решений и сокращение времени планирования;
- осуществление эффективного наглядного оперативного управления имеющимися заказами, повышение операционной эффективности;
- своевременное извещение операторов о возможных проблемах (отсутствие заказов на будущее, расхождение между «планом» и «фактом», неэффективное использование транспортных средств и т.д.);
- получение соответствующей комплексной систематизированной отчетности, снижение издержек по составлению и систематизации отчетной документации.

Основной задачей мультиагентной системы в рассматриваемой области является построение и поддержание баланса интересов заказчиков, самой компании, водителей и других участников процесса принятия решений. При этом каждый участник взаимодействия может иметь свои потребности и возможности, реализуемые через роли заказов и ресурсов участников, постоянно стремящихся найти друг друга и установить связи между собой, бронируя ресурсы под заказы.

Основные отличия выбранного решения, основанного на использовании мультиагентных технологий, от существующих подходов по разработке и внедрению систем управления транспортом состоят в следующем. Функциональность этих систем учетной направленности заключается в оперативной выдаче сведений о местонахождении груза и сроках его доставки по запросу менеджера. При решении задача подбора «оптимального» ресурса для выполнения каждой заявки традиционные системы не могут учесть всех реально существующих параметров и требований, которые налагают на них грузоотправители, грузополучатели, особенностей загрузки транспортных средств и реальной пропускной способности дорог. По мере возрастания числа бизнес ограничений (график работы объектов, характеристики транспортных средств, маршрутов и т.д.) уменьшается наглядность и прозрачность схемы взаимодействия, и, как следствие, выбор оптимального решения становится крайне сложным.

Одним из известных достоинств распространенных систем управления транспортом являются развитые средства сценарного анализа «что если», которые позволяют по различным

критериям проанализировать различные сценарии грузоперевозок и принять экономически обоснованное решение по лучшему из них. Но основное решение остается за менеджером, что не снимает субъективности логики принятия решения по каждому конкретному заказу. Мультиагентная система не только позволяет моделировать сценарные варианты «что будет, если...», но и «подсказывать» пользователю действительно оптимальные варианты планирования транспортных ресурсов, а для большей части заказов производить планирование в автоматическом режиме, что приводит к повышению объективности принимаемых решений. Другим важным преимуществом мультиагентной системы является возможность реагировать на появляющиеся в реальном времени нештатные или чрезвычайные ситуации (например, при поломке транспортного средства или задержке при выполнении заявки, происходит автоматическое перепланирование всего расписания), в то время как классические системы данного направления лишь позволяют оценить неблагоприятность той или иной ситуации.

Рассмотрим, как функционирует мультиагентная система планирования ресурсов в транспортно-экспедиционной компании на примере планирования новой заявки. При заполнении регистрационной карточки заявки (см. рисунок 2) диспетчер обязан ввести обязательные данные, значения которые в дальнейшем будут учтены при планировании.

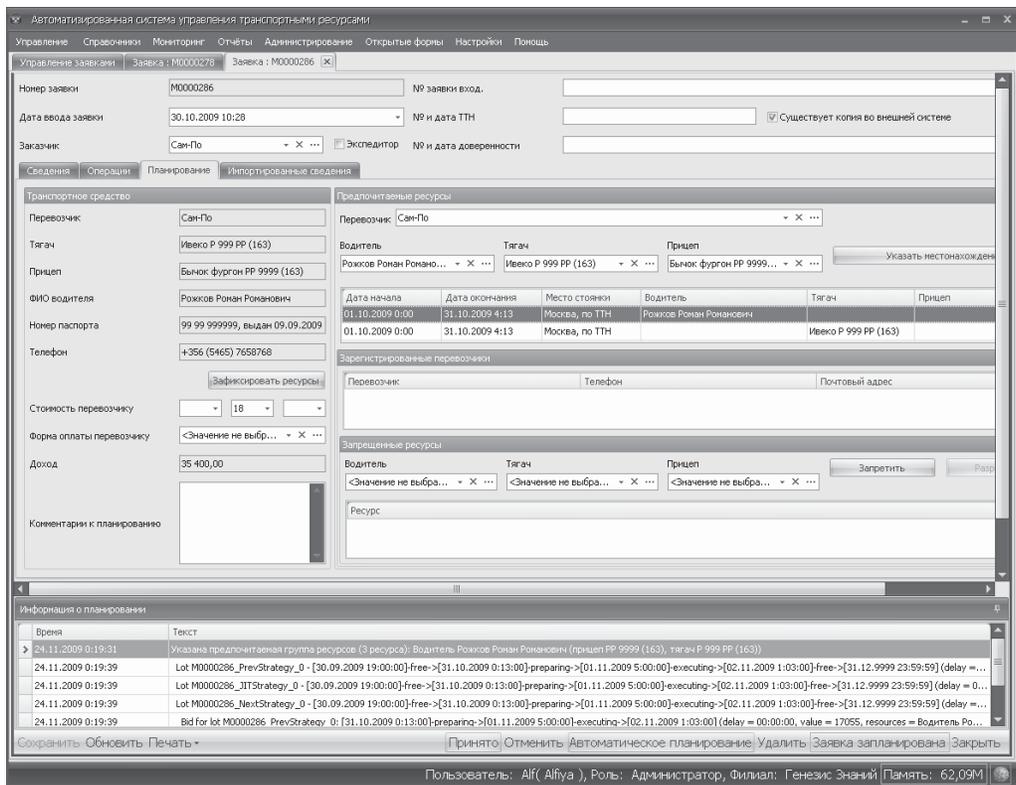


Рисунок 2 – Регистрационная карточка заявки с «подобранными» ресурсами

Помимо заявок в системе хранятся сведения о транспортных ресурсах (как собственных, так и привлеченных) с указанием их основных характеристик и текущего местоположения. На основании данных сведений мультиагентное ядро системы производит подбор наиболее подходящих ресурсов. Планирование осуществляется не путем централизованного и последовательного комбинаторного перебора возможных вариантов и выбора наилучшего из них, а на основе применения распределенного подхода, в котором одновременно и параллельно работа-

ет множество агентов с собственными интересами и критериями принятия решений. В ходе процесса переговоров агентов строится квази-оптимальный, сбалансированный по различным критериям план транспортировки с учетом индивидуальных ограничений и предпочтений.

Поскольку в этом случае весьма важно реализовать механизм объяснения принимаемых решений, в системе реализована возможность просмотра лога переговоров агентов, что позволяет пользователю выяснить, почему тот или иной ресурс был назначен на выполнение конкретной заявки. Кроме того, в системе предоставлена возможность просмотра занятости ресурса того или иного типа (транспортного средства или водителя) (см. рисунок 3).

