

ISSN 1998-0663

№1(23)—2013

<http://bijournal.hse.ru>

БИЗНЕС- ИНФОРМАТИКА

МЕЖДИСЦИПЛИНАРНЫЙ НАУЧНО-ПРАКТИЧЕСКИЙ ЖУРНАЛ НИУ ВШЭ

BUSINESS INFORMATICS

Учредитель:
Национальный
исследовательский университет
«Высшая школа экономики»

Редакционная коллегия

Абдульраб А.
Авдошин С.М.
Алескеров Ф.Т.
Бабкин Э.А.
Баранов А.П.
Беккер Й.
Белов В.В.
Грибов А.Ю.
Громов А.И.
Гурвич В.А.
Джейкобс Л.
Зандкуль К.
Ильин Н.И.
Калягин В.А.
Каменнова М.С.
Кузнецов С.О.
Мальцева С.В.
Миркин Б.Г.
Моттль В.В.
Пальчунов Д.Е.
Пардалос П.
Силантьев А.Ю.
Таратухин В.В.
Ульянов М.В.
Шалковский А.Г.

В ЭТОМ НОМЕРЕ:

ПРОИЗВОДСТВЕННО-
СБЫТОВЫЕ
СИСТЕМЫ

U-СПЕКТРЫ

ПЛАГИАТ
НЕ ПРОЙДЕТ

О ПРОСТРАНСТВЕННЫХ
ДАННЫХ

В соответствии с решением президиума ВАК РФ журнал «Бизнес-информатика» с 19.02.2010 включён в Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней кандидата и доктора наук.

*Журнал зарегистрирован в «Роскомнадзоре».
Свидетельство ПИ № ФС 7752404 от 28 декабря 2012 г.*

БИЗНЕС- ИНФОРМАТИКА

№1(23)–2013

СОДЕРЖАНИЕ

Принятие решений и бизнес-интеллект

В.Г. Чеботарев, А.И. Громов

Роль субъектности в бизнес-процессах..... 3

Моделирование социальных и экономических систем

М.А. Ивлев

Математические основы теории производства-потребления:
определение вида, структуры и параметров моделей 10

В.К. Абросимов, В.В. Лебидько

Имитационное моделирование организации
массовых мероприятий (на примере XXII зимних
Олимпийских игр 2014) 19

А.В. Федотова, М.В. Овсянников, В.В. Таратухин

Применение метода программирования в ограничениях
при решении задач планирования технического
обслуживания и ремонта воздушных судов 28

Математические методы и алгоритмы бизнес-информатики

В.В. Мокеев, С.В. Томилов

О решении проблемы выборки малого размера
при использовании линейного дискриминантного
анализа в задачах распознавания лиц 37

В.Е. Гай

Алгоритмы формирования спектрального представления
звукового сигнала на основе U-преобразования 44

В.В. Дягилев, А.А. Цхай, С.В. Бутаков

Надежное обнаружение плагиата малым числом
поисковых запросов 50

Информационные системы и технологии в бизнесе

Д.В. Исаев

Информатизация корпоративного планирования
и бюджетирования 58

Дискуссионный клуб

Ф.В. Краснов, Р.Э. Яворский

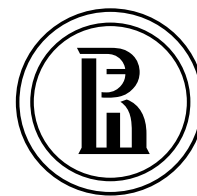
Измерение уровня зрелости профессионального
сообщества 64

Обзоры

К.В. Латышев, В.Н. Сидоренко

Применение современных информационных технологий
для анализа пространственно распределенных
социально-экономических данных 68

Annotations 76



БИЗНЕС-ИНФОРМАТИКА

№1(23)–2013

Междисциплинарный
научно-практический журнал
НИУ ВШЭ

Журнал рекомендован ВАК
для научных публикаций

Подписной индекс издания
в каталоге агентства
«Роспечать» – 72315

Учредитель:
Национальный
исследовательский университет
«Высшая школа экономики».
Выходит 4 раза в год.

Главный редактор
Голосов А.О.

Заместители главного редактора
Горбунов А.Р., Исаев Д.В.

Научный редактор
Лычкина Н.Н.

Технический редактор
Осипов В.И.

Дизайн обложки
Борисова С.Н.

Компьютерная вёрстка
Богданович О.А.

Администратор веб-сайта
Проценко Д.С.

Адрес редакции:
105187, г. Москва,
ул. Кирпичная, д. 33/5.
Тел. +7 (495) 771-32-38,
e-mail: bijournal@hse.ru

За точность приведённых сведений
и содержание данных,
не подлежащих открытой публикации,
несут ответственность авторы

При перепечатке ссылка на журнал
«Бизнес-информатика» обязательна

Тираж 500 экз.

Отпечатано в типографии НИУ ВШЭ
г. Москва, Кочновский проезд, 3.

© Национальный
исследовательский университет
«Высшая школа экономики»

РОЛЬ СУБЪЕКТНОСТИ В БИЗНЕС-ПРОЦЕССАХ

В.Г. Чеботарев,

кандидат технических наук, доцент кафедры моделирования и оптимизации бизнес-процессов Национального исследовательского университета «Высшая школа экономики»

А.И. Громов,

кандидат химических наук, профессор, заведующий кафедрой моделирования и оптимизации бизнес-процессов Национального исследовательского университета «Высшая школа экономики»

E-mail: vchebotarev@hse.ru, agromov@hse.ru

Адрес: г. Москва, ул. Кирпичная, д. 33/5

Статья посвящена исследованию роли субъектности, субъектной структуры и рефлексивности в управлении бизнес-процессами предприятия. Данное исследование проводилось при финансовой поддержке Правительства Российской Федерации (Минобрнауки России) в рамках договора № 13.G25.31.0096 о «Создании высокотехнологичного производства кросс-платформенных систем обработки неструктурированной информации на основе свободного программного обеспечения для повышения эффективности управления инновационной деятельностью предприятия в современной России».

Ключевые слова: субъектная структура, бизнес-процесс, системное мышление, интерактивное моделирование, бизнес-архитектура.

По мнению многих авторитетных ученых человечество стоит на пороге больших перемен, требующих пересмотра взглядов, как на природу организации, (предприятия), так и на метод научного познания организации. Эти перемены происходят и во внешней среде, и на самих предприятиях, нередко они требуют пересмотра системы понятий, ценностей и целей. В чем

причины потери конкурентного преимущества передовых предприятий и как эти причины могут повлиять на наше понимание того, что есть предприятие (организация), и какие научные подходы нужно использовать для его исследования? Какие сущности, роли отводятся сотрудникам предприятия, как использовать потенциал сотрудников на благо предприятия и что для этого необходимо

изменить? Как должно выглядеть предприятие 21 века? Поиску ответов на поставленные вопросы посвящено данное исследование.

1. О терминологии

Для корректности исследования чрезвычайно важно определить базовые термины исследования. К ним относятся следующие термины: «субъект», «субъектность», «субъектная структура», «рефлексия», «рефлексивность». Эта терминология затрагивает несколько научных областей (философия науки, психология, экономика и др.), в каждой из которых смысл этих терминов имеет свои особенности.

Субъект — носитель деятельности, сознания и познания [1].

Субъектность — это характеристика человека как субъекта познания, отражающая полноту его актуальных и потенциальных возможностей и степень их реализации не только в настоящем, но и в потенции, в будущем [2, 3]. В потенциале субъекта выделяются два аспекта: потенциал как данность и потенциал как возможность, т.е. потенциал субъекта может быть представлен как совокупность его актуальных и потенциальных возможностей, рассматриваемых с точки зрения нижнего (актуального), и возможного высшего, потенциального, предела развития и реализации.

Субъектная структура включает в себя ценности, цели, знания и навыки [3].

Рефлексия рассматривается в контексте деятельности и с точки зрения средств деятельности: 1) рефлексия как процесс и особая структура деятельности и 2) рефлексия как принцип развертывания схем деятельности [4].

Понятие рефлексивность тесно связано с понятиями «рефлексия» и «субъект». Джордж Сорос так объясняет, что такое рефлексивность [5]: «...С одной стороны, участники пытаются понять ситуацию, в которой они участвуют. Они пытаются создать картину, соответствующую реальности. Я называю это пассивной, или когнитивной, функцией. С другой стороны, они пытаются оказать влияние, подделать реальность под их желания. Я называю это активной функцией, или функцией участника. Когда реализуются одновременно обе функции, — я называю такую ситуацию рефлексивной». Отметим, что обе функции реализуются субъектом, дадим уточненные определения терминов «субъект» и «субъектность», учитывающие активную роль

участника при выполнении функции.

Субъект — носитель деятельности, сознания, познания и рефлексии. (1)

Дополнение определения «рефлексией» даёт определению субъекта необходимые и достаточные признаки активного носителя деятельности.

Разница между рефлексией и рефлексивностью состоит в том, что термин «рефлексия» будем использовать по отношению к деятельности или к принципу деятельности, а термин «рефлексивность» — к свойству субъекта быть носителем рефлексии.

Понятие «субъектность» так же, как и понятие «субъект», будет зависеть от рефлексивности, т.к. рефлексивность человека напрямую влияет на степень реализации его возможностей. С учетом этого факта, уточненное определение субъектности может быть представлено в следующем виде.

Субъектность - это характеристика субъекта (как носителя деятельности, сознания, познания и рефлексии), отражающая полноту его актуальных и потенциальных возможностей и степень их реализации в настоящем и будущем. (2)

Приведем также соответствующее определение понятия «субъектная структура». Формально оно мало отличается от исходного, но смысловое отличие значительно. Оно связано с уточнением термина «субъект».

Субъектная структура - структура ценностей, целей, знаний и навыков субъектов (3)

Отметим, что в состав субъектной структуры входят 4 элемента: ценности, цели, знания и навыки, а в состав объектной структуры — всего два: предмет (исходный материал) и результат (предмет как «опредмеченная» цель) [3].

В дальнейшем под терминами «субъект», «субъектность» и «субъектная структура» будем понимать выражения, представленные определениями (1), (2) и (3). Это значит, что в рамках данного исследования *сотрудники предприятия, не соответствующие определению (1), не рассматриваются как субъекты, не обладают субъектностью и не связаны с субъектной структурой предприятия.*

2. Сдвиг парадигмы

Анализируя причины потери конкурентного преимущества передовыми компаниями, Джемшид

Гараедаги выделил 4 скрытых фактора [6]:

- ИЕРАРХИЧЕСКАЯ ПОДЧИНЕННОСТЬ ↓
- ◆ имитация, копирование другими компаниями успешных отличий и технологий лидеров;
 - ◆ инерция, сопротивление передовых компаний необходимости перемен в своих успешных отличиях и передовых технологиях;
 - ◆ оптимизация отдельных частей, которые передовые компании продолжают считать своими конкурентными преимуществами, незаметно превращая свои «плюсы» в «минусы»;
 - ◆ изменение правил игры, успех компании приводит к изменению правил игры, правил конкурентной борьбы в окружающей среде, неучет этого фактора – путь к краху успешной компании.

Суммарный эффект от действия 4 факторов приводит нас к пониманию того, что вероятно общепринятые представления о системе понятий, ценностей и целей утратили свою состоятельность и исчерпали свой потенциал. Таким образом, мы приходим к мысли о необходимости изменения (сдвига) парадигмы, причем сразу в двух направлениях (рис. 1).

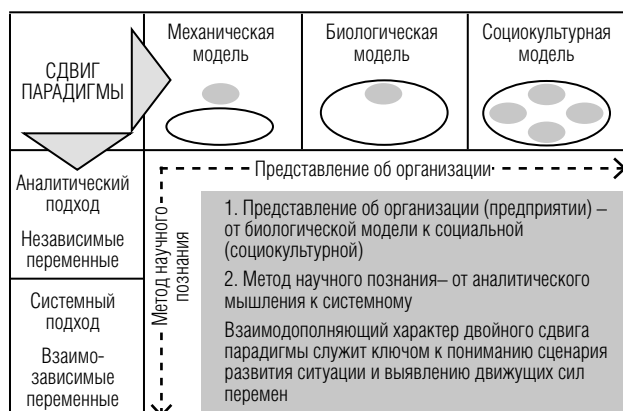


Рис. 1. Направления двойного сдвига парадигмы по Гараедаги

Первое направление сдвига парадигмы означает перемены в подходе к изучению предприятия, т.е. перемены в методе научного познания: от аналитического (механистического) подхода к системному.

В основе различия в подходах к методам исследования организационной системы лежит понятие «система». Аналитический подход основан на упрощенном представлении о независимости частей системы, а системный подход считает взаимозависимость частей важнейшей особенностью понятия «система» [7].

Аналитический (или механистический) подход – мыслительный процесс, состоящий из трёх этапов: 1) система (предприятие) мысленно разбивается на отдельные части; 2) изучается и объясняется поведение каждой части; 3) синтез полученных суждений для объяснения целого. На первых двух этапах используется метод анализа, а на третьем – метод синтеза. В основе подхода лежит гипотеза о независимости частей системы, при этом явление эмерджентности фактически игнорируется.

Системный подход – мыслительный процесс, состоящий из трёх этапов: 1) идентификация целого (системы); 2) объяснение поведения или свойств целого (системы); 3) объяснение поведения или свойств части системы по его роли (ролям) или функции (функциям) в содержащем его целом. На первых двух этапах используется метод синтеза, на третьем этапе – метод анализа. Системный подход позволяет учесть явление эмерджентности.

Второе направление сдвига парадигмы является отражением перемен, которые произошли в представлениях о так называемой «разумности» организации, которая отражается моделью организации [6]. Характерная особенность модели – наличие и месторасположение органа, принимающего решения, т.н. «разума». Перемены означают переход от механистической модели к биологической модели и далее к социальной (социокультурной) модели.

Механистическая модель представляет собой так называемое «неразумное» предприятие, т.к. на предприятии нет органа, принимающего решения. Решения принимает внешний «разум», в роли которого выступает владелец (хозяин) предприятия. Такое предприятие фактически является инструментом хозяина для получения прибыли. Субъектная структура, как и субъекты на предприятии отсутствуют.

Биологическая модель – предприятие (по аналогии с биологическим организмом) содержит единственный принимающий решения орган, один «разум» (человек или коллективный орган). Смысл существования – выживание в окружающей среде. «Разум» предприятия заинтересован в точном и эффективном выполнении процессов, например, с помощью правил и регламентов, выполняющих роль своеобразных генетических кодов. Управляющие воздействия передаются «сверху-вниз» в соответствии с иерархической организационной структурой. При необходимости предприятие способно к трансформации, но со значительным запаз-

дыванием. Неудовлетворительная работа предприятия связывается с недостаточностью информации и помехами в каналах связи, поэтому для решения проблем используется поиск дополнительной информации и улучшение каналов связи. Субъектная структура предприятия определяется субъектной структурой «разума».

Социальная (социокультурная) модель представляет собой так называемое «мультиразумное» предприятие, внутри которого находится множество «разумов», т.е. множество людей или коллективных органов, каждый из которых принимает свои решения и имеет собственные цели [6]. Главной целью «мультиразумного» предприятия является согласование целей участников («разумов», субъектов). В основе выполнения процессов предприятия лежит сетевое взаимодействие. При этом организационная структура нужна для создания благоприятных условий и коммуникаций при сетевом взаимодействии целенаправленных субъектов. Социальные системы самоорганизуются с помощью общей культуры, поэтому точнее их называть социокультурными. Субъектность социокультурного предприятия является его главной характеристикой, а субъектная структура является источником для создания общей культуры и гарантом успешной деятельности. Можно предположить, что неудовлетворительная работа социальной модели будет связана с недостаточным уровнем его субъектности и социальной структуры.

Отметим, что в механистических и социокультурных системах связи между элементами имеют разную природу. Элементы механистической системы связаны между собой энергетически (например, связь между водителем и автомобилем), а элементы социокультурной – информационно (например, связь между наездником и лошадью). По мнению

ряда исследователей [6], именно социокультурная модель является представлением о предприятии 21 века.

Далее выполним сравнение моделей между собой с позиций их характерных особенностей, влияющих на субъектную структуру. Для сравнения выберем следующие характеристики: 1) принятие решения (кто принимает, где находится); 2) смысл существования предприятия; 3) движущая сила (драйвер) совершенствования; 4) передача управляющих воздействий (принципы, механизмы); 5) определяющий фактор для развития субъектной структуры предприятия.

Результаты сравнения моделей (*табл. 1*) демонстрируют принципиальные отличия во взглядах на природу предприятия. Из всех представленных моделей только социокультурная («мультиразумная») модель основана на субъектном взгляде. При выполнении бизнес-процессов на «мультиразумном» предприятии осуществляются сетевые взаимодействия субъектов или коллективов (сообществ) субъектов. Для такого предприятия субъектная структура – основа для общей культуры и совершенствования деятельности. На практике, в первую очередь это относится к творческим, инновационным предприятиям, тем предприятиям, которые сознательно начали движение в сторону социокультурной ориентации и «мультиразумности». Существует авторитетное мнение, что успешная инновационная деятельность возможна только на платформе субъектности [8].

3. Рефлексивное управление

Большинство современных подходов к управлению предприятием построено на управлении функциями или процессами. Функции или процессы яв-

Таблица 1.

Сравнение моделей предприятия

Характерная особенность	Механистическая модель	Биологическая модель	Социокультурная модель
Принятие решения	Вне предприятия	Субъект внутри предприятия	Множество целенаправленных субъектов внутри предприятия
Смысл существования предприятия	Получение прибыли хозяином предприятия	Выживание в окружающей среде и рост	Постоянная балансировка целей субъектов и требований внешней среды
Драйвер совершенствования	Прямые воздействия хозяина предприятия	Управляющие воздействия «сверху-вниз»	Стремление субъектов к достижению своих целей
Передача управляющих воздействий	Прямое воздействие, иерархий	Иерархия	Сетевые взаимодействия
Определяющий фактор развития субъектной структуры предприятия	Субъектная структура на предприятии отсутствует	Субъект, принимающий решения	Множество взаимодействующих целенаправленных субъектов

ляются объектами управления, а всё остальное, что есть на предприятии, относится к ресурсам, необходимым для выполнения функций или процессов. Это правило справедливо и для персонала, который рассматривается как трудовой ресурс.

Многие современные предприятия не могут отойти от пресловутого принципа бихевиоризма, рассматривая трудовой ресурс в виде множества «черных ящиков», на вход каждого из которых подаются стимулы, а на выходах снимаются реакции. Для воздействия на трудовой ресурс используются различные формы подкреплений, например, реакция руководителей или сотрудников предприятия на результат. Более современный, альтернативный бихевиоризму, когнитивный подход (когнитивная психология) основан на аналогии между преобразованием информации вычислительным устройством и между познавательным процессом человека. Человек рассматривается не как «черный ящик», а как система поэтапной переработки информации. При этом стимулы внешнего мира проходят внутри психики через ряд последовательных преобразований. Те качества человека, которые не поддаются объяснению, считаются недопустимыми или просто отбрасываются. В качестве примера таких неудобных человеческих свойств можно привести интуицию, настроение, субъективизм, обман, проявление рефлексивности.

В шестидесятых годах прошлого века психолог и математик, участник Московского методологического кружка В.А. Лефевр пришел к выводу о том, что традиционный естественнонаучный подход недостаточен для изучения таких сложных систем, как человек и его психика. Стремясь преодолеть пропасть, разделяющую мир науки и мир субъективности, Лефевр ввёл понятия рефлексивной системы, рефлексивного процесса и рефлексивного управления. В рамках созданной им теории рефлексии разработал специальный математический

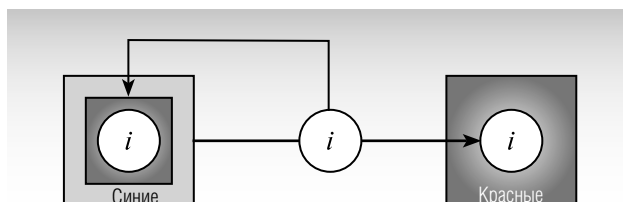
аппарат для исследования рефлексивных процессов и рефлексивного управления.

В качестве эмпирии для разработанного аппарата Лефевр использовал анализ конфликтных ситуаций между противниками (субъектами).

Принцип рефлексивного управления наглядно может быть представлен моделью «Красных и Синих» (рис. 2). Смысл модели чрезвычайно прост: Синим ничего не известно о том, какой информацией владеют Красные; Синие передают некую информацию (i) Красным; теперь Синим известно, какой информацией (i) владеют Красные, остается только подумать, как информация (i) может повлиять на поведение Красных. Это так называемая исходная схема рефлексивного управления, потому, что её можно продолжить: Красным известно, что Синие знают о том, что Красные владеют информацией (i), остается подумать о том, какие действия могут предпринять Синие и т.д. Вложенность схемы рефлексивного управления может быть сколь угодно сложной.

Процесс передачи оснований для принятия решения одним из персонажей другому Лефевр называет *рефлексивным управлением*. Любые «обманные движения», провокации, интриги, маскировки, создание ложных объектов и вообще ложь произвольного типа представляют собой рефлексивное управление. Ложь может иметь сложное строение: например, передача противнику правдивой информации, чтобы он, считая ее ложной, принял соответствующее решение.

Дальнейшее рассуждение требует понимание «рефлексивной системы» и «рефлексивного процесса». Лефевр предложил следующую смысловую аналогию: рефлексивная система – это система зеркал, многократно отражающих друг друга. Если в одном зеркале отразится, например, падающий со стола карандаш, то в других зеркалах траектория его падения отразится с различными искажениями. Каждое зеркало – это аналог субъекта, наделённого своей особой позицией. А весь сложнейший поток отражений зеркал друг в друге будет аналогом рефлексивного процесса. Когда для достижения своих целей субъекты объединяются в коллектив или сообщество, они создают рефлексивную систему, которая при определенных условиях может превратиться в саморефлексивную систему. Наличие рефлексивных связей позволяет создавать, изменять и использовать механизм имитации рассуждений как особое средство координации деятельности без



Модель Красных и Синих:
 1. Синие передают информацию i Красным
 2. Синие знают, что Красные знают i
 и используют это для управления Красными

Рис. 2. Исходная схема рефлексивного управления по Лефевру [9]

непосредственных информационных контактов между субъектами. Коллектив (сообщество) считается сформировавшимся, когда все субъекты обладают возможностями имитации принятия решений другими членами коллектива. Когда появляются знаковые средства планирования деятельности коллектива (сообщества) как единого целого, то происходит его превращение в саморефлективную систему [9].

Исходя из вышеизложенного, определение рефлексивного управления сообществом сформулируем в следующем виде.

Рефлексивное управление сообществом – создание возможностей для преобразования сообщества в саморефлективную систему, а также для поддержки и развития рефлексивных способностей сообщества и субъектов. (4)

Механизм имитации принятия решений таит в себе возможность рефлексивных конфликтов [9], для их разрешения могут быть использованы нормирующие институты, например, общая культура.

Сделаем два важных замечания: 1) рефлексивное управление сообществом, представленное выражением (4), хорошо укладывается в рамки социокультурной («мультиразумной») модели предприятия; 2) понятие субъект может быть обобщено на макроуровень, т.е. на отдельное предприятие. В этом случае могут быть использованы понятие саморефлексии предприятия (как макросубъекта) и понятие рефлексии предприятия по отношению к другим макросубъектам [10].

Исходя из вышеизложенного, определение рефлексивного управления предприятием сформулируем в следующем виде.

Рефлексивное управление предприятием – создание возможностей для преобразования предприятия в саморефлективную систему, а также для поддержки и развития рефлексивных способностей предприятия, его сообществ и субъектов. (5)

4. Субъектный подход к управлению бизнес-процессами

В последние годы была создана субъектно-ориентированная методология S-BPM (Subject-oriented Business Process Management) [11, 13] и соответствующая инструментальная система Metasonic S-BPM Suite, которые используют субъектный взгляд на управление бизнес-процессами. S-BPM

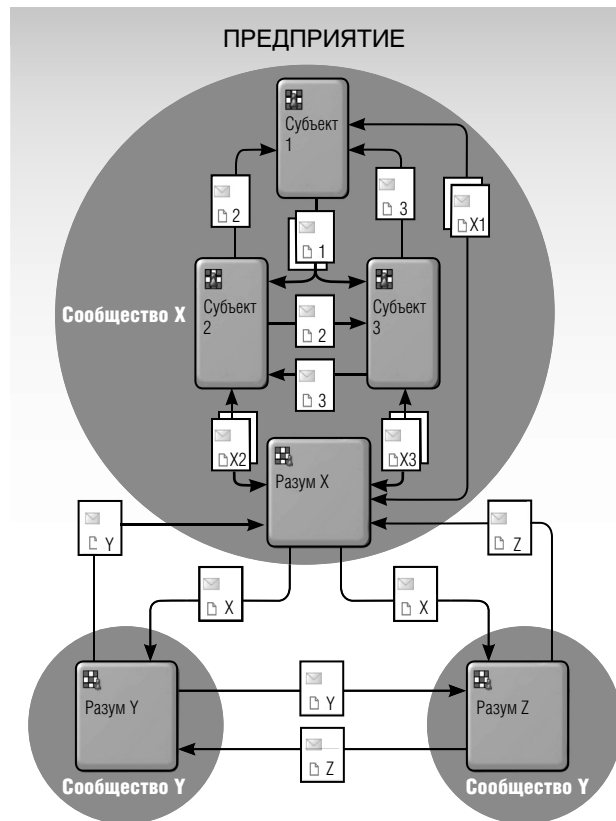


Рис. 3. Пример описания сетевых взаимодействий на предприятии

уже получило международное признание в качестве разновидности BPM. Отношение к человеку, как к субъекту, а не ресурсу, позволило кардинально изменить технологии моделирования и совершенствования бизнес-процессов, получить значительные выгоды при автоматизации бизнес-процессов по сравнению с традиционными методами [12-14].

Но главное преимущество автоматизированной системы класса S-BPM, по нашему глубокому убеждению, заключается в том, что с её помощью уже сегодня можно сделать первые шаги к созданию предприятия XXI века, как социокультурной, «мультиразумной» системы, основанной на главенствующей роли субъектности при выполнении бизнес-процессов.

Для иллюстрации приведем пример модели, описывающей сетевые взаимодействия (обмен сообщениями) субъектов «мультиразумного» предприятия. Эта модель типа «менеджер процесса» построена в инструментальной среде Metasonic S-BPM Suite. Второй тип модели «менеджер субъекта» (дополняющий «менеджер процесса») здесь не приведен, он используется для прямого описания состояний

субъекта и возможных переходов субъекта из одного состояния в другое.

Отметим, что в классическом моделировании также используются состояния, но состояния не субъекта, а системы или среды. Для отражения состояния используются, например, события в eEPC модели процесса (ARIS).

Конечно, описание взаимодействий субъектов в «менеджере субъектов» и описание состояний в «менеджере субъекта» Metasonic S-BPM Suite не предполагает моделирование субъектности (для этого необходимо описывать сознание, познание и рефлексию). Отметим, что фактор субъектности лежит в основе моделирования бизнес-процесса, выполнения и валидации изменений путем имитации или получения реакции других субъектов процесса.

5. Заключение

По мере становления сообществ субъектов (творческих коллективов) на предприятии, субъектно-ориентированные модели начнут играть роль средства планирования деятельности коллективов, постепенно превращая их в саморефлективные системы.

Субъектность и субъектная структура становятся (возможно, уже стали) важнейшим фактором успешной деятельности предприятия. В первую очередь это касается творческих, инновационных предприятий, сознательно движущихся в сторону социокультурной ориентации и «мультиразумности».

Достаточно обоснованным выглядит мнение о том, что реальных успехов в инновационной деятельности можно добиться только на платформе субъектности. ■

Литература

1. Новая философская энциклопедия: В 4-х томах. Том третий / Ред. совет: В.С.Степин, А.А.Гусейнов, Г.Ю.Семигин, А.П.Огурцов и др. — М.: Мысль, 2010.
2. Стерледев Р.К. К вопросу о понятии субъектности. Исторические, философские, политические и юридические науки, культурология и искусствоведение. Вопросы теории и практики. — Тамбов: Грамота, 2011. № 6 (12): в 3-х ч. Ч. II. — С. 160-163.
3. Степин В.С. Философия науки: общие проблемы: учебник для аспирантов и соискателей ученой степени кандидата наук. — М.: Гардарики, 2006.
4. Щедровицкий Г.П. Коммуникация, деятельность, рефлексия // Исследование рече-мыслительной деятельности. — Алма-Ата, 1974.
5. Сорос Дж. Кризис мирового капитализма. Открытое общество в опасности / Пер. с англ. — М.: ИНФРА-М, 1999.
6. Гараедаги Д. Как управлять хаосом и сложными процессами. Платформа для моделирования архитектуры бизнеса. — Минск: Гревцов Букс, 2010.
7. Громов А.И., Чеботарев В.Г. Применение системного подхода к идентификации процессов организации // Информационные технологии в проектировании и производстве, №3. — М: Изд. ФГУП «ВИМИ», 2008.
8. Лепский В.Е. Исходные посылы и онтологии субъектно-ориентированного подхода к инновационному развитию // Рефлективные процессы и управление. — 2007. — №2. — С. 5-28.
9. Лефевр В.А. Конфликтующие структуры. Издание второе, переработанное и дополненное. — М.: Советское радио», 1973.
10. Лепский В.Е., Зорина Г.И. Рефлективное предприятие XXI века // Рефлективные процессы и управление. — 2005. — №2. — С. 21-40.
11. Gartner. Publication Date: 25 July 2011/ID Number: G00214214. Hype Cycle for Business Process Management, 2011.
12. Active Compliance Management with Subject-oriented Business Process Management. On the way to service-oriented business. 2009 S.A.R.L. Martin, 6 rue Paul Guiton, 74000 Annecy, France.
13. Чеботарев В.Г., Громов А.И. Эволюция подходов к управлению бизнес-процессами // Бизнес-информатика. — 2010. — №1 (11). — С. 14-21.
14. Чеботарев В.Г., Бородин Е.Г., Григорьева Д.М. Особенности применения субъектно-ориентированного моделирования бизнес-процессов // Бизнес-информатика. — 2010. — №2 (12). — С. 54-59.

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ТЕОРИИ ПРОИЗВОДСТВА-ПОТРЕБЛЕНИЯ: ОПРЕДЕЛЕНИЕ ВИДА, СТРУКТУРЫ И ПАРАМЕТРОВ МОДЕЛЕЙ

М.А. Ивлев,

*кандидат технических наук, доцент Нижегородского государственного
технического университета им. Р.Е. Алексеева*

E-mail: ivlev-ma@yandex.ru

Адрес: г. Нижний Новгород, ул. Минина, д. 24

В рамках теории производства-потребления товаров и услуг в открытых социально-экономических системах предложены математические модели продукции, описывающие её соответствие запросам потребителя. Основой моделей является конечный оргграф с потоками. Обоснован способ минимального разбиения графа на покрывающие его множества независимых дуг, сформулированы правила вершин. На основе анализа компонентов оргграфа определены целевые функции и необходимые ограничения для решения ряда оптимизационных задач.

Ключевые слова: теория производства-потребления, задачи оптимизации, конечный плоский оргграф с потоками, целевые функции и ограничения.

1. Предметная область и задачи математического моделирования

Как показано в работах [1-3], в задачах управления современными субъектами общественного производства в их состав следует включать как необходимый компонент категорию потребителей, что позволяет ввести в обиход новый класс производственных в широком смысле систем, определенный как системы производства-потребления. Ключевые процессы, определяющие эффективность их функционирования, описываются предложенной автором парадигмой

производства-потребления дифференцированной продукции – концепции, рассматривающей системы производства-потребления как открытые социально-экономические системы [3]. В этих системах управленческие воздействия имеют вид решений, самостоятельно принимаемых индивидуумами – лицами, принимающими решения (ЛПР). Указанная выше парадигма (далее – теория производства-потребления) рассматривает ЛПР двух групп. К первой из них относятся потребители продукции, приобретающие её в соответствии с желаемым качеством жизни, с одной стороны, и

покупательской способностью, с другой стороны. Ко второй группе ЛПР относятся руководители и специалисты производственной подсистемы социально-экономической системы – собственно производственного предприятия, определяющие виды его деятельности, модели продукции, её параметры и объемы выпуска – характеристики, определяемые нами как концептуальные.

Предприятие как хозяйствующий субъект открытой социально-экономической системы будет иметь долговременную экономическую устойчивость и успешно развиваться при условии, что ЛПР второй категории будут действовать с учетом решений, принимаемых и изменяемых ЛПР первой категории. При выполнении этого условия будет реализован актуальный принцип управления производством продукции на основе «голоса потребителя» [4]. Итак, областью функционирования исследуемых систем является социально-экономическая сфера, а их активными компонентами являются ЛПР, которые в той или иной степени обладают свободой выбора и поступают в соответствии с личными предпочтениями.