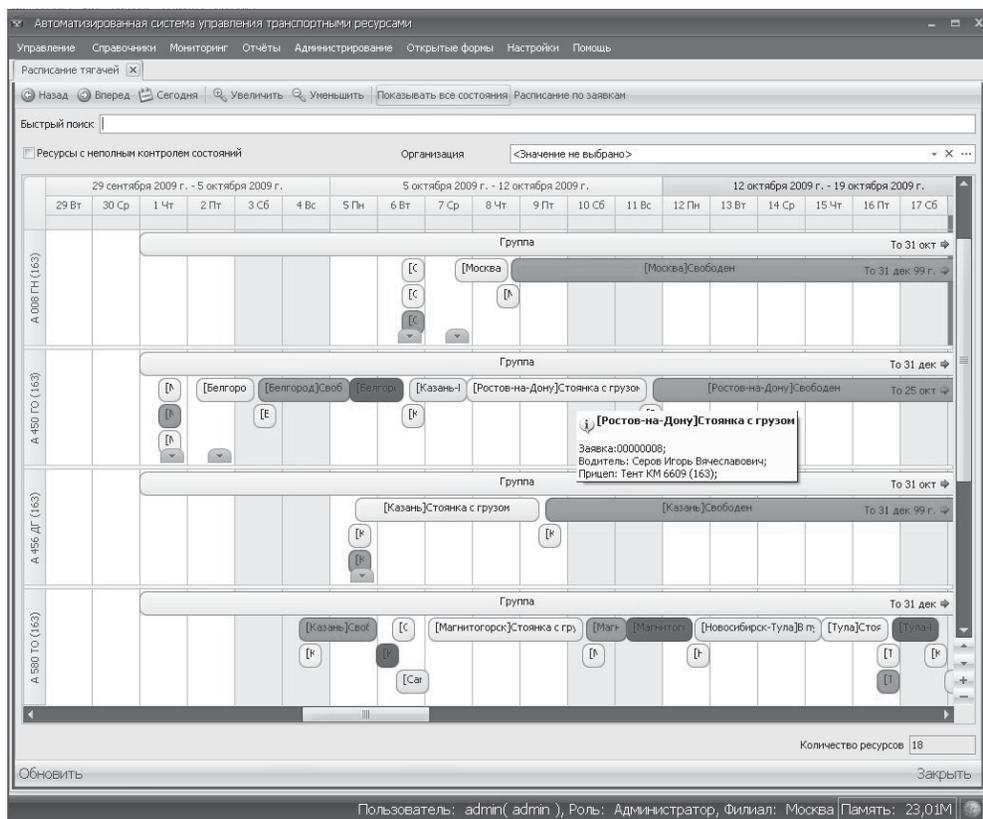


Рисунок 3 – Расписание транспортных ресурсов

3 Организация взаимодействия пользователей мультиагентной системы управления транспортными ресурсами

Использование современных мультиагентных технологий для адаптивного управления транспортно-экспедиционной компанией связано с обеспечением интерактивного взаимодействия менеджеров компании, диспетчеров и водителей. Для этого необходимо, во-первых, реализовать связь между всеми участниками взаимодействия с помощью GPS и мобильных устройств связи, а, во-вторых, предусмотреть распределенную поддержку принятия решений в мультиагентной системе управления транспортными ресурсами.

На рисунке 4 представлена физическая архитектура системы. Заказчик (грузоотправитель) через Интернет-портал (также могут использоваться и другие средства связи) формирует заявку на перевозку груза. Данная заявка поступает в систему, где на основании сведений об имеющихся транспортных ресурсах, их загруженности, местоположении и прочих данных

осуществляется подбор и ранжирование ресурсов для выполнения заявки. Первому в этом списке (т.е. самому подходящему) из списка выбранных водителей на терминал, установленном в кабине тягача, присылается сообщение с информацией о заявке (адрес погрузки/разгрузки, дата, время), после подтверждения приема заявки по GPS-передатчику осуществляется отслеживание местоположения водителя.

Актуальную информацию о состоянии заявки (на погрузке, на разгрузке, транспортируется и пр.) заказчик может отследить также через Интернет-портал. Эти же данные поступают с системы финансового учета, установленные на предприятии (например, продукты компании 1С) и используются для получения разнообразных отчетов.

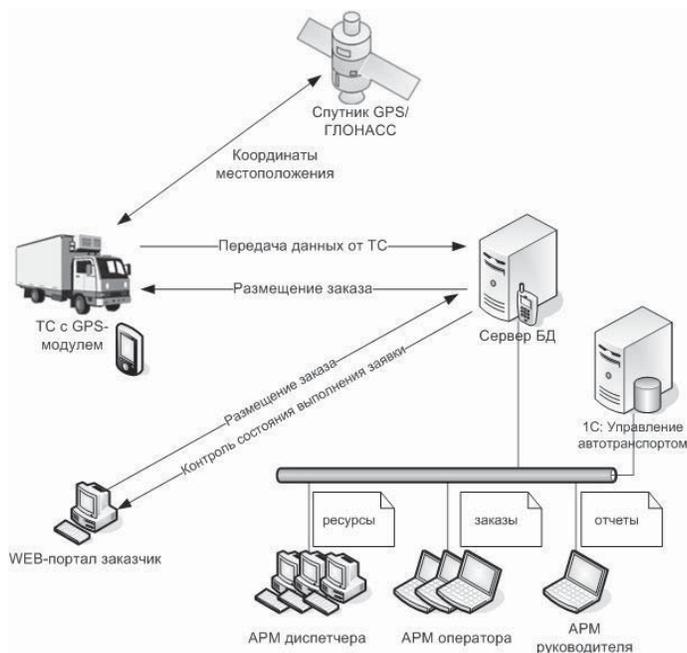


Рисунок 4 – Диаграмма развертывания системы

Помимо планирования ресурсов другой важной задачей является осуществление мониторинга исполнения заявки, включающего контроль местоположения транспортного средства и состояния груза в процессе перевозки. В системе реализована возможность отслеживания тягача по данным GPS-навигатора: диспетчер может просмотреть как историю перемещения тягача, так и запланированные «будущие» перемещения при выполнении заявок. Для организации обратной связи с водителями предлагается использование специального терминала, устанавливаемого в кабине тягача, с помощью которого водитель может информировать систему о своем актуальном статусе, а также осуществлять взаимодействие с диспетчером при получении нового заказа и в случае возникновения непредвиденных ситуаций.

Интерфейс мобильного приложения содержит:

- табличную часть, в которой приводится детальная информация об операциях, запланированных на данного водителя (тип операции (погрузка/разгрузка), дата и время прибытия, адрес погрузки/разгрузки, фамилия диспетчера, передавшего заказ);
- кнопки, предназначенные для регистрации водителем своего статуса по контрольным точкам исполнения заявки (прибыл на погрузку/разгрузку, закончил погрузку/разгрузку);
- кнопки детализации данных по операции, предназначенные для получения в случае необходимости подробных сведений для осуществления операции;
- тревожную кнопку для связи с диспетчером.