Развитая теория производства-потребления [2] позволяет на семантическом уровне интуитивно доступно проектировать и производить дифференцированную продукцию, свойства которой наилучшим образом отвечают запросам потребителей. Однако для исключения эвристических ошибок в стратегическом планировании производственной деятельности, в том числе в задачах определения ряда численных оценок перспективной продукции и требуемых ресурсов промышленного предприятия, необходимо применить строгую технологию оптимизации. Постановка и процедура решения оптимизационной задачи в общем случае представляют собой поиск точного верхнего (нижнего или заданного) значения некоторого функционала (во многих случаях – целевой функции) на множестве допустимых решений [5].

Указанная теория производства-потребления и её интерпретация, предложенная автором в виде цепочечных бинарных структур [6], предоставляют возможность сделать первый шаг математического моделирования. Таковым являются выбор и обоснование соответствующего этой структуре вида математической модели, на основе анализа которой на следующих шагах конструируется целевая функция, формируются необходимые ограничения и разрабатывается алгоритм поиска оптимального решения.

2. Выбор вида и структуры математической модели

Для решения этой задачи рассмотрим ключевые особенности неформального описания предметной области [1-3].

Теория производства-потребления оперирует элементами качества жизни потребителя, «переносимыми» продукцией, и определяет вид оптимального решения её разработчика и изготовителя в виде вектора $X=\{x_i\}$, $i = \overline{1, n}$, компоненты которого принимают значения 0 или 1. Каждый компонент решения устанавливает факт *включения* ($x_i=1$) или *игнорирования* ($x_i=0$) конкретного элемента качества жизни потребителя в конечный набор элементов, определенный в работе [7] как профиль качества продукции. Математическая интерпретация семантической теории должна отображать *два варианта решения по каждому его компоненту*.

Другой характерной особенностью теории производства-потребления является вид описываемого ею процесса – процесса *направленного преобразования* множества потенциальных потребителей продукции в несколько непересекающихся множеств.

Количество указанных множеств зависит от принятого числа критериев, влияющих на выбор потребителем того или иного варианта продукции, определенной как дифференцированная продукция [3]. Таких критериев в настоящей работе принято два: качество жизни потребителя, достигаемое потреблением продукции, и стоимость приобретения последней. Следовательно, непересекающимися множествами будут:

- ◆ множество потребителей, удовлетворенных профилем качества и стоимостью продукции с определенным набором элементов качества жизни потребителя;
- ◆ множество потребителей, предпочитающих другой профиль качества продукции, несмотря на доступную стоимость (с поправкой на прибыль производителя и наценку предприятия, реализующего товар на рынке, под стоимостью здесь и далее будем понимать цену товарной продукции – затраты потребителя на её приобретение);
- ◆ множество потребителей, считающих профиль качества продукции соответствующим их запросам, но отказывающихся от её потребления по причине высокой стоимости;

♦ гипотетическое четвертое множество – множество потребителей, которые оценивают профиль качества продукции неудовлетворительным, а стоимость его приобретения – недостижимой, далее принимаем пустым (сокращение набора необходимых потребителю элементов качества жизни приводит к снижению его стоимости, а не к её росту) и в решении поставленных задач далее не учитываем.

В работе [6] предложена адекватная графическая интерпретация указанного направленного преобразования с помощью эквивалентной схемы, позволяющей оценить рыночную востребованность различных профилей качества продукции [7].

Рассмотренные особенности теории производства-потребления, а также показанные в перечисленных работах её практические приложения, делают целесообразным и возможным для решения поставленной оптимизационной задачи применение элементов теории графов. Введение адекватной графовой модели опирается на следующие высказывания.

Отмеченная направленность процесса преобразования множества потребителей обуславливает применение модели вида *ориентированного графа*.

Процессу преобразования множества потребителей поставим в соответствие абстрактное понятие *потока*, протекающего по узлам и ребрам графа, – понятия, используемого, в частности в работе [8]. Поток в графе $N(V, A)$ принято считать целочисленную функцию φ , определенную на множестве дуг A , а целое число $\varphi(a)$ – потоком по дуге $a \in A$. Вершины V различаются по характеру их воздействия на поток φ . Применяемыми в теории графов видами воздействия вершин на поток являются: создание, поглощение и сохранение потока. Соответствующие этим воздействиям вершины называют источниками, стоками и внутренними (промежуточными) вершинами. Указанные понятия и виды воздействий, как показано ниже, не составляют теоретическую платформу, достаточную для решения поставленной задачи оптимизации, однако могут послужить основой для требуемой конкретизации видов воздействий вершин на поток и видов вершин.

Векторная форма области существования возможных решений с конечным числом n бинарных компонентов определяет цепочечную структуру *конечного графа с n звеньями*.

Приведенные положения позволяют конкретизировать вид математической модели продукции (математической абстракции теории производства-потребления) как *конечный плоский цепочечный орграф с потоками*. Возможные варианты путей потоков, соединяющих внешние вершины графа, определяют множество допустимых решений. На *рис. 1* представлен простой пример орграфа (для случая числа компонентов вектора решения $n = 5$), служащий для демонстрации свойств и возможностей «полных» моделей.

Предложенная модель позволяет решить следующие актуальные задачи, в том числе указанные в [8]:

Задача максимизации потока между двумя заданными вершинами при условии, что поток источника конечен.

Задача нахождения потока минимального качества, при котором будет обеспечен сток на одной из вершин не менее заданной величины, при заданной величине потока источника.

Задача определения потоков максимальной стоимости на одной из вершин при заданных потоке источника и стоимости единицы потока.

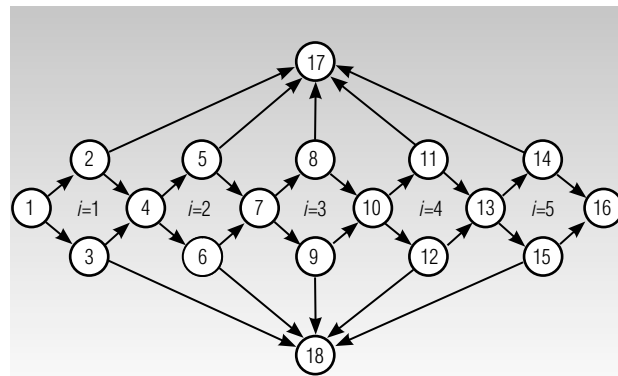


Рис. 1. Модель теории производства-потребления – конечный плоский (двумерный) орграф с потоками (для случая числа компонентов вектора решения $n=5$)

С целью формирования алгоритмов решений перечисленных задач оптимизации на основе введенного конечного плоского орграфа с потоками необходимо описать его компоненты и сформулировать для него правила вершин и дуг.

3. Определение свойств и характеристик компонентов модели

Эту задачу целесообразно решить путем разбиения дуг и вершин графа на множества, характери-

зующиеся определенным структурным типом. Ряд важнейших практических задач состоит в разбиении дуг (ребер) и вершин графа на наименьшие числа соответственно дуг и вершин [8].

3.1. Разбиение дуг

Задачей этого процесса является определение минимальных рёберных разбиений, покрывающих ориентированный граф. Известен алгоритм разбиения дуг в графах без потоков [8], однако его применение в данном случае даёт неверное решение.

Последнее обстоятельство обусловлено тем, что в известном алгоритме рёберных разделений не учтены особенности ряда дуг предложенной графовой модели. Особенности состоят в том, что не все дуги орграфа (рис. 1), положительно или отрицательно инцидентных одной вершине, представляют независимые варианты путей движения потока. Указанные дуги определим как альтернативные. К ним относятся все дуги широко распространенных сетевых моделей, и на «альтернативном» понимании дуг основаны такие методы моделирования, как метод критического пути, ПЕРТ и другие. Такими дугами в приведенном примере являются, в частности, пара дуг $a_{1,2} \cong (1,2)$ и $a_{1,3} \cong (1,3)$. В графах с потоками часть дуг не представляет альтернативных направлений и, следовательно, не увеличивает количества возможных путей движения потока по графу. К подобным дугам на орграфе (рис. 1) относятся, в частности, пара дуг $a_{2,4} \cong (2,4)$ и $a_{2,17} \cong (2,17)$. Для них невозможным является выбор пути либо по дуге $a_{2,4}$, либо по дуге $a_{2,17}$ – путь по ним будет хотя и разветвленный, но один. Таким образом, несколько «неальтернативных» дуг (назовем их блокированными дугами) по числу вариантов путей эквивалентны одной традиционной дуге. Тогда расчет числа r минимальных разбиений дуг орграфа должен выполняться по следующим выражениям

$$r = k - \sum_{v \in C} (M_v - 1) \delta^-(v) \quad (1)$$

$$k = \sum_{v \in A} [\delta^+(v) - \delta^-(v)] = \sum_{v \in B} [\delta^-(v) - \delta^+(v)] \quad (2)$$

Здесь k – количество разбиений без учета блокировки дуг [8], $\delta^+(v)$ и $\delta^-(v)$ – количества дуг, соответственно положительно и отрицательно инцидентных вершине v ; A и B – множества вершин орграфа, таких что

$$A = \{v \in V | \delta^+(v) > \delta^-(v)\}, \quad (3)$$

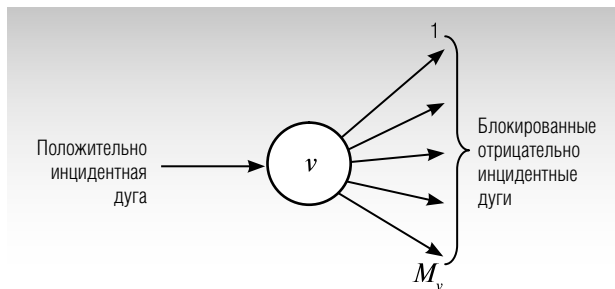


Рис. 2. Блокированные дуги $\{1, \dots, M_v\}$ вершины v орграфа с потоками

$$B = \{v \in V | \delta^+(v) < \delta^-(v)\}, \quad (4)$$

множество $C \subseteq B$ – множество вершин, имеющих отрицательно инцидентных вершине $v \in D$ блокированные дуги a_1, \dots, a_{M_v} ; M_v – количество инцидентных вершине v блокированных дуг. Второе слагаемое в выражении (1) учитывает сокращение вариантов путей потоков в орграфе с потоками по сравнению с сетевыми моделями. Потоки всех блокированных дуг вершины v представляют поток положительно инцидентной ей дуги, расщепленный на части (рис. 2).

Расчет минимального числа разбиений дуг, выполненный по выражениям (1-4) для примера, приведенного на рис. 1, дает следующие результаты: множества вершин:

$$A = \{16, 17, 18\}, \quad B = \{1, 2, 3, 5, 6, 8, 9, 11, 12, 14, 15\}, \\ C = \{2, 3, 5, 6, 8, 9, 11, 12, 14, 15\};$$

количество блокированных дуг для каждой вершины $M_v = 2, v \in C$; количество дуг, отрицательно инцидентных вершинам $v - \delta^-(v, v \in C) = 10$; количество разбиений без учета блокировки дуг $k = 12$, количество разбиений с учетом блокировки дуг $r = 2$. На рис. 3 приведено собственно минимальное

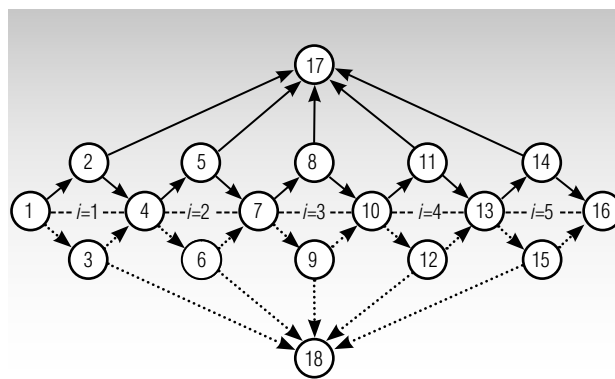


Рис. 3. Минимальное разбиение дуг орграфа с потоками, приведенного на рис. 1

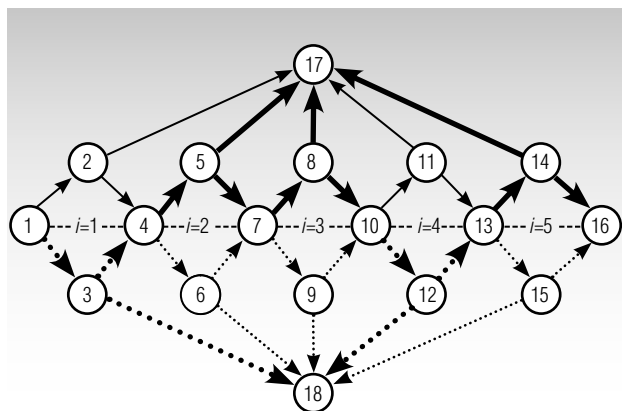


Рис. 4. Вариант путей потоков орграфа (рис. 1), соответствующий решению $X' = \{0, 1, 1, 0, 1\}$

разбиение дуг. Штриховая линия (назовем ее *экватор графа*) делит плоскость графа на две полуплоскости: верхняя соответствует компонентам $x_i = 1$, нижняя – $x_i = 0$.

Проведенное минимальное разбиение дуг орграфа с потоками дает наглядное представление о характере решений, принимаемых ЛПР в производстве продукции. Эти решения («включать» – $x_i = 1$ или «не включать» – $x_i = 0$ компонент качества жизни потребителя в профиль продукции) будут формировать траекторию движения потока по орграфу. Один из вариантов решений по выбору профиля качества продукции, вектор X' которого имеет компоненты $X' = \{0, 1, 1, 0, 1\}$, выделен на рис. 4 жирными дугами.

3.2. Разбиение вершин

Ряд задач оптимизации, решаемых на основе графовых моделей, связан с разбиением вершин графа на наименьшее количество множеств независимых вершин – множеств, не содержащих смежных вершин (задачи раскраски графа). В таких задачах указанное наименьшее количество независимых

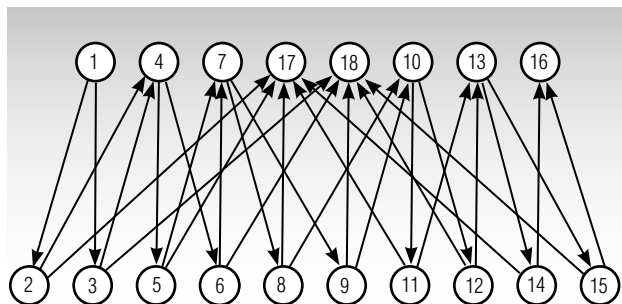


Рис. 5. Иллюстрация свойства двудольности орграфа с потоками

множеств называют хроматическим числом графа. Нетрудно убедиться, что хроматическое число введенного орграфа равно двум, а сам орграф является двудольным (рис. 5).

Приведенные варианты представления введенного орграфа позволяют выделить группы вершин с характерными структурными свойствами и сформулировать правила вершин для каждой из этих групп. Вершины классифицируются по их воздействию на поток. Для поставленной задачи поиска оптимальных решений на основе введенной графовой модели эти воздействия связаны с компонентами решений $x_i, i = \overline{1, n}$, следовательно, правила вершин должны содержать значения x_i как независимые параметры.

3.2.1. Внутренние вершины

а) Вершины с альтернативными потоками (alternative flows)

Обозначим такие вершины вершинами *A*-типа. К ним на примере рис. 1 относятся вершины 4, 7, 10, 13. Фрагмент плоского орграфа, содержащий вершину *A*-типа и инцидентные ей дуги и показанный на рис. 6, разделен вертикальной и горизонтальной штриховыми линиями на две соответствующие пары полуплоскостей. Левая по-

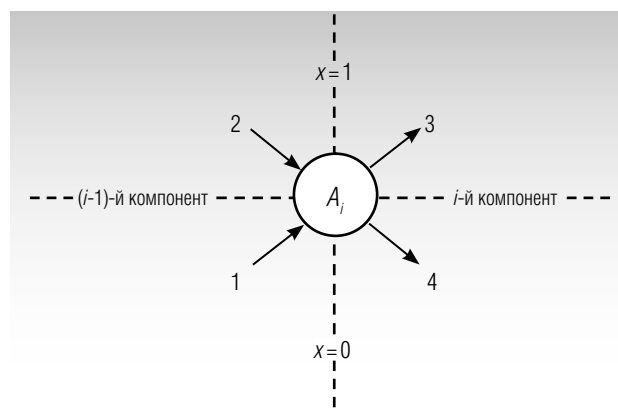


Рис. 6. Вершина *A*-типа с инцидентными ей дугами 1-4

Таблица 1.

Правила (логика) *A*-вершины

Компоненты решения		Потоки в дугах			
x_{i-1}	x_i	1	2	3	4
0	0	$\varphi(1)$	0	0	$\varphi(1)$
0	1	$\varphi(1)$	0	$\varphi(1)$	0
1	0	0	$\varphi(2)$	0	$\varphi(2)$
1	1	0	$\varphi(2)$	$\varphi(2)$	0

луплоскость соответствует $(i-1)$ -ому компоненту вектора решения, правая – i -ому. Верхняя полуплоскость фрагмента графа соответствует значению $x_i=1$, нижняя – $x_i=0$.