Последовательность действий водителя при работе с мобильным приложением выглядит следующим образом:

- при выходе на линию водитель активизирует данное устройство, при этом он становится виден в системе как свободный ресурс;
- при поступлении новой заявки, назначенной на данного водителя, данные заявки отображаются в табличной части;
- при начале выполнения заявки водитель путем нажатия кнопки «выехал», фиксирует собственное начало выполнения заявки;
- при прибытии на место погрузки водитель нажимает на кнопку «прибыл»; о начале погрузки водитель информирует диспетчера нажатием на кнопку «начал погрузку» и т.д. Организация взаимодействия с заказчиками осуществляется по двум направлениям:
- оперативное получение заявки без возможного искажения приведенных сведений;
- информирование клиента (грузоотправителя) о текущем статусе выполнения заявки.

Для обеспечения оперативного взаимодействия с заказчиком можно предложить создание специализированного Интернет-портала, с помощью которого заказчик сможет как регистрировать заявку, так и отслеживать изменение ее состояния. При размещении заявки грузоотправитель указывает такие параметры заявки как наименование, место погрузки, дата и время погрузки, требуемый тип погрузки, вес, количество и т.п. На основании введенных данных система управления транспортными ресурсами подберет наиболее подходящее транспортное средство. После поступления заявки в базу данных системы ей присваивается уникальный номер, который пересылается на e-mail заказчика вместе с уведомлением о приеме заявки к планированию. Также на портале предусмотрена возможность просмотра текущей стадии выполнения заявки, для чего зарегистрированному заказчику необходимо указать номер интересующей его заявки в соответствующей строке поиска. Данные могут обновляться в реальном времени на основании сведений, поступающих с терминала водителя.

Заключение

В данной статье приводится описание результатов построения автоматизированной системы управления распределением мобильных ресурсов на базе мультиагентной платформы. Полученные результаты позволяют сделать вывод о необходимости совершенствования организации согласованного взаимодействия лиц, принимающих решения для обеспечения наиболее эффективной работы мультиагентной системы и своевременной обработки событий, касающихся процесса обработки заявок и происходящих в реальном времени.

К основным преимуществам, достижению которых будет способствовать внедрение и эксплуатация мультиагентной системы управления транспортными ресурсами, можно отнести улучшение показателей эффективности работы транспортно-экспедиционной компании за счет автоматического контроля местонахождения и автоматического контроля соблюдения расписания, ускорение процессов долгосрочного и краткосрочного планирования перевозок, контроль работы привлеченных транспортных ресурсов, возможность получения оперативного среза работы ТС на любой момент времени и снижение потерь рабочего времени и информации при взаимодействии служб и сотрудников.

Список литературы

- [1] Дязитдинова А.Р., Иващенко А.В., Мартышкин Д.М., Скобелев П.О., Сурнин О.Л., Царев А.В. Разработка мультиагентной платформы для планирования в сфере транспортной логистики // Труды XI Международной конференции «Проблемы управления и моделирования в сложных системах» Самара: Самарский научный центр РАН, 2009. – с. 608 – 623.
- [2] Виттих В.А., Скобелев П.О. Метод сопряженных взаимодействий для управления распределением ресурсов в реальном масштабе времени / Автометрия. – 2009. – № 2. – с. 78 – 87.

КОНЦЕПЦИЯ МУЛЬТИАГЕНТНОЙ СИСТЕМЫ ИНТЕРАКТИВНОГО ПОСТРОЕНИЯ ПРОГРАММЫ ПОЛЕТА И ПЛАНИРОВАНИЯ ГРУЗОПОТОКА МЕЖДУНАРОДНОЙ КОСМИЧЕСКОЙ СТАНЦИИ

А.Р. Диязитдинова¹, А.В. Иващенко¹, Е.В. Симонова¹,
П.О. Скобелев¹, М.В. Сычева², И.И. Хамиц², А.В. Царев¹

¹ГК «Генезис знаний»
443080, Самара, ул. Санфириковой 95, Россия
ivashenko@kg.ru, <http://www.kg.ru>
тел/факс: +7 (846) 332-21-01

²ОАО РКК «Энергия»
141070, Московская область, Королев, ул. Ленина, 4-а, Россия

Ключевые слова: планирование грузопотока, программа полета, МКС, распределение ресурсов, динамическое планирование, мультиагентная технология, онтология

Abstract

This paper describes a vision of multi-agent technology application to build a solution of International Space Station flight program, cargo flow and resources optimization scheduling. New principles are provided of how to organize an interaction of decision makers on the basis of ontology and multi-agent platform.

Введение

Сегодня можно привести достаточно большое количество примеров успешного применения мультиагентных технологий для динамического управления распределением ресурсов в режиме реального времени [1, 2]. При этом среди преимуществ использования именно мультиагентных технологий следует отдельно выделить возможность обеспечения согласованного планирования с учетом сложности и многообразия существующих связей в условиях нестабильности и высокой динамики. Одной из таких задач, при решении которой важно обеспечить применение наиболее современных технологий автоматизации планирования и управления распределением ресурсов является планирование программы полета и грузопотока Международной космической станции (МКС).

Для поддержания жизнедеятельности МКС необходимо решить множество взаимосвязанных задач по поддержанию жизнедеятельности экипажа, доставке лабораторного оборудования, различных материалов и инструментов. Кроме этого, заказчиков космических стартов интересует не только грузоподъемность и надежность космических кораблей, но и стоимость транспортировки. По сравнению с авиаперевозками, отношение цена/масса доставленного груза в космической индустрии весьма высока, в связи с чем весьма актуальна проблема оптимизации распределения грузов.

В данной статье рассматриваются основные принципы работы мультиагентной системы для поддержки принятия решений в управлении деятельностью по планированию грузопотока МКС. Особенностью предлагаемого мультиагентного подхода является возможность построения программы полетов в форме динамически формируемой сети потребностей и возможностей ресурсов МКС, конфигурация (временное равновесие) агентов которой в любой момент может изменяться под действием внешних или внутренних событий, включая изменение приоритетов, предпочтений или ограничений.

Интерактивный характер взаимодействия системы и ее пользователей (специалистов по планированию грузопотока) означает, что диалог с системой может быть инициирован каждой

из сторон в любой момент времени, что позволит адаптивно, т.е. без полной перестройки заново, доставлять или перестраивать планы по мере возникновения важных событий.

1 Постановка задачи

Для успешной работы МКС требуется решение задачи планирования грузопотока, включая доставку на борт станции лабораторного оборудования, различных материалов и инструментов, пополнение запаса воды и продовольствия для космонавтов и эвакуации продуктов их жизнедеятельности, с учетом обеспечения безопасности существующей на МКС замкнутой системы жизнеобеспечения.