Правила вершин A -типа (связь входящих и исходящих потоков) сформулируем в виде *табл. 1*. Потоки $\varphi(1)$ и $\varphi(2)$ в дугах 1 и 2 не зависят от состояния вершины A_i . Потоки $\varphi(3)$ и $\varphi(4)$ в дугах 3 и 4 зависят как от потоков $\varphi(1)$ и $\varphi(2)$, т.е. от значения x_{i-1} , так и от состояния вершины A_i (т.е. от значения x_i).

б) Вершины, разветвляющие (делящие) входные потоки

Обозначим такие вершины вершинами D -типа (*Dividing*). Вершина D -типа с инцидентными ей дугами показана на *рис. 7*, здесь дуги 2 и 3 являются заблокированными.

К этому типу вершин в примере *рис. 1* относятся вершины 2,3,5,6,8,9,11,12,14,15. Правила вершин для вершины D -типа i -го компонента решения представлены в *табл. 2*. Вершина D_i разделяет проходящий через нее поток φ на два потока: поток $p_i\varphi$ одной отрицательно инцидентной дуги и поток $(1-p_i)\varphi$ второй отрицательно инцидентной дуге. Величину p_i можно трактовать как *вес дуги 2*, а величину $(1-p_i)$ – как *вес дуги 3* (*рис. 7*).

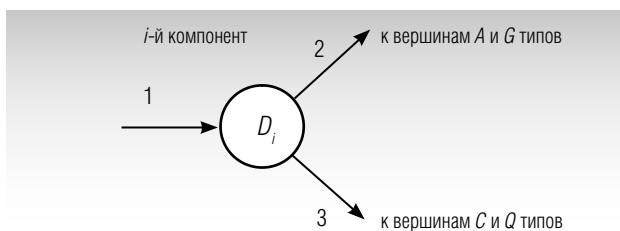


Рис. 7. Вершина D -типа i -го компонента модели с инцидентными дугами 1-3

Таблица 2.

Правила D -вершины

Коэффициент передачи потока	Потоки в дугах		
	1	2	3
p_i	$\varphi(1)$	$p_i\varphi(1)$	$(1-p_i)\varphi(1)$

3.2.2. Внешние вершины

а) Вершина без входящих потоков – исток (source).

Обозначим такие вершины вершинами S -типа. К этому типу вершин в примере *рис. 1* относится вершина 1. Для вершин без входных потоков будем указывать величину исходящего потока как *мощ-*

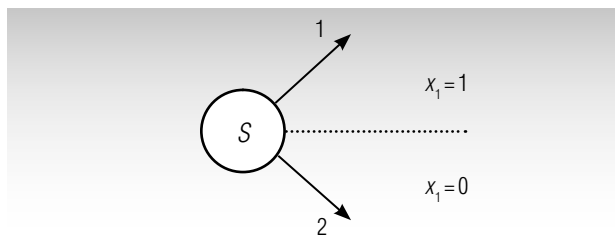


Рис. 8. Вершина S -типа ($i=1$)

Таблица 3.

Правила S -вершины

Мощность вершины S	Компонент решения x_i	Потоки в альтернативных дугах	
		1	2
$\varphi(S)$	0	0	$\varphi(S)$
	1	$\varphi(S)$	0

ность вершины $\varphi(S)$. Вершина S -типа показана на *рис. 8*, где отрицательно инцидентные ей альтернативные дуги 1 и 2 соответствуют двум вариантам решения, соответственно $x_1=1$ и $x_1=0$. Правила вершин для неё приведены в *табл. 3*.

б) Вершины без исходящих потоков

Вершина с альтернативными входными потоками – сток (Gutter). Обозначим такие вершины вершинами G -типа. К ним на примере *рис. 1* относится вершина 16. Вершина G -типа показана на *рис. 9*, где инцидентные ей дуги 1 и 2 соответствуют двум вариантам n -го компонента решения, соответственно $x_n=1$ и $x_n=0$. Правила вершин для неё приведены в *табл. 4*. Отметим очевидный факт, что эта вершина не оказывает воздействие на потоки в дугах 1 и 2, и наоборот, потоки в дугах 1 и 2, а также компонента решения x_n определяют её мощность $\varphi(G)$.

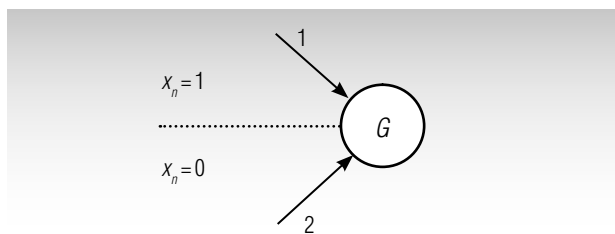


Рис. 9. Вершина G -типа с инцидентными дугами 1 и 2

Таблица 4.

Правила G -вершины

Компонент решения x_n	Потоки в альтернативных дугах		Мощность вершины G $\varphi(G)$
	1	2	
1	$\varphi(1)$	0	$\varphi(1)$
0	0	$\varphi(2)$	$\varphi(2)$

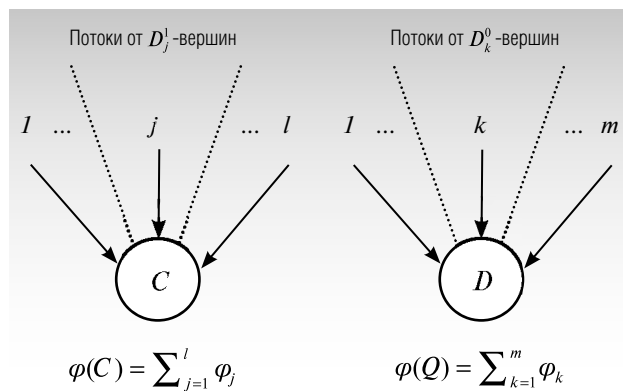


Рис. 10. Вершины C и Q-типов с положительно инцидентными дугами и соответствующие правила вершин: $x_j = 1, j=1, \dots, l; x_k = 0, k=1, \dots, m; l+m=n$

Вершины с объединяемыми входными потоками. Такие потоки будут определяться либо компонентами $x_i = 1$, либо компонентами $x_i = 0$. Обозначим вершины, соответствующие $x_i = 1$, C-вершинами и вершины, соответствующие $x_i = 0$ вершинами Q-типа. На примере рис. 1 вершиной C является внешняя вершина 17, а вершиной Q – вершина 18. Вершины C- и Q-типов и соответствующие им правила показаны на рис. 10.

Вершина C-типа суммирует потоки дуг, отрицательно инцидентных вершинам D_j^1 - D-вершинам с ненулевой мощностью, находящимся в верхней полуплоскости орграфа (на примере рис. 1 такими вершинами являются вершины 2,5,8,11,14). Вершина Q-типа суммирует потоки дуг, отрицательно инцидентных вершинам D_j^0 - D-вершинам с ненулевой мощностью, находящимся в нижней полуплоскости орграфа (на примере рис. 1 такими вершинами являются вершины 3,6,9,12,15). Орграф, представленный на рис. 1 с учетом обозначений вершин введенных структурных типов, принимает

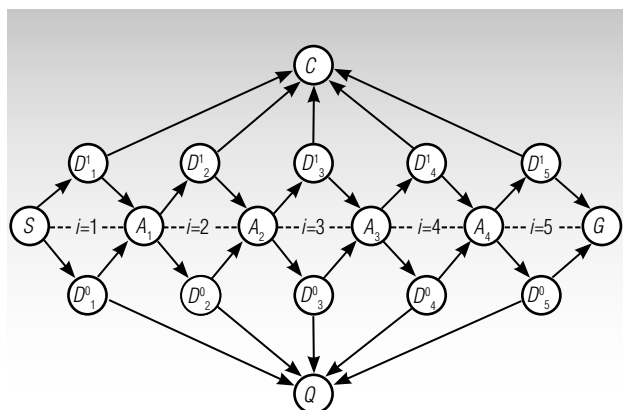


Рис. 11. Орграф с размеченными типами вершин

вид (рис. 11). Правила вершин, размеченных здесь, приведены в табл. 1-4 и на рис. 10.

4. Целевые функции и ограничения

Модель (рис. 11), приведенные правила для её вершин (табл. 1-4, рис. 10) и представление её ориентированных реберных покрытий (рис. 3) позволяют сформулировать целевые функции и ограничения для различных практических приложений.

4.1. Задача максимального потока, идущего к вершине G

Решается прикладная задача формирования профиля качества продукции с максимальной долей рынка потребителей (максимальным количеством потребителей при фиксированном объеме рынка P) – потребителей, удовлетворенных её качеством и ценой. В отличие от традиционного целочисленного понимания потока, потоком будем считать любое неотрицательное вещественное число, означающего часть целого, количество людей, стоимость (цену) продукции и др. В первом варианте данной задачи поток в компонентах графа – доля потока потребителей, идущего от вершины S к внешним вершинам и принятого за единицу, т.е. $0 \leq \varphi_i \leq 1, i = 1, n$.

Воспользуемся логикой постановки и решения одной из задач дискретного программирования – задачи о назначениях. Расширим введенное ранее понятие профиля качества продукции. Будем считать, что в профиль качества продукции назначаются i-тые элементы двух видов: собственно элемент качества (индекс $j=1$) и «отсутствие» такового ($j=2$). Введем соответствующие переменные x_{ij} и p_{ij} :

$$x_{ij} = \begin{cases} 1, \text{ если в профиль назначен } ij\text{-ый элемент} \\ 0, \text{ в противном случае} \end{cases} \quad (5)$$

p_{ij} – коэффициент передачи потока i-той D-вершиной для ij-го элемента решения, $0 \leq p_{ij} \leq 1, j=1, 2; i = 1, n$.

Математическая модель задачи формирования профиля качества продукции, привлекательной и доступной для максимальной доли рынка потребителей $C(X)$ имеет следующий вид:

$$C(X) = \prod_{i=1}^n \sum_{j=1}^2 p_{ij} x_{ij} \rightarrow \max; \quad (6)$$

при условии:

$$\sum_{j=1}^2 x_{ij} = 1, i = \overline{1, n} \quad (7)$$

Здесь вектор $X = \{x_{ij}\}$ – вектор решений производителя по профилю качества продукции, величина p_{i1} – доля потребителей, не учитывающих при оценке привлекательности продукции наличие i -го элемента в профиле её качества, величина p_{i2} – доля потребителей, считающих доступным приобретение продукции с i -тым элементом качества.

Вариант рассматриваемой задачи о максимальном количестве потребителей P^+ дополняется выражением:

$$P^+ = PC(X) = \varphi(S) C(X). \quad (8)$$

Здесь мощность истока графовой модели $\varphi(S)$ имеет смысл количества P потенциальных потребителей данной продукции.

Выражения (5) и (6) показывают, что данная задача относится к классу задач нелинейного целочисленного программирования. Наиболее распространенным способом их решения является метод «ветвей и границ».

4.2. Задача максимальной стоимости потока, идущего к вершине G

Решается прикладная задача формирования профиля качества продукции, при котором достигается максимальный доход производителя. Математическая модель данной задачи имеет вид:

$$C(X) = P^+ [C_0 + \prod_{i=1}^n c_i \sum_{j=1}^2 p_{ij} x_{ij}] \rightarrow \max, \quad (9)$$

$$\sum_{j=1}^2 x_{ij} = 1, i = \overline{1, n} \quad (10)$$

Здесь C_0 – цена базовой комплектации продукции (постоянная составляющая цены, не зависящая от элементов качества), c_i – цена i -го элемента.

4.3. Задача минимизации потока к вершине G до заданного порога при нулевом потоке к вершине C

Решается однокритериальная задача формирования минимального профиля качества продукции с заданным уровнем предпочтения (с долей рынка

потребителей, не удовлетворенных качеством продукции, не более $\varphi_0(Q)$). Математическая модель данной задачи имеет вид:

$$C(X) = \prod_{i=1}^n \sum_{j=1}^2 p_{ij} x_{ij} \rightarrow \min, \quad (11)$$

$$p_{i1} = 1, \sum_{j=1}^2 x_{ij} = 1, i = \overline{1, n} \quad (12)$$

при ограничении

$$\prod_{i=1}^n \sum_{j=1}^2 p_{ij} x_{ij} \geq 1 - \varphi_0(Q). \quad (13)$$

4.4. Задача минимизации потока к вершине G до заданного порога при нулевом потоке к вершине Q

Решается однокритериальная задача формирования минимального профиля качества продукции, соответствующего доле её потребителей, не удовлетворенных стоимостью продукции, не более $\varphi_0(C)$. Симметрия графа относительно его экватора позволяет использовать в качестве целевой функции выражение (11), в котором величины p_{i2} и x_{ij} удовлетворяют условию:

$$p_{i2} = 1, \sum_{j=1}^2 x_{ij} = 1, i = \overline{1, n} \quad (14)$$

Задача решается при следующем ограничении:

$$\prod_{i=1}^n \sum_{j=1}^2 p_{ij} x_{ij} \geq 1 - \varphi_0(C). \quad (15)$$

5. Выводы

Обоснована математическая модель семантической теории производства-потребления продукции в открытых социально-экономических системах – плоский орграф с потоками.

Предложена методика расчета минимального реберного разбиения, учитывающая особенности формирования вариантов путей в орграфе с потоками по сравнению с традиционным разбиением ребер в сетевых моделях.

Дана классификация вершин орграфа и сформулированы их правила, положенные в основу расчетов потоков модели во внешние вершины различных видов.

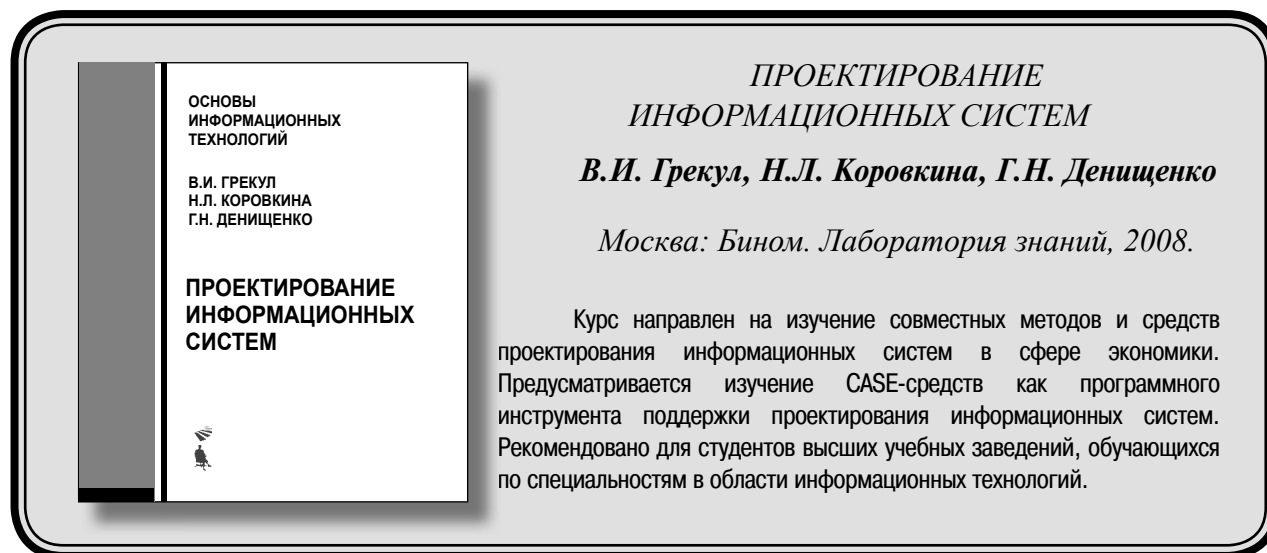
Сформулированы целевые функции и ограничения, необходимые и достаточные для поиска решений, оптимальных по ряду критериев.

Прикладные задачи расчета характеристик социально-экономических систем относятся к классу задач нелинейного целочисленного программирования (целевая функция нелинейная, ограничения на решения – линейные/нелинейные).

Введенная математическая модель может при необходимости быть развита на область задач с количеством факторов, определяющих целевую функцию, более двух. При этом граф будет не плоским, а M -мерным (M – число факторов). ■

Литература

1. Ивлев М.А. Инвариантные аспекты производственных систем в задаче адаптивного управления // Организатор производства. – 2009. – №4. – С. 84-89.
2. Ивлев М.А. Открытые производственные системы // Перспективы науки. – 2010. – №2 (04). – С. 120-124.
3. Ивлев М.А. Парадигма производства-потребления дифференцированной продукции // Перспективы науки. – 2010. – №5 (07). – С. 109-113.
4. Перспективное планирование качества продукции и план управления. APQP. Ссылочное руководство / Пер. с англ. – Н.Новгород: ООО СМЦ «Приоритет», 2010.
5. Системный анализ и принятие решений: Словарь-справочник: Учебное пособие для вузов / Под ред. В.Н.Волковой, В.Н.Козлова. – М.: Высшая школа, 2004.
6. Ивлев М.А. Бинарные цепочечные структуры как модели дифференцированной продукции // Бизнес-информатика. – 2010. – № 4 (14). – С. 21-28.
7. Ивлев М.А. Расчет концептуальных параметров открытых производственных систем на основе бинарных моделей // Организатор производства. – 2010. – № 3. – С. 9-14.
8. Басакер Р., Саати Т. Конечные графы и сети. – М.: Наука, 1973.



ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ОРГАНИЗАЦИИ МАССОВЫХ МЕРОПРИЯТИЙ (НА ПРИМЕРЕ XXII ЗИМНИХ ОЛИМПИЙСКИХ ИГР 2014)

В.К. Абросимов,

доктор технических наук, старший научный сотрудник,
начальник отдела компании «Стэп Лоджик»

В.В. Лебидько,

кандидат физико-математических наук,
ведущий аналитик компании «Стэп Лоджик»

E-mail: v_abrosimov@step.ru, v_lebidko@step.ru

Адрес: г. Москва, ул. Полковая, д. 3, стр. 3

Рассмотрены основные положения моделирования организации массовых мероприятий с использованием многоагентной технологии. Выделены характерные особенности массовых мероприятий, показана практическая значимость моделирования. Посетители массовых мероприятий и обслуживающий персонал представлены как агенты, проявляющие в процессе достижения целей свойство коллективного поведения с учетом собственной автономности. На примере модели посещения спортивного мероприятия на Олимпийском стадионе в г. Сочи продемонстрированы возможности имитации последствий возникновения различных критических ситуаций.

Ключевые слова: моделирование, массовое мероприятие, многоагентный подход, Олимпийские игры, Сочи, очередь, толпа.

1. Общая задача моделирования подготовки и проведения массовых мероприятий

Во многих практических случаях возникают задачи оперативного обслуживания значительного количества людей ограниченной группой обслуживающих лиц. Как правило, такие задачи связаны с организацией целенаправленного

управления поведением больших групп людей. Но при объединении людей в группы спонтанно возникает эффект так называемого коллективного поведения — образуются формальные и неформальные лидеры, проявляется эффект толпы, повышается влияние внутренних социальных установок и др. В силу многочисленных особенностей процессы моделирования поведения значительных групп людей очень редко могут быть описаны в строгой ма-

тематической форме. Такие попытки предпринимались (см, например, [1]), однако любые строгие закономерности должны находиться в соответствии с адекватными реальности законами поведения групп людей, что сделать достаточно проблематично в силу общей неформализуемости таких процессов.

К числу основных задач моделирования, возникающих при организации массовых мероприятий, следует отнести имитацию прохождения потоков людей через различные ограничители движения (препятствия). К числу таких препятствий можно отнести металлорамки, различные турникеты, даже обычные двери (например, при входе в метрополитен) и др. К классу своего рода препятствий при моделировании можно отнести и кассы для продажи билетов, где посетители мероприятий должны эти билеты приобретать, полицейских, регулирующих (направляющих) поток людей при его движении. При проведении таких массовых мероприятий, как XXII зимние Олимпийские игры 2014 в г. Сочи (далее, Олимпийские игры), для оказания помощи посетителям активно будут привлекаться волонтеры, в этой связи их нахождение на территории Олимпийского парка также будет являться препятствием, замедляющим движение посетителя к цели, которое также можно моделировать.

Наряду с моделированием препятствий целесообразно также рассматривать и моделировать классы событий, возникновение которых может существенно изменить как порядок проведения мероприятия, так и направление движения людей. К наиболее типичным событиям можно отнести внезапный отказ от проведения мероприятия (перенос сроков проведения или полная отмена), прекращение мероприятия в процессе проведения, последствия террористических актов и др. Эти задачи относятся к классу задач «моделирование толпы», в результате решения которых получены достаточно интересные результаты [2 - 4 и др.].

Выделим некоторые характерные особенности процессов организации массовых мероприятий:

- ♦ посетители массовых мероприятий должны иметь некоторые удостоверяющие их намерения документы (например, билет на спортивное соревнование), позволяющие пройти на мероприятие с обеспечением требований по безопасности;

- ♦ отсутствие статистических характеристик поступления заявок на обслуживание в практических задачах приводит к необходимости рассматривать данные процессы как, с одной стороны, случайные

с неизвестным законом распределения, а с другой – как, своего рода, нечеткие;

- ♦ в процессе движения к цели посетитель может столкнуться с рядом ограничений, связанных, главным образом, с ограниченными возможностями ответственных лиц обслужить потоки с высокой интенсивностью поступления «заявок» или с конструктивными особенностями помещений;

- ♦ в процессе подготовки и проведения мероприятия возможно возникновение некоторых событий, существенно меняющих стратегии поведения посетителей.

Основным методом исследования ситуаций организации массовых мероприятий в этом контексте является имитационное моделирование. Наиболее теоретически обоснованной и практически используемой в задачах, сущность которых составляют процессы обслуживания клиентов, считается теория массового обслуживания. При этом посетителей массовых мероприятий можно рассматривать как своего рода «заявки» на обслуживание, а турникеты, металлорамки, кассиров, волонтеров и полицейских можно рассматривать как «приборы», обрабатывающие «заявки». Массивы «заявок», находящиеся в ожидании обработки «прибором» («приборами») образуют «очереди». Время «жизни» «очереди» определяется скоростью обработки «заявки» «прибором» и своевременным покиданием «заявки» данной конкретной «очереди». Момент покидания «очереди» определяется фактом полного обслуживания «заявки». В случае если «заявка» обработана «прибором» не полностью, возникает накопление «очереди», последствия влияния которой на моделируемую систему определяются связью с остальными элементами системы. Тем самым, основными параметрами несущими элементы недетерминированности и случайности в построенной модели являются поток (плотность и скорость потока) «заявок» (посетителей Олимпийских игр) и продолжительность (время) их «обслуживания».

2. Агентный подход к моделированию подготовки и проведения массовых мероприятий

В настоящей работе предполагаются реализуемыми две основных гипотезы.

Первая гипотеза. Группы движущихся к единой цели людей имеют возможность активно обмениваться информацией друг с другом относительно изменений окружающей среды, своего целеполагания и поведения.

Вторая гипотеза. Поведение групп людей, движущихся к единой цели, определяется не заранее определенным планом, а управляется происходящими в процессе движения событиями.

По существу, обмен сообщениями между движущимися к цели людьми изменяет их поведение, влияя как на состоявшееся событие, участниками которого они являются, так и на ещё не наступившее. Анализируя условия решения таких задач, можно прийти к выводу, что для их решения необходимо использовать понятия, которые находятся как бы «над» традиционными подходами к их решению. Как показывает исследование различных парадигм представления и описания поведения активных объектов в изменяющихся внешних средах (адаптивные, самоорганизующиеся системы, конечные автоматы и др.), целесообразно для моделирования таких ситуаций использовать парадигму агентно-ориентированных систем, использующих понятие «интеллектуальный агент», как высокоуровневую абстракцию для формализации сложных недоопределенных ситуаций [5]. При этом организацию коллективного движения целесообразно описывать как перемещение в пространстве так называемых «интеллектуальных агентов».

Интеллектуальные агенты — есть новый класс некоторых программно-аппаратных сущностей, которые действуют от имени пользователя, чтобы находить и обрабатывать информацию, поддерживать решение трудных задач, вести самостоятельные переговоры в программных системах, автоматизировать выполнение рутинных операций и сотрудничать с другими программными агентами при возникновении сложных проблем, снимая тем самым с человека избыточную информационную нагрузку [5-8].

В нашем случае под агентом мы будем понимать лицо, стремящееся посетить массовое мероприятие (посетителя), обладающего возможностью оценивать возникающую ситуацию, взаимодействовать с другими посетителями и принимать самостоятельные решения. Зафиксируем также способность агента перемещаться в пространстве в соответствии с собственными целями, критериями и предпочтениями, в зависимости от складывающейся ситуации в соответствии с информацией, имеющейся в собственной базе знаний, полученной от других агентов, либо с учетом их мнения.

Задача такого агента состоит в том, чтобы:

- ◆ двигаясь в группе себе подобных агентов,
- ◆ по заданной начальной позиции (прибытие в район стадиона, например),

- ◆ конечной цели (проход на мероприятие),
- ◆ при использовании собственной базы знаний об окружающей среде (необходимость иметь билет, иные документы),
- ◆ потенциальных пространственно-временных ограничений (необходимо пройти через металлодетекторы и турникеты в условиях, когда время до начала мероприятия ограничено),
- ◆ с учетом мнения других агентов, выполняющих аналогичные задачи (например, информации об отмене мероприятий, образовании очереди, предложениях о покупке/продаже билетов и др.)

построить оптимальную пространственно-временную траекторию движения к цели, избежав возможных очередей.

Будем считать, что движение агента не зависит от его ориентации, пройденного пути, но существенно зависит от возможных конструктивных и организационных ограничений на движение. Математической формой агента может быть просто точка в четырехмерном пространстве Минковского (определяющем 3 пространственные координаты и одну временную), для простоты моделирования преобразованном в трехмерное пространство. В качестве определяющей концепции поведения агента будем закладывать концепцию рациональности. Рациональный агент действует «правильно», то есть с максимальной эффективностью. Оценка же рациональности зависит от показателей эффективности его движения (прежде всего вероятности достижения цели), его знаний о среде (прежде всего, ограничений на движение), оценки текущей ситуации (прогноза последствий) и собственных решений. При этом агент выбирает решения на основе только текущей ситуации, не игнорируя предварительно закладываемые в базу знаний данные, использования on-line информации о состоянии среды.

3. Посетитель Олимпийских игр, как интеллектуальный агент

Примером такой задачи может являться обслуживание гостей Олимпийских игр в период подготовки и проведения спортивных мероприятий. Априори ясно, что создание разнообразных моделей (Олимпийского парка, Олимпийской деревни, трафика в информационных сетях, транспорта и др.) занимает важное место в общих процессах подготовки и проведения игр. Так, например, при проведении Летних Олимпийских игр в Лондоне (2012 год) по данным средств массовой информации, ор-

ганизаторы, построив необходимые компьютерные модели с мощными средствами визуализации, отвели моделированию сотни тысяч часов (разумеется, при пересчёте модельного, значительно меньшего времени, фигурирующего в имитационных экспериментах в реальные астрономические часы).

Организаторы Игр в Российской Федерации уделяют большое внимание удобству посетителей Игр: наличию возможности как заблаговременно приобрести электронный или обычный билет, получить его по результатам аукциона, приобрести его перед началом соревнования. При этом значительное внимание уделяется вопросам безопасности. Так, организаторы Игр в Сочи для этих целей разрабатывают так называемую единую идентификационную систему (ЕИС), которая призвана не только обеспечить совершенно новый уровень технологических сервисов для клиентов олимпийского проекта, но также и решать вопросы безопасности, разграничения доступа, получения электронных билетов, навигации, получения информации об Играх и др. [2]. Очевидно, что в рамках реализации ЕИС должно быть предусмотрено создание необходимого количества центров выдачи карт ЕИС (в настоящее время имеет название паспорт болельщика), размещаемых в удобных для зрителей местах. Заметим при этом, что время выдачи таких карт существенно превышает время приобретения билета, так как связано с разнообразными временными процедурами.

Описание элементов системы *посетитель Игр – персонал, обслуживающий Игры – олимпийский объект* и их взаимодействий друг с другом выглядит следующим образом. Все посетители массового спортивного мероприятия образуют по пути на олимпийский объект квазинепрерывный поток. Объектами моделирования являются группы людей, следующих на мероприятие (посетители) и обслуживающих его. При этом выделяются следующие два корневых класса активных объектов.

Первый класс – *посетители Игр*, в котором есть условно представители трех групп:

Группа K_{11} , имеющих билет и карту ЕИС;

Группа K_{12} , имеющих билет, но не имеющие карту ЕИС;

Группа K_{13} , не имеющих и билета, и карты ЕИС.

Второй класс – *обслуживающие Игры лица*, в котором также есть условно представители трех групп:

Группа K_{21} – волонтеры;

Группа K_{22} – кассиры;

Группа K_{23} – сотрудники службы ЕИС.

В наиболее общем случае i -го агента j -го класса можно определять вектором состояния вида:

$$A_{ij} = \{B_{ij}, G_{ij}, S_{ij}, N_{ij}, U_{ij}, E_{ij}\}, i=1, j=1, 2$$

где j – номер корневого класса;

B_{ij} – база знаний i -го агента;

G_{ij} – множество целей (желаний) агента;

S_{ij} – множество стратегий поведения агента;

N_{ij} – структура намерений агента (подмножество множества S_{ij} ($N_{ij} \subset S_{ij}$), то есть совокупность стратегий поведения, выбираемых i -тым агентом для достижения цели gs из множества G_{ij});

U_{ij} – структура обязательств агента (подмножество множества G_{ij} ($U_{ij} \subset G_{ij}$), то есть совокупность целей gk из множества G_{ij} , которые могут корректироваться i -тым агентом по результатам взаимодействия с другими агентами для достижения общей цели);

E_{ij} – многомерная матрица, определяющих описание внешних связей i -того агента с другими агентами j -того класса, включая их состояние, цели, намерения и др.

Каждый агент из соответствующего корневого класса имеет собственные значения параметров, собственное поведение, функционирует с одной стороны независимо от других объектов и экземпляров, имея цель (попасть на соревнование), с другой – зависимо от внешней среды и эффекта «толпы». Вместе с тем, многие характеристики и поведение агентов группы в выбранном классе одинаковы по своей сути, так как цели преследуются по существу одинаковые. Параметрами агента в данной постановке является только его принадлежность к определенному классу, из чего вытекает стратегия его поведения.

Модельное время здесь рассматривается как непрерывное. Для спецификации действий агента используются таймеры, запускаемые на определенный интервал времени и по окончании этого интервала выполняющие определенное действие. Так, по таймеру в модели предусматривается фиксация некоторого события (например, начало продажи билетов на соревнование). Таймеры в модели также запускают изменение поведения людей в зависимости от времени начала мероприятия (например, за 15 минут до начала мероприятия скорость движения людей увеличивается).

Выше отмечалось, что агенты могут взаимодействовать с окружением – другими агентами своего класса или с агентами других классов. Агенту из среды, в которой он определен, можно послать сообщение, и по-

лучить ответ, что оно получено. Это взаимодействие определяется так называемыми стейчартами. Они в данной постановке определяют действия агента (его поведение) во время посещения спортивного мероприятия, правила его перемещения или переходы в другое состояние (состояние имеющих и не имеющих: билетов на Игры, карт ЕИС и др.) во времени под воздействием событий и условий.

Рассмотрим свойства посетителя массового мероприятия Игр, как интеллектуального агента.

1. База знаний агента (только по вопросам прохода на стадион) содержит заранее полученную информацию, определяющую его поведение:

- ◆ необходимо получить карту ЕИС, в противном случае на Игры посетители не допускаются;
- ◆ можно задавать вопросы волонтерам и полиции, по ответам которых корректировать свое поведение;
- ◆ при отсутствии очередей и форс-мажорных обстоятельств требуемое время получения необходимых документов и прохода на стадион не превышает 20 минут.

2. Множество целей агента вырождается в 4 элемента: купить билет, получить карту ЕИС, продать лишний билет, пройти на соревнование.

3. Множество стратегий поведения агента вырождается в следующие производственные правила:

- ◆ если нет билета, найти кассу и купить его;
- ◆ если нет карты ЕИС, то найти пункт выдачи карт ЕИС и получить карту;
- ◆ если есть и билет, и карта ЕИС, то пройти на стадион;
- ◆ если есть небольшая очередь, то встать в нее и дожидаться своей очереди.

При этом при затруднении реализации стратегии используются следующие средства:

- задать вопрос волонтеру;
- получить информацию из инфомата;
- общаться с другими агентами для получения информации устно (обмен сообщениями);
- обратиться к полицейскому с сигналом «тревога» в случае ЧП и, получив помощь, продолжать выполнение цели.

4. В данной постановке упростим задачу — намерения агента и его обязательства по отношению к другим агентам в своем классе и по отношению к другим классам будем считать отсутствующими.