Необходимость проведения работ обуславливается возрастающей сложностью процессов планирования и отслеживания исполнения программы полета и грузопотока МКС, что вызывается, в том числе, следующими причинами:

- рост числа грузов, доставку и возвращение которых необходимо планировать с учетом накопления грузов (в том числе не используемых) на МКС;
- рост интенсивности запусков космических кораблей с международным участием, что требует работы с самыми различными грузами от самых разнообразных организаций;
- значительное число непредвиденных событий, связанных, например, с неготовностью груза к полету или переносом сроков полета;
- необходимость учета индивидуальных особенностей пусков космических кораблей и каждой единицы оборудования;
- усложнение грузов, наличие у них разнообразных особенностей и широкого спектра характеристик и параметров;
- необходимость синхронизации использования и доставки необходимо различных грузов;
- необходимость исполнения регламентов обслуживания и правил использования грузов;
- необходимость нахождения оптимального баланса транспортировки топлива, воды, других ресурсов системы жизнеобеспечения, и прочих грузов.

Процесс принятия решения по включению того или иного груза в план полета конкретного космического корабля характеризуется двунаправленной зависимостью: при изменении программы полета должно происходить изменение грузопотока (поскольку необходимо учитывать индивидуальные условия транспортировки груза) и наоборот. Высокая динамика развития этих процессов приводит к невозможности совершенствования технического и информационного обеспечения стандартными методами. В подобных условиях использование «классических» методов и приемов при планировании грузопотока зачастую оказываются весьма «тяжеловесным» и затратным.

Для решения указанной задачи предлагается использование мультиагентной технологии [3], которая позволяет учитывать специфику предметной области. В данной системе каждому грузу и космическому кораблю присваивается программный агент, который, действуя по поручению и от лица каждого груза, корабля и/или его владельца, договаривается с другими агентами и планирует оптимальную доставку груза на конкретном корабле «точно-в-срок».

Особенностью описываемого мультиагентного подхода является возможность построения плана грузопотока в результате функционирования динамически формируемой сети потребностей и возможностей [4, 5] ресурсов МКС, конфигурация (временное равновесие) агентов которой в любой момент может изменяться под действием внешних или внутренних событий, сведения о которых поступают от центра управления полетами (ЦУП).

2 Описание мультиагентного подхода к решению поставленной задачи

Предположим, имеется упрощенная модель МКС со всего лишь одним отсеком, заполненным рядом объектов, и в текущий момент планируется три запуска, причем корабли могут

иметь различный объем грузового отсека (см. рисунок 1). Подсистемы станции, имеющие потребность в разных грузах, показаны здесь как объекты различной формы, цвета и размера. Каждая такая подсистема МКС требует соответствующих грузов, места под которые бронируются в запусках, для этого объекты, размещенные на борту, могут самостоятельно генерировать запрос на материалы и регламентные обслуживания. Известен также ряд заказчиков, предполагающих размещение своих объектов на борту станции в момент времени назначенного пуска. Аналогичным образом могут генерироваться запросы на своевременный возврат использованного оборудования или других объектов со станции. Например, для рассматриваемой задачи модель ПВ-сети [4, 5] может включать агентов полетов, грузов, космических кораблей (транспортных средств) и т.д.

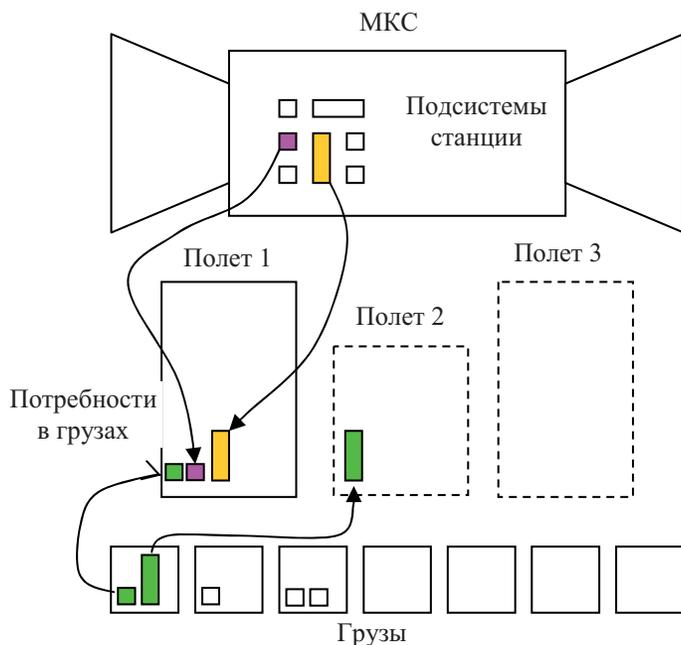


Рисунок 1 – Структура ПВ-сети МКС

При этом груз постоянно ищет себе лучший вариант размещения на космическом корабле, а космический корабль – наилучшую загрузку, удобные сроки и т.п. Получая предложения от разных полетов (возможностей), заказ на транспортировку груза (потребность) может решить, какой из имеющихся полетов ему лучше всего подходит. Но, с другой стороны, и сам полет (возможность) получает предложения от разных грузов (потребностей) на перевозку и решает, какие именно из грузов для него являются более предпочтительными. Эти решения не являются один раз и навсегда принятыми, они могут пересматриваться и изменяться по мере того, как меняется ситуация и возникают события в реальном времени. При этом установление новых связей между агентами вызывает изменение условий функционирования для других агентов, и, тем самым, определяет процесс самоорганизации системы, приводящей к перестройке расписания в ответ на возникающие события. Результат считается достигнутым и система завершает свою работу в том случае, когда ни у одного агента нет больше возможностей улучшить свое состояние.

Сложность модели ПВ-сети увеличивается как с ростом числа типов программных агентов, представляющих разнообразные интересы, предпочтения и ограничения различных участников, необходимых для решения задачи, так и ростом числа возможных взаимодействий между агентами разных типов.

Конфликты, порождаемые событиями (например, приход новой потребности, для которой в текущей ситуации нет открытой возможности), могут разрешаться агентами заказов и ресурсов путем переговоров и взаимных уступок, направленных на достижение приемлемых для всех компромиссов. Разрешение конфликта может вызывать целую цепочку операций перепланирования (включая переход заказов (грузов) на другой ресурс (полет), сдвиг заказов вправо или влево по шкале времени, обмен заказами между ресурсами и т.д.), глубина которой может быть ограничена допустимым временем ответа или другими факторами. В то же время, если имеется запас времени, решение о выделении ресурса или сформированное расписание использования ресурса может подвергаться непрерывной, в том числе, и классической оптимизации. Постоянная активность всех агентов сети, причем как со стороны потребностей, так и возможностей, вызывает многосторонние переговоры в системе, идущие асинхронно и квазипараллельно. Таким образом, реализуется способность системы оперативно реагировать на заранее непредвиденные события.