В таких допущениях и предположениях была создана условная модель организации проведения

спортивного мероприятия. Модель разработана на основе программного продукта AnyLogic 6 компании XJTechnologies (Санкт-Петербург) с использованием так называемой «пешеходной библиотеки». Результаты моделирования приведены как экранные скриншоты рабочей области модели (рис. 1–3) с соответствующими пояснениями.

Модель отражает движение посетителей Олимпийских Игр в Сочи-2014 на центральный стадион Олимпийского парка.

Посетители прибывают группами (форма группы — групповой поток). Каждая группа посетителей прибывает согласно заданной интенсивности прибытия (эквивалентной экспоненциально распределенному времени между прибытиями со средним значением, равным 1/интенсивность). Максимальная интенсивность прибытия группы посетителей в единицу времени (группы формируются из двух/трех человек) задана как 300 групп/час. Время между прибытиями посетителей, входящих в состав одной группы удовлетворяет экспоненциальному распределению и составляет $\exp(10/\text{сек.})$. Задается скорость, с которой посетитель будет комфортно двигаться при отсутствии внешних факторов (в модели — 0,5 м/с). Модельное время составляет 1 час. Случайные процессы (обращение к волонтеру за информационной услугой, приход в кассу за билетом и др.) моделируются различными законами распределения (равномерный, Пуассона и др.).

Предполагается, что посетитель может столкнуться с пятью видами потенциальных ограничений: прохождение через металлорамку, встреча для получения информации с волонтером, получение билета на мероприятие, получение карты ЕИС, турникеты стадиона. По всем ограничениям задается закон распределения времени задержки и допускается вариация его параметров. Так, например, максимальная интенсивность прохождения посетителей через объект «металлорамки» составляет 30 чел/мин, время обслуживания агента волонтером — 20 секунд, кассиром — 3 минуты и др.

В зависимости от статуса посетителя, задаваемого так называемым коэффициентом предпочтения, посетители следуют на один из пяти выходных «портов»: проходя через металлодетектор (в виде рамки), двигаются по направлению к турникетам, расположенным возле входа на стадион, либо приобретают необходимые разрешительные документы (билет на мероприятие и карту ЕИС).

В модели по определенному таймеру реализуется ряд событий. К числу важнейших отнесены следующие (см. табл. 1).

Таблица 1.

Моделируемые события

Время регистрации события	Событие	Реакция модели
Начало отсчета моделирования	Приход поезда с посетителями массового мероприятия на Олимпийском стадионе	Формирование потока посетителей из дискретных групп
15 минута с момента начала моделирования	Закрытие одной из билетных касс по техническим причинам	Увеличение времени выдачи карт ЕИС в 2 раза с образованием очереди
20 минута с момента начала моделирования	Сообщение 50% агентам группы K_{13} об образовании очереди на продажу билетов	Ускорение движения 50% агентов потока на 25%
35 минута с момента начала моделирования	Техническая неисправность в пункте выдачи ЕИС	Увеличение времени выдачи карт ЕИС в 2 раза с образованием очереди
50 минута с момента начала моделирования	Сообщение 20% агентам группы K_{12} и 20% агентам группы K_{13} об образовании очереди на выдачу карт ЕИС	Ускорение движения 50% агентов потока на 15%
1 час 10 минут с момента начала моделирования	Нехватка волонтеров возле металлорам, сообщающих посетителям маршрут движения в зависимости от наличия билета и карты ЕИС	Образование большой очереди возле металлорам с возможностью перехода в затор
1 час 20 минут с момента начала моделирования	Фиксация времени, когда до начала мероприятия остается 10 минут	Ускорение движения 100% потока на 40%
Случайное число в интервале от 20 до 40 минут до начала массового мероприятия	Задержание потенциального террориста возле металлорам	Сокращение пропускной способности прохода через металлорамки на 50%
Случайное число в интервале от 10 до 30 минут до начала массового мероприятия	Срабатывание взрывного устройства возле металлорам либо на территории Олимпийского парка	Хаотическое движение больших групп людей в локализованной области пространства

В модели также реализована матрица взаимных связей агентов различных классов (в терминологии AnyLogic – стейчарты). Она включает в себя обмен сообщениями различного рода между агентами различных классов и внутри классов, возможность обращения к волонтерам и полицейским и получения обратной связи от них (оказание информационной услуги, получение сигнала тревоги и др.). Посредством стейчартов и таймеров моделируются разнообразные события.

Все объекты инфраструктуры (часть 3D-визуализации Олимпийского парка Сочи и Олимпийского стадиона) уменьшены относительно их реального размера; группы посетителей игр представляются в виде кружков разных цветов (цвета меняются в зависимости от статуса посетителя) и для наглядности увеличены. Модель является динамической, может работать как в реальном, так и в модельном времени.

Такая модель позволяет описывать поток следующих на массовое спортивное мероприятие посетителей Игр, имеющих и не имеющих соответствующие разрешительные документы для безопасного прохода на соревнования, возникновение очередей и, как предельный случай, толпы при организации массового мероприятия в зависимости от работы

транспорта и характеристик «приборов» обслуживания: количества турникетов и металлорам, оперативности выдачи разрешительных документов, а также эффективности работы волонтеров в части времени правильного обслуживания посетителей Игр.

4. Примеры результатов моделирования

В силу ограниченности объема статьи, приведем лишь три примера результатов моделирования.

Поток людей (показаны кружками черного цвета) проходит сквозь металлорамки, часть людей следует сразу на стадион, не имеющие разрешительных документов следуют либо к кассам, либо к пункту выдачи карт ЕИС, либо последовательно к обоим объектам, далее, после приобретения, направляются на стадион, волонтеры (более светлые (бордовые) кружки) осуществляют диспетчеризацию следующих на мероприятие людей в зависимости от наличия/отсутствия билетов и карт ЕИС.

В нижней части рисунка – процессная схема модели в нотации UML (Unified Modeling Language), реализованной в программном продукте AnyLogic 6.

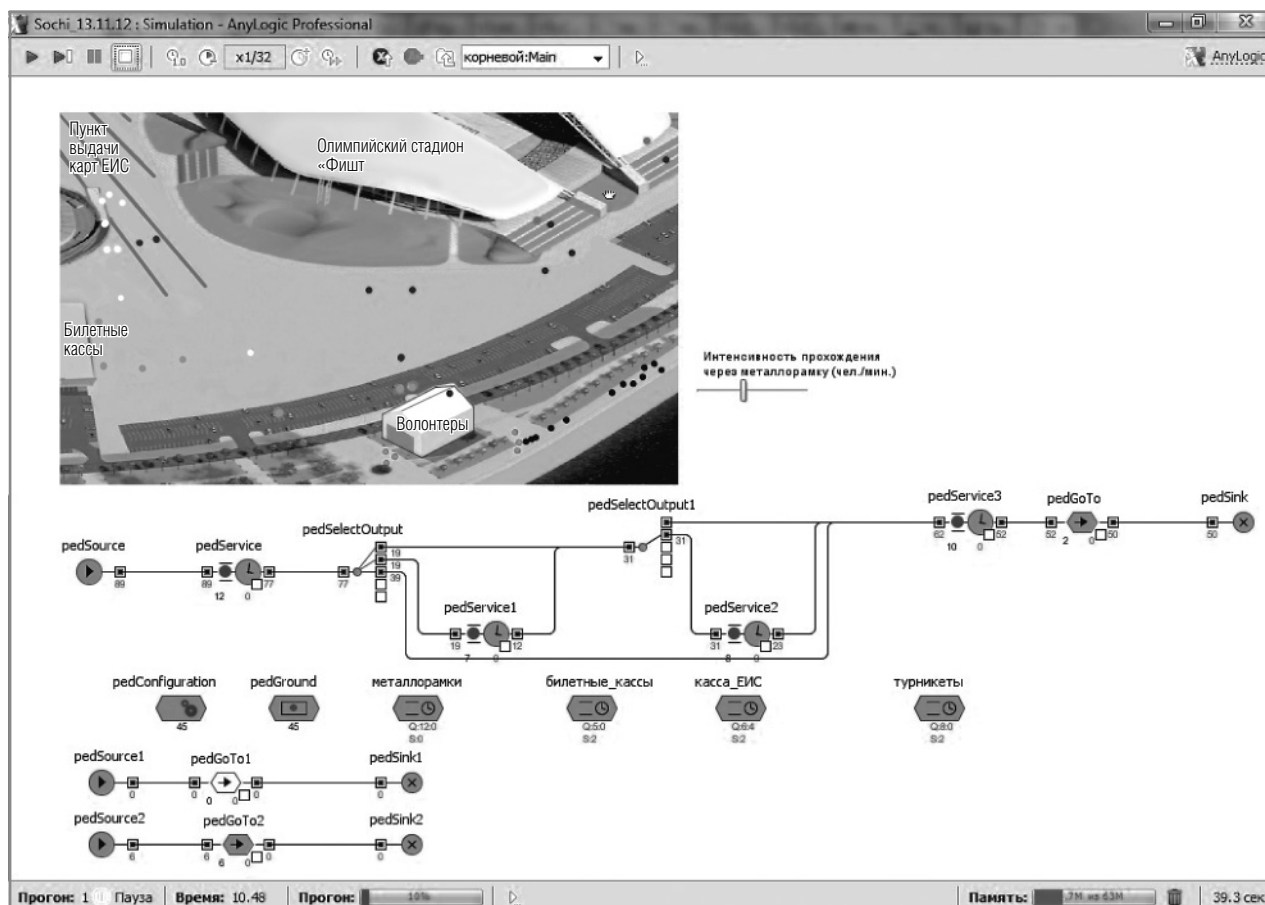


Рис. 1. Стандартная ситуация прохода на массовое мероприятие

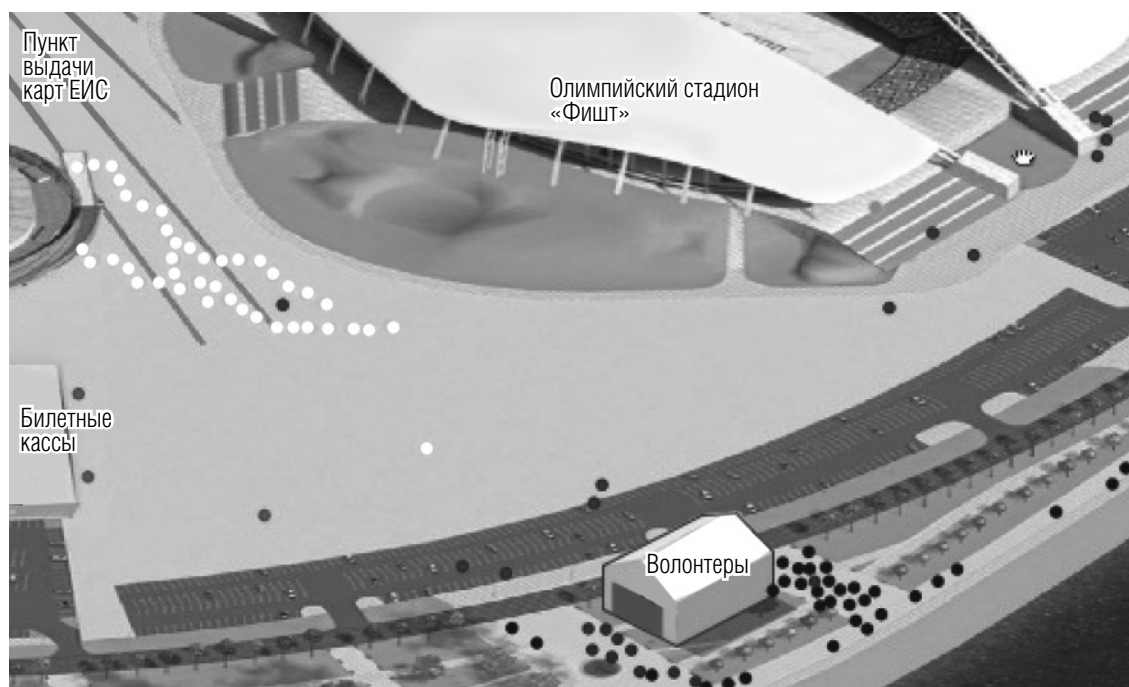


Рис. 2. Кризисная ситуация: образование очереди

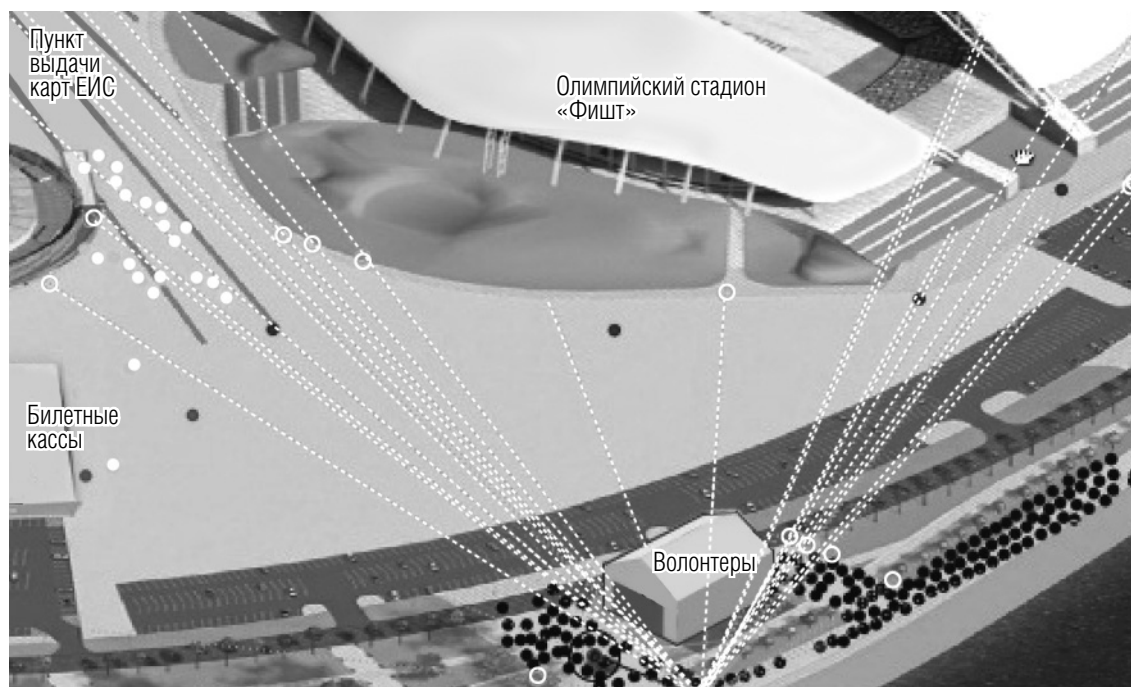


Рис. 3. Экстремальная ситуация

Вследствие фиксации кризисного события (например, выхода из строя части оборудования пункта выдачи карт ЕИС) увеличился временной период выдачи карт ЕИС, что привело к образованию существенной очереди («цепочка» из светлых кружков). Обмен информацией между посетителями, стоящими в очереди, и другими группами посетителей, подходящих к металлорамкам, вызвал изменение параметров движения групп людей (увеличение скорости движения и распространение информации на следующие группы). Оба события повлекли за собой образование небольшого «сгустка» посетителей на входе перед металлорамками, что существенно увеличило нагрузку на волонтеров, а ограниченное их количество еще более усугубило ситуацию.

Иллюстрация форс-мажорной ситуации (например, одновременный приход к олимпийскому объекту нескольких видов транспорта, обнаружение и задержание полицией потенциального террориста и др., срабатывание взрывного устройства). На входе в Олимпийский парк образовалась толпа с соответствующими последствиями (в разработанной модели управление дальнейшими процессами передается через интеграционную шину и, далее, в модель поведения толпы). Проходимость металлорамки не обеспечивает потребности в движении посетителей. Волонтеры не могут справиться со своими обязанностями. В пункте выдачи карт ЕИС

за счет указания от полиции об ужесточении режима прохода на олимпийских объектах увеличено время выдачи карт и образовалась очередь. Билетные кассы простаивают. Начало мероприятия под угрозой срыва.

5. Обсуждение результатов

Несмотря на значительные возможности, имитационное моделирование при решении практических задач до настоящего времени используется недостаточно активно. Вместе с тем, создание таких моделей в ряде случаев является жизненно необходимым, так как позволяет не только анализировать происходящие процессы в динамике, но также и моделировать критические ситуации с целью принятия соответствующих мер по их предотвращению и недопущению.