3 Поддержка принятия согласованных решений

Задачу динамического планирования предлагается формулировать как задачу поиска и поддержания баланса интересов различных служб, подсистем корабля, внешних заказчиков и отдельных грузов.

В основе предлагаемой мультиагентной системы лежит архитектурная концепция постоянно действующего «виртуального круглого стола» агентов, которые участвуют в переговорах. Целью «круглого стола» является выявление спектра мнений по поставленной задаче с разных точек зрения и поиск какого-либо коллегиального решения, в той или иной мере удовлетворяющего все стороны. Методология «круглого стола» заключается в обеспечении активного взаимодействия лиц, принимающих решение – участников круглого стола, основанного на использовании единого понятийного аппарата и организованного в едином информационном пространстве. На рисунке 2 показано взаимодействие агентов, которое обеспечивается работой с единой моделью ситуации. За каждым из этих агентов на первом этапе может стоять пользователь, принимающий решения в ручном режиме, или находиться целая подсистема планирования, автоматически вырабатывающая решения для согласования с другими агентами.

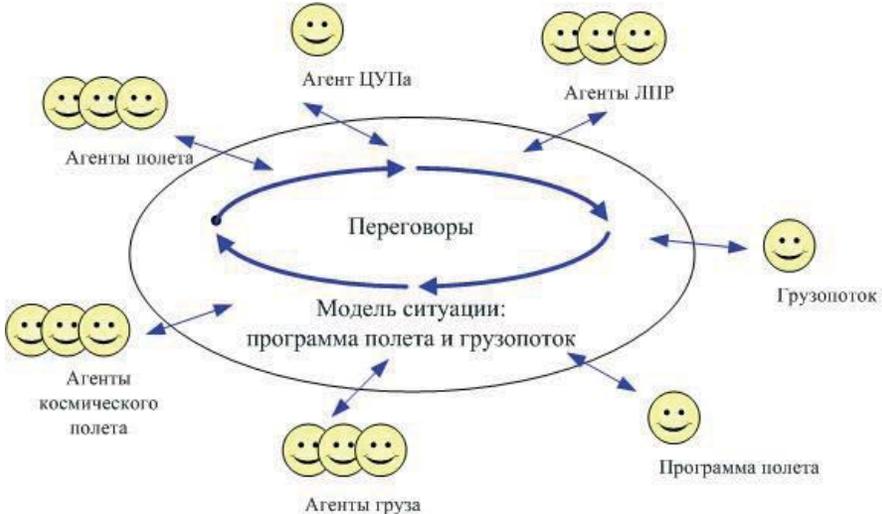


Рисунок 2 – Логическая архитектура виртуального «круглого стола»

Такой принцип построения мультиагентных систем, когда за агентом может быть скрыта целая мультиагентная подсистема, называется «холоническим» (от лат. Holon – элемент, соединяющий в себе свойства части и целого). При этом обеспечиваются открытость, масштабируемость, высокая производительность и надежность решения, а также практически неограниченные возможности развития и, таким образом, решение все более сложных и динамичных задач в реальном времени.

Описанный метод во многом интегрирует современные идеи оптимального планирования, реализуемого в мета-эвристиках, фактически создавая среду конкурирующих и кооперирующих алгоритмов (агентов). Так, агенты могут запоминать и избегать плохих решений за счет использования своей памяти, информировать друг друга о промежуточных опциях, при близости опций принимать решения случайно, прекращать поиск при наличии ограничений по времени принятия решений и т.д.

При этом за счет представления задачи в форме, близкой к естественной, логика принятия решений системы становится более прозрачной как для программистов, так и для пользователей, что позволяет встраивать большее число эвристик без увеличения сложности кода и уменьшает общее время разработки системы, а также делает результаты системы доступными для понимания пользователем.

4 Пример мультиагентного планирования при построении программы полета

Рассмотрим пример (см. рисунок 3) решения задачи динамического планирования нескольких полетов на один ресурс (стыковочный порт МКС). Когда три заказа уже разместились в расписании порта на соответствующих интервалах времени (такт 1), поступает четвертый заказ, который хочет разместиться в расписании в определенный временной период.

Такт 2. Агент полета 4 (П4) пытается встать на требуемое ему место в расписании и, как видно из рисунка, сразу обнаруживает конфликт с П2 с перекрытием на дельту Д1. Тогда агент решает для себя, не может ли он уйти от конфликта, сдвинувшись влево, но тут наталкивается на П1, при этом все-таки уменьшается зона конфликта (дельта Д2), пусть при этом не возникает отклонение от предпочтения П4, заданного, например, достаточно широкими окнами, и соответственно нет неудовольствия этого агента.

Такт 3. Теперь П4 решает, что делать дальше. Либо попросить П1 подвинуться влево на дельту Д2, либо попросить П2 сдвинуться вправо на ту же дельту Д2, причем можно попробовать оба варианта последовательно, либо запустить эти опции параллельно. Поскольку исследования вариантов взаимодействия между агентами будут продолжаться, то при этом весь смысл «параллельности» будет состоять в том, чтобы не перебирать все варианты до конца, один за другим, а начать их исследовать все сразу, но постоянно сопоставлять прогресс и «обрывать» нити вариантов как можно быстрее, по мере того, как в ходе переговоров выявляются серьезные ограничения.

Такт 4. Пусть П2 работает в таком параллельном режиме и делает запросы одновременно к П1 на подвижку влево и П2 на подвижку вправо на Д2. П2 соглашается рассмотреть вариант своей подвижки вправо, но сразу обнаруживает, что свободного места вправо с учетом необходимости соблюдения некоторого промежутка времени между стыковками недостаточно для нужного сдвига на Д2, о чем он немедленно сообщает П4, чтобы тот мог сравнить варианты и решить, в каком направлении следует продолжать поиск вариантов. Пусть П1, со своей стороны, сообщает, что он готов двигаться влево, но тут нарушаются его предпочтения и потому его удовлетворенность упадет, что потребует компенсации. В результате сопоставления этих вариантов П4 решает продолжить исследование варианта подвижки П2 вправо. В этом случае П4 просит П2 продолжить переговоры и выяснить возможность продолжения сдвига вправо. Пусть П3 категорически отказывается от подвижки вправо, поскольку очень существенно нарушаются его предпочтения, и тогда ему останется только уходить на другой ресурс, напри-

мер, на другой день. Тогда П2 сообщает П4, что может уйти вправо только на небольшую дельту и тогда П4 остается только попросить П1 сдвинуться влево на дельту Д3 за компенсацию.

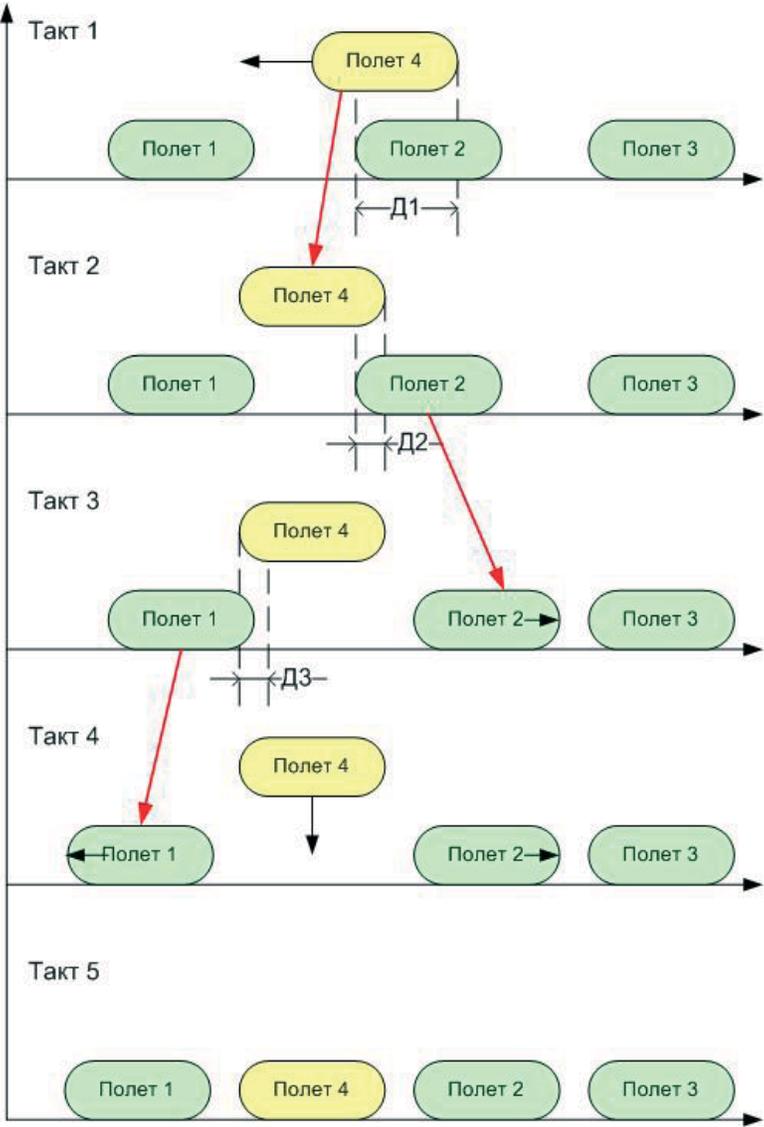


Рисунок 3 – Последовательность переговоров

Когда процесс переговоров завершается согласием всех участников (такт 5), т.е. решение задачи найдено – диспетчером завершается фаза планирования системы и дается команда агентам на переход к фазе исполнения, чтобы агенты могли исполнить намеченные планы и осуществить все подвижки.

В результате проведенных переговоров и достигнутого согласия П1 сдвигается вправо на Д3, П2 уходит влево на небольшую дельту, т.е. П1 и П2 расступаются для предоставления П4 возможности занять требуемое ему место, но это также требует от П4 компромисса, отступа от его первоначального предпочтения, и сдвига влево на небольшую дельту.

Заметим, что при планировании практически не использованы возможности глобального анализа ситуации по сцене – ведь П4 в ходе выявления конфликта с П2 мог «пробежаться» по всей цепочке вправо, оценить наличие там свободного места и увидеть, что справа уже есть большой «хвост» задач, которые потребуются двигать. И если нужно ответить как можно быстрее, то следует начинать переговоры с П1, даже если вариант будет не оптимальным и общая неудовлетворенность в системе окажется больше. Полученный результат может быть поправлен на этапе проактивности, когда П1 попросит П4 попытаться сдвинуться вправо, если у него есть такая возможность, поэтому приведенный выше результат может быть получен и путем такого маневра.

Реализацию описанной выше логики мультиагентного планирования иллюстрирует рисунок 4. Интерактивное редактирование временных характеристик полетов активизирует работу мультиагентного планировщика, который подсказывает пользователю возможные варианты решения проблемных ситуаций. Изменения программы полета влияют на грузопоток – при переносе времени старта грузы могут быть перепланированы. В свою очередь, план по грузопотоку может повлиять на программу полета: например, в случае недозагрузки космического корабля полет может быть отменен, а оставшиеся грузы перераспределены по более поздним, или более ранним стартам.

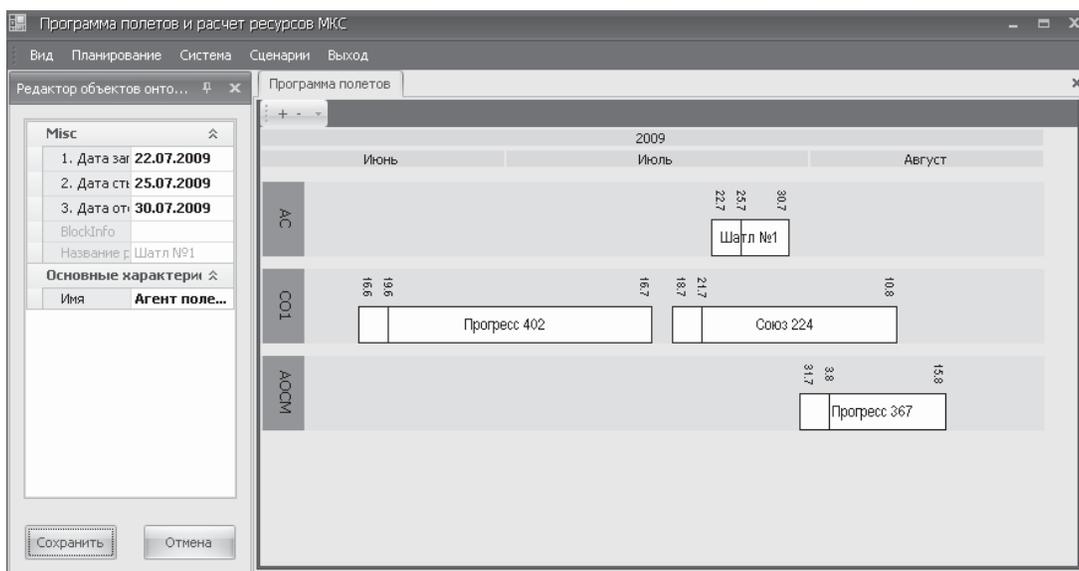


Рисунок 4 – Основное окно программы полета

5 Применение онтологий для построения базы знаний

Для работы агентов в предлагаемом подходе необходима формализованная модель ситуации, которая будет описывать начальное состояние всех заказов, план по запускам, потребности подсистем корабля и других ресурсов, а также все другие необходимые сведения в любой заданный момент времени. В этих целях в описанном подходе создается и используется редактор онтологий планирования грузооборота МКС, предназначенный для отображения и редактирования основных концептов и взаимосвязей предметной области.