При проведении массовых мероприятий, в частности спортивного характера, значительное внимание необходимо уделять не только организационным мерам по обеспечению безопасности, но также и процессам имитационного моделирования поведения потоков людей, моделированию образования очередей, последствий возникновения чрезвычайных ситуаций различного характера.

Количество публикаций по вопросам имитационного моделирования процессов движения потоков людей при проведении в различных странах

Олимпийских игр исчисляется единицами. Вместе с тем, значимость такого моделирования трудно переоценить. Даже описанная модель, построенная на простейших решающих правилах, показывает возможность возникновения крупных проблемных ситуаций с обслуживанием посетителей будущих Олимпийских игр в Сочи. Но ценность модели состоит в том, что она является динамической, может функционировать как в реальном, так и в модельном времени и позволяет путем варьирования ключевыми характеристиками, параметрами законов распределения случайных величин, входными дан-

ными выбрать оптимальные условия проведения массового мероприятия, определить потребные ресурсы обслуживающего персонала.

Модели, подобные разработанной, должны увязываться в единый комплекс с другими имитационными моделями – транспортных потоков, образования и поведения толпы, потребного информационного трафика и обмена данными и др., что позволит наиболее полно и адекватно исследовать сложные организационно-технические процессы, объективно возникающие при подготовке и проведении массовых мероприятий. ■

Литература

1. Akopov S., Beklaryan A. Simulation of human crowd behavior in extreme situations // International Journal of Pure and Applied Mathematics, v. 79, № 1, 2012, p. 121-138.
2. Аптуков А.М., Брацун Б.А., Моделирование групповой динамики толпы, паникующей в ограниченном пространстве // Вестник Пермского Университета, серия Математика, Механика, Информатика, Вып. 3(29), 2009.
3. Степанцов М.Е., Моделирование динамики движения группы людей на основе решеточного газа с нелокальными взаимодействиями // Известия Высших учебных заведений. Прикладная нелинейная динамика, 7:5 (1999), 44–46.
4. Карпов Ю. Имитационное моделирование систем. – Спб.: БВХ- Петербург, 2005.- 400 с.
5. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям. Серия «Науки об искусственном». - Едиториал УРСС 2002, 352 с.
6. Поспелов Д.А. От коллектива автоматов к мультиагентным системам // Proc. Of the International Workshop «Distributed Artificial Intelligence und Multi-Agent Systems», DIAMAS' 97, St. Peterburg, 1997. P. 319–325.
7. Жуков С.Ю. Навигация интеллектуальных агентов в сложных синтетических пространствах. Диссертация на соискание степени кандидата физико-математических наук, Санкт-Петербург, 2000.- 135 с.
8. Швецов А.Н. Агентно-ориентированные системы: от формальных моделей к промышленным приложениям / Всероссийский конкурсный отбор обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы», 2008. - 101 с.

ПРИМЕНЕНИЕ МЕТОДА ПРОГРАММИРОВАНИЯ В ОГРАНИЧЕНИЯХ ПРИ РЕШЕНИИ ЗАДАЧ ПЛАНИРОВАНИЯ ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ И РЕМОНТА ВОЗДУШНЫХ СУДОВ

А.В. Федотова,

*соискатель кафедры компьютерной автоматизации производства
Московского государственного технического университета им. Н.Э. Баумана*

М.В. Овсянников,

*кандидат технических наук, доцент кафедры компьютерной автоматизации
производства Московского государственного технического университета
им. Н.Э. Баумана*

E-mail: afedotova@act.org, mvo50@mail.ru

Адрес: г. Москва, ул. 2-я Бауманская, д. 5

В.В. Таратухин,

*доктор философии, кандидат технических наук, заведующий базовой кафедрой SAP
Национального исследовательского университета «Высшая школа экономики»,
руководитель научной группы Европейского исследовательского центра
в области информационных систем (ERCIS) Мюнстерского университета
(University of Muenster), Германия*

E-mail: victor.taratoukhine@ercis.uni-muenster.de

Адрес: Germany, Munster, Leonardo-Campus, 3, University of Muenster, ERCIS – Headquarters

В статье описываются анализ методов планирования ТОРО (технического обслуживания и ремонта оборудования), постановка задачи ТОРО, формирование ограничений, приведены методы удовлетворения ограничений, подходящие для решения задачи планирования периодических процессов обслуживания. Выбраны методы и сформулирована методика решения задачи, основанная на комбинации методов перебора и предварительного ограничения пространства поиска. Предложены алгоритмы реализации методов.

Ключевые слова: ТОРО, методы планирования ТОРО, удовлетворение ограничений, CSP, периодические процессы обслуживания, комбинаторная задача, методы поиска решений, поиск с возвратами, аксиомы, онтология предметной области.

1. Введение

На конкурентоспособность воздушных судов оказывает влияние целый ряд факторов, в том числе действующая система технического обслуживания и ремонта. Техническое обслуживание и ремонт или ТОРО - комплекс операций или операция по поддержанию работоспособности или исправности изделия при использовании по назначению, ожидании, хранении и транспортировании [1].

Основными параметрами, характеризующими совершенство системы ТОРО, являются объем и периодичность выполнения работ по техническому обслуживанию, оказывающие прямое влияние на эксплуатационные расходы и интенсивность эксплуатации воздушного судна (ВС). Совершенно очевидно, что чем меньше объем работ и больше периодичность их выполнения, тем меньше эксплуатационные расходы и выше интенсивность эксплуатации ВС. Доля затрат на ТОРО может составлять от 10 до 50 процентов от общих эксплуатационных расходов [2]. На объем и периодичность ТОРО влияют конструктивные особенности ВС, методы технической эксплуатации, нормативные и методические требования, качество выполнения работ, параметры эксплуатации, условия окружающей среды и другие факторы.

Объем работ и периодичность их выполнения определяются программой ТОРО. Необходимо отметить, что требования к содержанию программы ТОРО в РФ отличаются от требований, действующих за рубежом. В РФ программа ТОРО является документом разработчика, общим для типа ВС, в то время как программа ТОРО за рубежом является документом, разрабатываемым эксплуатантом на каждое отдельно взятое ВС с учетом условий и параметров его эксплуатации, а также принятой эксплуатантом системы ТОРО, на основе минимальных требований к техническому обслуживанию (ТО), определенных в процессе сертификации типа.

В работе рассматриваются методы планирования и обслуживания воздушного судна.

2. Анализ методов планирования ТОРО

Цели эффективного плана технического обслуживания изделия авиационной техники (АТ) заключаются в следующем:

- ◆ обеспечение заданных уровней надежности и безопасности;

- ◆ восстановление надежности и безопасности до заложенных при проектировании уровней в случае их понижения;

- ◆ получение информации, необходимой для улучшения конструкции тех изделий, надежность которых оказалась недостаточной;

- ◆ достижение этих целей с минимальными суммарными затратами, включая затраты на техническое обслуживание и затраты, вызванные отказами.

В общей постановке при наличии всей необходимой исходной информации количественное подтверждение эффективности плана ТО состоит в обеспечении заданных вероятностных показателей безопасности воздушного судна (ВС) при минимизации удельных затрат на ТО (например, стоимости ТО на один час налета). При таком подходе основной трудностью является количественное подтверждение выполнения требований к ВС, для чего необходима адекватная математическая модель, позволяющая оценить влияние ТО на показатели безопасности и эффективности ВС.

Реальной альтернативой количественному подходу является рациональное сочетание качественного инженерного анализа по выбору методов технической эксплуатации (ТЭ) и работ по ТО ВС с количественной оптимизацией периодичности ТО на основе математической модели влияния планового ТО на надежность и безопасность бортовых систем. Это позволяет формализовано анализировать влияние возможных отказов систем и их компонентов на безопасность, регулярность полетов и экономичность ВС [3].

Качественный подход был предложен в 70-х годах авиакомпанией United Airlines и стал позднее широко известен как методический документ по разработке программы планового ТО – ATA MSG-3 [4]. Этот документ признан FAA США, JAA и EASA Евросоюза и широко используется разработчиками и эксплуатантами ВС.

Как показала практика, методика MSG-3 имеет ряд существенных недостатков. Поэтому принципы MSG-3 были доработаны и получили развитие в руководстве для конструкторов и эксплуатантов [5], которым конструкторы и специалисты по ТО могут пользоваться при создании и сопровождении эксплуатации ВС.

Методика планирования ТО по MSG-3 описывает метод формирования планового ТО. Неплановые ТО обуславливаются плановыми работами,

замечаниями в процессе нормальной эксплуатации или анализом эксплуатационных данных.

Плановые ТО разрабатываются с использованием подхода целенаправленной логики принятия решений и реализуются в виде программы, ориентированной на состав работ по ТО. Эта логика является базой для методики анализа каждого из элементов систем, важных в плане ТО (систем, подсистем, модулей, агрегатов, блоков, составных частей и т.д.), с использованием располагаемых технических данных. Принципиально анализ базируется на рассмотрении функциональных отказов изделий и их последствий [4].

В Российской Федерации над данной задачей работают специалисты Научно – исследовательского центра CALS-технологий «Прикладная логистика». Ими была предложена «**Методика выбора периодичности работ по ТО на основе количественных методов анализа**».

Периодичность плановых работ по ТО устанавливается таким образом, чтобы минимизировать естественное ухудшение заложенного в конструкцию уровня безопасности и надежности ВС, его систем и элементов, не допуская выхода уровня летной годности ВС за пределы, установленные нормами летной годности.

Объем планового ТО ВС определяется рабочими группами, включающими представителей авиакомпаний и Разработчика и представляется для утверждения.

При определении периодичности работ по ТО ВС рассматриваются типовые режимы его эксплуатации: средняя продолжительность типового полета, среднегодовой налет [3, 6].

В настоящее время планирование графика ТОРО основывается на паспортных данных изделия, стандартов организации, требований отрасли и т.д. При этом на оси времени жёстко фиксируется время выполнения регламентных работ путём назначения интервалов между выполнением ТО. Интервалы между ТО задаются циклически. Состав работ при каждом ТО формируется в зависимости от его типа. Для каждого типа ТО предусматривается перечень работ, задаваемый документацией, который необходимо выполнить в рамках проведения обслуживания. К основным недостаткам данного метода можно отнести: невозможность внесения изменений в график ТО; практически полное отсутствие обратной связи с обслуживаемой системой; отсутствие накопления

данных об обслуживаемой системе; низкая адаптивная способность.

Цель данной работы: создать метод и алгоритм планирования выполнения работ, универсальный как по процедурам обслуживания, так и по видам изделий, а также позволяющий без перепрограммирования производить настройку на конкретные условия, что значительно упростит поддержку системы в целом и расширит области ее использования, а также позволит сократить затраты на ТО.

3. Постановка задачи

Выделяется два типа задач построения конфигурации систем: отбор и/или размещение, построение иерархии системы. Наша задача относится к первому типу, на входе: множество элементов/компонентов, которое разделено на подмножества.

ТО представляет собой совокупность процедур $P = \{p_i | i = \overline{1, n}\}$, выполняемых на объекте для поддержания его технической исправности. Общее число процедур ТО, выполняемых на временном интервале $T = [t_1, t_n]$, задаваемом начальной и конечной точкой (t_1 и t_n соответственно), формирует множество процедур P . Иллюстрация данного факта приведена на рис. 1.

Имеется множество работ $W = \{w_k\}$, которые могут быть выполнены при ТО.

Процедура ТО представляет собой множество работ: $W = \{w_k\}$, где w_k – k -я работа, W_n – множество работ, входящих в состав процедуры ТО p_n и представлена на рис. 1.

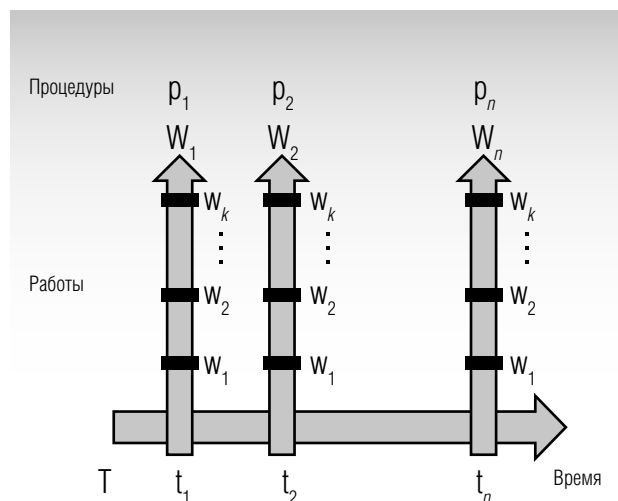


Рис. 1. Задача планирования ТОРО

Задача алгоритма планирования ТО состоит в формировании множества Т и определения множеств $W_i, i = \overline{1, n}$, определяющих состав работ при проведении работ в процедуре ТО. При этом при формировании множества ТО и множества входящих в него работ следует учитывать ограничения, накладываемые различными источниками.

Существует m единиц оборудования, для которого требуется провести работы по поддержанию исправности и работоспособности. Существует множество работ $w_k - W_n, k = [1, k_{max}]$ ТО и планового ремонта специального оборудования, выполняемых в ходе процедур ТО.

Постановка задачи звучит следующим образом:

Определить состав и время проведения процедур ТО при которых выбранный критерий будет принимать экстремальное (минимальное) значение, т.е. определить все x_{kn} , при которых

- 1) $S \rightarrow \min$,
- 2) $R = \max(R_n) \rightarrow \min$,

где x_{kn} – логическая переменная, значение которой определяется входимостью k -ой работы в n -ую процедуру ТО: если k -ая работа входит в n -ую процедуру, то $x_{kn} = 1$, иначе $x_{kn} = 0$:

$$x_{kn} = \begin{cases} 1, & \text{если } k\text{-я работа входит в } n\text{-ю процедуру ТО} \\ 0, & \text{если } k\text{-я работа не входит в } n\text{-ю процедуру ТО} \end{cases}$$

S – общие материально-технические затраты на производство ТО:

$$S = \sum_{kn} s_k \cdot x_{kn},$$

где s_k – материально-технические затраты на k -ую работу, $k = [1, k_{max}]$, где k_{max} – общее количество работ в системе;

R_n – трудоёмкость n -ой процедуры ТО:

$$R_n = \sum_k r_k \cdot x_{kn},$$

где r_k – трудоёмкость k -ой работы, $k = [1, k_{max}]$, где k_{max} – общее количество работ в системе.

N					1				
...			1						1
7		1			1				
6									1
5				1					
4	1		1				1		
3						1			
2									1
1	1		1	1	1				
	1	2	3	4	5	6	7	...	K

Рис. 2. Представление решения x_{kn}

Возможное решение x_{kn} представлено на рис. 2.

Поставленная задача должна быть решена с учетом специальных ограничений, которые будут сформулированы ниже в п.5.

Для реализации поставленной цели сводим задачу к общей задаче проектирования конфигурации (**configuration design**), где под конфигурацией понимается распределение конкретных работ по процедурам.

Суть метода программирования в ограничениях заключается в описании предметной области набором правил и ограничений. Ограничения могут иметь вид уравнений, неравенств, логических выражений и т.п.

Система работ представляет собой набор данных, на основе которых формируется состав работ в каждом ТО. Система должна содержать:

- ♦ информацию о функциональной декомпозиции работ в виде иерархической структуры компонентов, образующих состав работ;
- ♦ формальное описание закономерностей, определяющих возможность включения данной работы в состав текущего ТО;
- ♦ механизм, позволяющий проверить полученное множество W_n работ на соответствие требованиям.

Построение системы работ является основой для дальнейшего многократного формирования множества графиков ТО.

4. Построение онтологии предметной области

Для описания функциональной декомпозиции работ в виде иерархической структуры компонентов, образующих состав работ, представим онтологию предметной области.

Онтология представляет собой несколько областей: справочник изделий, к которым применяется ТО; справочник работ, из которых формируются процедуры, формирующие ТО; взаимосвязь всех компонентов с учетом принадлежностей, свойств и ограничений.