Онтология является определенного рода словарем терминов для сцены. Сцена – это набор экземпляров классов, описывающих некоторую ситуацию, и каждый из экземпляров набора связан с некоторым концептом онтологии отношением вид-род. В центре онтологии находятся

классы, у каждого из которых могут быть подклассы. Классы онтологии образуют несколько иерархических конструкций, описывающих понятия предметной области. У классов могут быть определены атрибуты, которые описывают свойства и внутреннюю структуру понятий, лежащих в основе классов. Все подклассы наследуют атрибуты родительских классов. Тип значения атрибута описывает, какие типы значений может содержать атрибут, например строку или число.

Для решения описываемой задачи использование онтологий обусловлено необходимостью иметь возможность изменения используемых при автоматизированном принятии решений знаний. Конфигурация, состав оборудования и потребности в грузах станции постоянно изменяются. Предлагается описывать предметные подобласти станции отдельными онтологиями, которые затем будут использованы для построения моделей станции в ее конкретных конфигурациях и будут отражать состояние станции сценами («моментальными фотографиями») объектов и отношений. Это позволит отделить знания об объектах и их связях от программного кода системы и пополнять описания мира станции по мере необходимости, без перепрограммирования системы в каждом случае изменений.

Онтология планирования грузопотока МКС основана на методе семантических сетей, позволяющем формализовано описать знания предметной области. Семантическая сеть предметной области образуется классами понятий и отношений, которые играют роль связей. Для рассматриваемой предметной области примерами таких классов понятий могут быть «полет», «космический корабль», «запуск», «заявка», «груз», а отношений – «принадлежит» (груз – заявке), «забронирован» (груз – запуск) и другие. Понятия и отношения могут иметь свои свойства и атрибуты, например, атрибутами понятия «запуск» могут являться время и тип запланированного транспортного средства, а атрибутами отношения «забронирован» – момент времени, когда было выполнено такое бронирование.

Работа пользователя с редактором онтологий заключается в выполнении двух основных задач:

- формирование собственно семантической сети, т.е. выделение и ввод концептов, задание их основных свойств и указание иерархических связей между различными концептами;
- задание взаимосвязей различного типа между концептами.

Приведем простой пример, поясняющий использование онтологии. Предположим, была создана онтология мира МКС и с помощью данной онтологии была построена сцена, которая связывает конкретный экземпляр понятия «МКС» с некоторой конкретной «установкой», размещенной на борту станции через отношение «принадлежит». В свою очередь, установка «нуждается» в определенных материалах и порождает заказ на груз, который далее «бронировать» себя в одном из конкретных запусков. Все эти объекты и связи определяют сцену, которая фактически и является собой расписание, поскольку в указанных отношениях сохраняются соответствующие параметры, указывающие, когда был порожден заказ, на какой полет он договорился и т.д.

Если теперь на станцию планируется завозить новое оборудование, которое требует запчастей, расходных материалов и энергии, то в онтологию будет введен новый объект, для которого будет специфицировано, что именно он требует, через указанное выше отношение «нуждается-в», причем для данного отношения имеются свои базовые подклассы, для которых можно задавать, кто, когда, что и сколько требует (регулярно, однократно и т.д.). Тогда общий агент нового оборудования сможет зачитать из онтологии конкретную специфику потребностей данного оборудования и создать под них связанных агентов, которые на разных кораблях (т.к. потребности различны) будут планировать необходимые грузы к требуемому сроку.

В этом случае придется перепрограммировать систему только в случае возникновения непредвиденных ранее новых подклассов данного отношения, причем после такого дополнительного введения соответствующее отношение пополнит общую библиотеку отношений и станет доступным для всех подобных ситуаций.

Таким образом, формализованное представление знаний в форме онтологий предоставляет следующие возможности:

- единство терминологии, т.е. описания объективных элементов действительности в единых, строго определенных образцах;
- использование агентами знаний о потребностях объектов в ходе работы и принятия решений;
- расширение знаний в случае появления нового знания;
- обеспечение гибкой перенастройки системы без перепрограммирования;
- возможность анализа альтернативных вариантов грузопотока, что позволяет проводить его оптимизацию.

Заключение

Описанный в данной статье мультиагентный подход к динамическому управлению распределением ресурсов в реальном времени для поддержки принятия решений в управлении деятельностью по формированию программы полета и планированию грузопотока для МКС позволяет повысить эффективность управления работой МКС, обеспечить гибкость планирования грузопотока, оперативность в решении поставленной задачи, удобство для пользователей, высокую производительность и надежность.

В результате предлагается программное решение нового типа, обеспечивающее:

- организацию коллективной работы всех групп пользователей по согласованию программы полета и грузопотока в рамках единого информационного пространства;
- полнофункциональные средства анализа программы полета и плана грузопотока с целью оптимизации и поиска возможных альтернативных вариантов плана;
- возможность моделирования различных ситуаций и быстрой (оперативной) проработки разных вариантов программы полета и плана грузопотока в случае необходимости.

В перспективе предлагаемый подход позволит создавать адаптивные сети динамических планировщиков, способных обеспечить поддержку коллективного согласования и принятия решений в реальном времени на различных этапах разработки, производства, транспортировки и запуска объектов космической техники, в разных подразделениях и организациях, работающих совместно над решением данных задач.

Список литературы

- [1] Multi-Agent Real Time Scheduling System for Taxi Companies / Andrey Glaschenko, Anton Ivaschenko, George Rzevski, Petr Skobelev // AAMAS 2009, Budapest, Hungary. – 2009. – р. 29 – 36.
- [2] Андреев В.В., Глащенко А.В., Иващенко А.В., Иноземцев С.В., Скобелев П.О., Швейкин П.К. Мультиагентные системы адаптивного планирования мобильных ресурсов в реальном времени / в сб. трудов четвертой международной конференции по проблемам управления – М.: Институт проблем управления им. В.А. Трапезникова РАН, 2009. – с. 1534 – 1542.
- [3] Андреев М.В., Иващенко А.В., Скобелев П.О., Царев А.В. Построение адаптивной системы управления предприятием с использованием мультиагентных технологий // Вестник Самарского государственного технического университета, Серия «Технические науки» № 1 (23) – 2009 – с. 5 – 14.
- [4] Виттих В.А., Скобелев П.О. Мультиагентные модели взаимодействия для построения сетей потребностей и возможностей в открытых системах / Автоматика и Телемеханика, 2003, №1. – с. 177 –185.
- [5] Виттих В.А., Скобелев П.О. Метод сопряженных взаимодействий для управления распределением ресурсов в реальном масштабе времени / Автометрия, 2009, № 2. – с. 78 – 87.

Author index

Индекс авторов

А

Абарин Д.В., 483
Андреев М.В., 670
Аникин И.В., 274, 281, 606
Апанович З.В., 518
Асанов А.З., 218, 224, 232
Асеева Н.В., 464
Ахатова Ч.Ф., 615

Б

Бабер И.С., 298
Бабкин Э.А., 464
Батищев В.И., 263
Батышкина В.В., 470
Бедняк О.И., 505
Беленький Л.Б., 328, 336
Белослудцев В.А., 310, 316
Блинова Д.В., 195
Боровик С.Ю., 316, 328, 336
Булгаков С.В., 512
Бульхин А.К., 412, 420
Буцких В.А., 435

В

Вайсблат А.В., 678
Васин Н.Н., 435
Визгунов А.Н., 458, 464
Вильвер П.Ю., 364
Виноградов И.Д., 158
Витгих В.А., 33, 452
Волков В.Ю., 470
Вольман С.И., 577, 586

Г

Габидулин М.А., 429
Гавриленко О.С., 120
Галяутдинова А.Н., 348
Гвоздев В.Е., 195
Гинзбург А.Н., 33, 298, 304, 501, 548
Головихина О.Е., 40
Горемыкина Г.И., 599
Губанов Н.Г., 263

Д

Давыдов А.Н., 126
Дегтярев С.В., 478
Дегтярева И.В., 176
Демьянов Д.Н., 218
Дёмкин В.М., 304, 501
Дилигенская А.Н., 165
Дилигенский Н.В., 126
Диязитдинова А.Р., 678, 686
Дрынова И.В., 478

Е

Елсуков В.С., 120

Ё

Ёлкин Д.А., 577

З

Загорулько Г.Б., 512
Загорулько Ю.А., 512
Зиновьев П.А., 268
Золотухин Ю.Н., 291
Зубов М.Л., 458

И

Иващенко А.В., 657, 670, 678, 686
Ившина Ю.В., 392
Игнатъев М.В., 548
Ильясов Б.Г., 176, 187
Инден У., 657

К

Каримов В.С., 224
Карпов Н.И., 298
Катасёв А.С., 615
Кижаяев С.А., 412, 420
Кислицина Т.А., 518
Кистанов А.М., 211
Кичигин В.В., 100, 105
Ключников В.Ф., 412, 420
Князькин В.П., 464
Козырев О.Р., 33
Колеватов А.П., 387
Колоденкова А.Е., 593

Кононенко И.С., 540
Кораблин М.А., 505
Кочемасов А.В., 382
Кочуров А.В., 586
Куликов А.В., 387
Курманалиева Р.Н., 240
Кутейникова М.М., 316, 323

Л

Ларин И.А., 204
Лачин В.И., 120
Липатов А.Ю., 370
Лукоянов В.С., 281

М

Маевский Д.П., 488
Макарова Е.А., 176
Максимов А.И., 100, 105
Мараканов И.Н., 570
Мартышкин Д.М., 670
Марьев С.П., 298
Масленникова Ю.С., 348
Мейко А.В., 268
Меньшиков В.Ф., 40
Микайылов Ф.Д., 137
Минаков И.А., 577, 586
Митрофанов О.Г., 298
Моржова С.В., 398
Мышкина И.Ю., 232

Н

Набиуллин О.Р., 464
Назаров А.Г., 40
Назаров А.Ш., 187
Нестеров А.А., 291
Николаев С.Г., 392

П

Павленко А.И., 644
Павлова А.Н., 176
Петрин К.В., 18
Печенков А.Ю., 100, 105
Позняк И.В., 100, 105
Пономарева М.А., 599
Порунов А.А., 348

Потапов А.С., 606
Прокофьев А.М., 141
Протасов А.В., 364
Пунда Д.И., 146

Р

Райков Б.К., 316, 328
Рапопорт Э.Я., 110

С

Сайтова Г.А., 187
Салмин А.А., 505
Самаха Б.А., 470
Секисов Ю.Н., 310, 316, 323
Серова Е.Г., 562
Сидорова Е.А., 534, 540
Симонова Е.В., 686
Ситников П.В., 553
Скачков О.В., 204
Скобелев О.П., 316, 328, 336
Скобелев П.О., 657, 670, 678, 686
Смирнов С.В., 33, 505, 526
Суворова Е.Г., 526, 553
Сычева М.В., 686
Сюсин И.А., 657

Т

Тарасов А.Д., 650
Теряев Е.Д., 18
Тулупова В.В., 316, 343

У

Уланова Л.В., 670

Ф

Фалдин Н.В., 370, 398
Филимонов А.Б., 18
Филимонов В.А., 488
Филимонов Н.Б., 18
Филиппов М.Н., 244
Фурасов В.Д., 443
Фурасов Д.В., 443

Х

Хамиц И.И., 686

Ц

Цапенко М.В., 126
Царев А.В., 657, 670, 678, 686

Ч

Чекалов Л.Л., 250, 256
Чефранова М.А., 171
Чигарьков Г.К., 382

Ш

Шатунов А.Н., 100, 105
Шелехов В.И., 622
Шиткова М.С., 464
Шишин А.В., 458

Я

Якушин А.В., 586
Яшков В.А., 376

В

Вааке Е., 59, 88
Babkin E.A., 495
Bijl A.C., 28
Blinov K., 88
Bulkin A.K., 404

D

Dughiero F., 67

F

Forzan M., 67

I

Inden U., 5

K

Kizhaev S.A., 404
Klyuchnikov V.F., 404
Korshikov S., 88

L

Lupi S., 67

N

Nacke B., 51, 74
Nicoletti F., 67

P

Pleshivtseva Yu.E., 74, 81

R

Rzevski G., 3, 633

S

Sharapova O., 88
Shemyakin Yu.V., 94

Научное издание

Труды XII Международной конференции
ПРОБЛЕМЫ УПРАВЛЕНИЯ
И МОДЕЛИРОВАНИЯ В СЛОЖНЫХ СИСТЕМАХ

21-23 июня 2010, Самара, Россия

Редакторы:

академик Е.А. Федосов
академик Н.А. Кузнецов
профессор В.А. Виттих

Proceedings of the XII International Conference
COMPLEX SYSTEMS:
CONTROL AND MODELING PROBLEMS

June 21-23, 2010, Samara, Russia

Editors:

academician E.A. Fedosov
academician N.A. Kuznetsov
professor V.A. Vittikh

Подписано в печать 3.06.2010.
Формат 70x100/16. Бумага офсетная. Печать офсетная.
Объем 56,87 усл. печ. л. Тираж 500 экз. Заказ № 894.

Отпечатано в типографии ООО «Офорт».
443080, г. Самара, ул. Революционная, 70, литера П.
Тел.: 372-00-56, 372-00-57.