Объекты предметной области:

<Изделие, ТО, Процедура, Работа>

Множество процедур: ПР <ПР1, ПР2, ..., ПРn>

Множество работ: РБ <РБ1, РБ2, ..., РБk>

Изделие: И <И1>

Оборудование: ОБ <ОБ1, ОБ2, ..., ОБg>

Техническое обслуживание: ТО

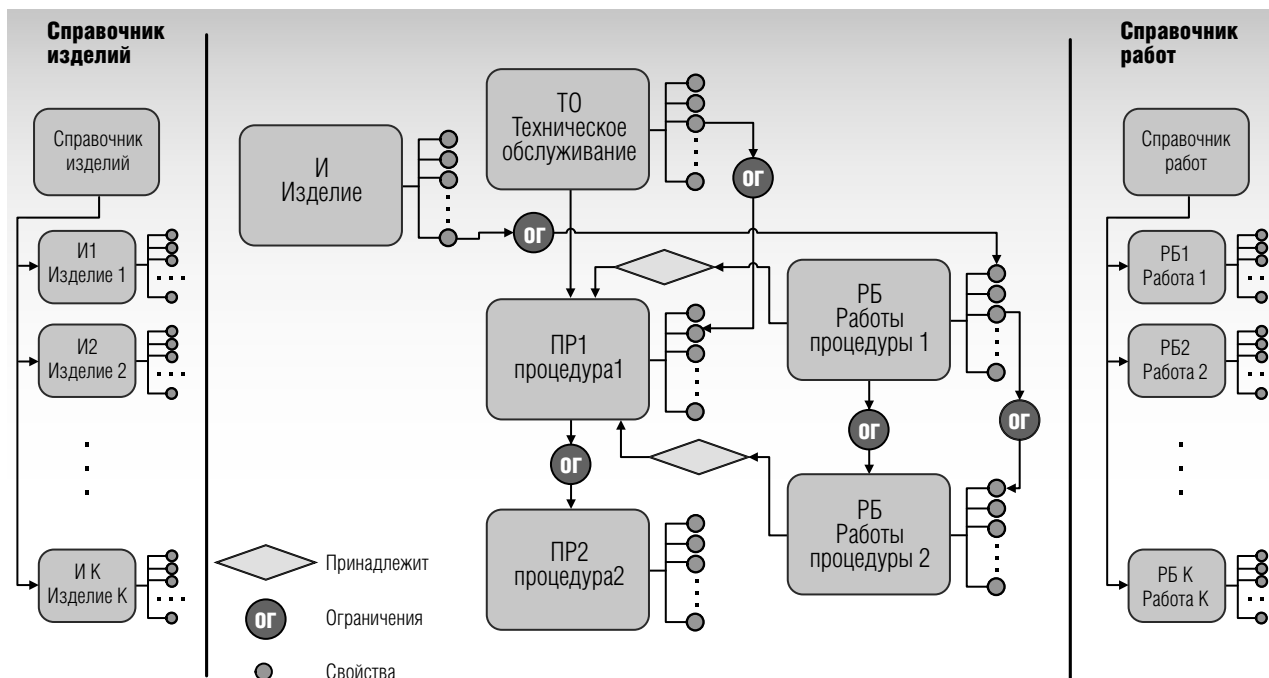


Рис. 3. Онтология предметной области

Объекты и их свойства:

Объект **ТО** имеет следующие свойства:

ТО.Трудоемкость, ТО.Затраты, ТО.Время_выполнения;

Объект **ПР** имеет следующие свойства:

ПР.Трудоемкость, ПР.Затраты, ПР.Время_выполнения, ПР.Тип;

Объект **РБ** имеет следующие свойства:

РБ.Трудоемкость, РБ.Затраты, РБ.Время_выполнения, РБ.Тип, РБ.Совместная_выполнимость;

Объект **И** имеет следующие свойства:

И.Кол-во_отказов, И.Время_наработки_на_отказ, И.Время_наработки, И.Тип, И.Время_простоя;

Объект **ОБ** имеет следующие свойства: ОБ.Время_наработки, ОБ.Кол-во_отказов, ОБ.Время_простоя, ОБ.Время_наработки_на_отказ.

Перечень требований **ТР**: {ТР}.Затраты_доп, {ТР}.Трудоемкость_доп, {ТР}.Надежность, {ТР}.Время_простоя_макс, {ТР}.Затраты_пр_мин, {ТР}.Затраты_пр_макс, {ТР}.Трудоемкость_пр_мин, {ТР}.Трудоемкость_пр_макс, {ТР}.Время_наработки_макс, {ТР}.Время_наработки_мин, {ТР}.Время_простоя_макс, {ТР}.Время_простоя_мин, {ТР}.Время_выполнения_макс, {ТР}.Время_выполнения_мин, {ТР}.Время_макс_м_раб.

5. Ограничения

Описание закономерностей, определяющих возможность включения данной работы в состав текущего ТО указываются в виде ограничений (аксиом).

Перечень ограничений (Constraint):

1. В одну процедуру ТО не могут одновременно входить работы, заданные следующей матрицей:

R_{Bi} ВХОДИТ_В_СПИСОК PR_n , R_{Bj} ВХОДИТ_В_СПИСОК PR_n (Совместимость=1).

$$z_{ij} = \begin{cases} 1, & \text{если } i\text{-я работа может выполняться вместе с } j\text{-й} \\ 0, & \text{если } i\text{-я работа не может выполняться вместе с } j\text{-й} \end{cases}$$

2. Затраты на проведение ТО не должны превышать заданной величины доп. ТО:

$$\text{Затраты} \leq \{\text{ТР}\}.\text{Затраты_доп}$$

$$S \leq S_{\text{доп}}$$

3. Затраты на каждой процедуре ТО должны быть не более максимального значения:

$$\text{ПР.Затраты} \leq \{\text{ТР}\}.\text{Затраты_пр_макс}$$

4. Нарботка оборудования в момент проведения процедуры ТО должна находиться в заданных пределах, $t_{\text{нартmin}} \leq t_{\text{нарт}i} \leq t_{\text{задmax}}$, где $t_{\text{нарт}i}$ - наработка m -го оборудования при проведении i -ой процедуры ТО.

ОБ.Время_наработки \leq {ТР}.Время_наработки_макс && ОБ.Время_наработки \geq {ТР}.Время_наработки_мин

5. Время простоя оборудования в момент проведения процедуры ТО должно находиться в заданном интервале времени: $t_{прост\ min} \leq t_{прост\ i} \leq t_{прост\ max}$, где $t_{прост\ i}$ – время простоя i -го оборудования при проведении i -ой процедуры ТО.

ОБ.Время_простоя $\leq \{TR\}$. Время_ простоя _ макс &&

ОБ. Время_простоя $\geq \{TR\}$. Время_ простоя _ мин

6. Промежуток времени между k -ми работами не должен превышать максимального значения

$$\begin{aligned} t_i &= t(w_k), w_k \in W_n; \\ t_{i+1} &= t(w_k), w_k \in W_{n+1}; \\ t_{min} &\leq (t_{i+1} - t_i) \leq t_{max} \end{aligned}$$

ПР n .Время_начала (РБ Входит_в_список ПР n) - ПР m .Время_начала (РБ Входит_в_список ПР m) $\leq \{TR\}$.Время_макс_м_раб

7. При наработке оборудования типа g на отказ выше заданного значения $t_{нараотказ_зад}$ в промежуток времени между предыдущей и текущей процедурой ТО должны быть проведены работы, определенные типом оборудования g при текущей процедуре ТО: если $\varepsilon_{gp}(n-1); (n) > 1$ $x_{kn} = 1$ для $k \in K_{gp}$.

Если ОБ g . Время_наработки_на_отказ $> \{TR\}$.
Время_наработки_макс,

РБ условия применимости: ВХОДИТ_В_СПИСОК

6. Решение задачи формирования процедур ТО

1. Поиск всех дополнительных вариантов распределения работ по процедурам ТО, по которым выполняются все ограничения.

2. Упорядочивание полученных решений по целевой функции.

3. Поиск решения в соответствии с наилучшими значениями целевой функции.

7. Методы поиска, существующие алгоритмы

Здесь можно привести следующий общий набор групп методов поиска решений:

◆ Методы и алгоритмы ограничения пространства поиска и распространения ограничений (Forward Checking, Node Consistence, Arc Consistence, Constraints Propagation и т.д.);

◆ Точные методы (Backtracking, Динамического

программирования, Back Jumping, Reset и т.д.);

◆ Приближенные методы.

Рассмотрим некоторые методы и алгоритмы поиска решений, наиболее интересные для решения задачи структурного синтеза.

Алгоритм неоптимизированного полного перебора

Наиболее простым алгоритмом подбора конфигураций с учетом ограничений является алгоритм полного перебора. Преимуществом использования данного алгоритма является гарантированный обход всех вариантов решений и наиболее точные результаты поиска в виде выборки всех фактически возможных конфигураций или установление факта отсутствия решения. Данный подход имеет существенный недостаток, связанный со слишком большим временем перебора всех возможных вариантов конфигураций.

Метод локального поиска

Адаптированный под данную задачу алгоритм локального поиска выглядит следующим образом:

◆ Построить начальную конфигурацию;

◆ В цикле:

○ Проверить условие завершения и если цель достигнута – на выход:

■ Условие завершения – Удовлетворение всех ограничений или достижение ограничения по времени исполнения.

○ Выбрать ОК для замены компонента.
Критерии:

■ Участие компонента ОК в наибольшем числе ограничений;

■ Компонент ОК нарушает наибольшее число ограничений;

■ Компонент с наименьшим доменом (числом компонентов для подстановки в данный ОК).

○ Выбрать компонент из справочника для подстановки:

Критерии:

■ Нарушает наименьшее количество ограничений;

■ Наиболее точно обеспечивает выполнение требований.

○ Сравнить качество полученной конфигурации с предыдущей:

Критерии:

■ Количество нарушенных ограничений;

- Степень близости параметров конфигурации требуемым;
- Сравнить текущую конфигурацию с лучшей:
 - Если текущая конфигурация лучше – объявить текущую конфигурацию лучшей.

Несмотря на то, что метод локального поиска широко применяется для решения комбинаторных задач, данный метод имеет недостаток в том, что не обеспечивает нахождения оптимального значения и может стремиться к локальному оптимуму.

8. Описание разработанной методики

На основе обзора существующих методов была сформулирована следующая методика поиска. Методика основана на комбинации метода перебора с методами предварительного ограничения пространства поиска.

В отличие от метода полного перебора данная методика предполагает предварительные этапы ограничения пространства поиска, чем обеспечивается значительное сокращение времени поиска решения.

Методика состоит из следующих этапов:

1. На **первом** этапе происходит **обеспечение совместности в узлах** путем проверки унарных требований к параметрам отдельных компонентов, с использованием алгоритма **NC-1**. На этом этапе удовлетворяются ограничения, которые связаны с уже известными параметрами компонентов (требования к параметрам компонентов);

Алгоритм обеспечения совместности в узлах NC-1

Алгоритм обеспечивает совместность в узлах и называется NC-1 [7,8,9].

Этот алгоритм позволяет удалить из областей определения всех переменных задачи значения, которые не удовлетворяют унарным ограничениям.

Достижение совместности в узлах сводится к просмотру каждого элемента в каждой области и проверке удовлетворяет ли это значение унарным ограничениям на эту переменную. Все значения, которые нарушают унарные ограничения, удаляются из этих областей.

После завершения алгоритма исходная задача будет сведена к задаче, которая является совместной в узлах.

2. На **втором** этапе производится **предварительная проверка** компонентов методом **Forward Checking**

(Метод предварительной проверки). На этом этапе происходит попарное сравнение компонентов справочников с целью удаления компонентов не совместимых между собой.

Метод предварительной проверки Forward Checking

С помощью метода предварительной проверки (*Forward Checking*) пространство поиска может быть значительно уменьшено или полностью разрешено [10,11].

Основную идею метода можно выразить следующим образом: если при назначении какой-либо переменной значения её домена набор неравенств становится несовместным, то следует удалить несовместные значения из доменов остальных переменных и продолжить поиск, а если текущий домен пуст, то происходит откат на предыдущий уровень [12]. Домен – набор компонентов (работ) из справочника.

Следует начинать поиск с переменных, у которых наименьшие по количеству значений домены.

Адаптированный под данную задачу алгоритм выглядит следующим образом:

1. Переменные сортируются в порядке возрастания количества значений в домене
2. Для каждой переменной x_i :
 - а. Переменной x_i ставится в соответствие каждое значение из ее домена
 - і. Для каждой переменной y_j :
 1. Переменной y_j ставится в соответствие каждое значение из ее домена
 - а. Проверяется выполнение бинарного ограничения на совместность значений параметра (компонентов) x_i и y_j
 - б. Если не выполняется то удаляется текущее значение x_i из соответствующего домена

3. На **третьем** этапе производится **поиск решений** с использованием метода **Back Tracking**, для этого происходит обход дерева и производится вычисление неизвестных значений параметров компонентов и проверка соответствия расчетных параметров с требованиями.

Поиск с возвратом (Backtracking)

Алгоритм поиска с возвратами для решения задач удовлетворения ограничений CSP составлен по принципу рекурсивного поиска в глубину [7, 13].

Данный метод предполагает перебор всех возможных комбинаций путем обхода дерева вариантов компоновки с возвратами при не выполнении ограничений, *рис. 4*.

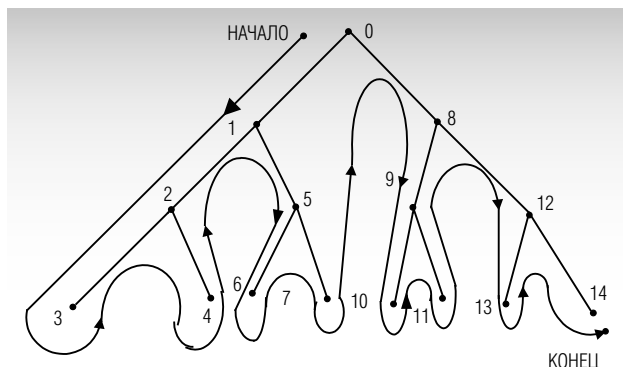


Рис. 4. Поиск с возвратом (Backtracking)

Метод позволяет произвести полный обход всех комбинаций компонентов и найти все возможные решения.

9. Заключение

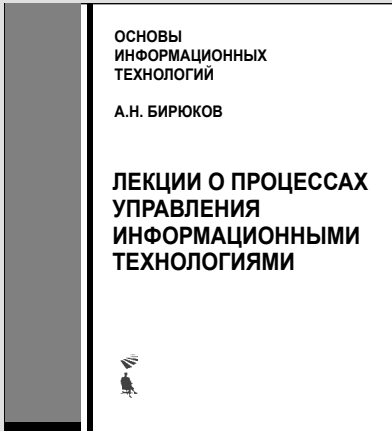
Реализация алгоритма предполагает наличие отдельных модулей: парсера (решателя) ограничений и генерации и поиска решений. Это позволяет без перепрограммирования производить настройку на конкретные условия, что значительно упрощает поддержку системы в целом и расширяет области ее использования.

Среди дальнейших планов стоит реализация данной методики в виде обобщенного решателя и алгоритмов формирования множества допустимых комбинаций в онтологических структурах. При разработке решателя необходимо исследовать и разработать язык и средства создания и использования ограничений, а также апробировать алгоритмы на более объемном примере и разработать модели и интерфейсы интеграции в существующие PDM и EAM системы. ■

Литература

- ГОСТ 18322-78. Система технического обслуживания и ремонта техники. Термины и определения.
- Airline Maintenance Program Development Seminar, Fleet Maintenance Seminars, Boeing. — Seattle, Washington: 2008.
- Методические указания «Руководство по проведению анализа логистической поддержки изделий авиационной техники», Научно-исследовательский центр CALS-технологий «Прикладная логистика». — М.: 2010.
- ATA MSG-3. Revision 2007.1. Operator/Manufacturer Scheduled Maintenance Development. ATA. — 2007.
- Руководство для конструкторов и эксплуатантов по разработке и сертификации программ технического обслуживания и ремонта воздушных судов гражданской авиации (РДК-Э). - ЛИИ им. М.М. Громова, ГосНИИ ГА. — 1993.
- Судов Е.В., Левин А.И., Петров А.В., Чубарова Е.В. Технологии интегрированной логистической поддержки изделий машиностроения. — М.: ООО Издательский дом «ИнформБюро», 2006.
- Семенов А.Л. Методы распространения ограничений: основные концепции // PSI'03/ИМРО. — Интервальная математика и методы распространения ограничений, 2003.
- Benhamou F., Older W. Applying Interval Arithmetic to Real, Integer and Boolean Constraints // Journal of Logic Programming. — 1997. — No. 32. — P. 1-24.
- Cleary J. Logical Arithmetic // Future Computer Systems. — 1987. — No. 2 (2). — P. 125-149.
- Gent I. Artificial Intelligence: Constraint Programming [Электронный ресурс]. — Режим доступа: <http://www.dcs.st-and.ac.uk/~ipg/AI/>. — Загл. с экрана.
- Bartak R. On-line guide to Constraint Programming [Электронный ресурс]. — Режим доступа: <http://ktlinux.ms.mff.cuni.cz/~bartak/constraints/>. — Загл. с экрана.
- Беляев С.А. Проблема разрешения constraint при решении задач комбинаторной оптимизации [Электронный ресурс]. — Режим доступа: http://constr.chat.ru/ARTICLES/cresprob_20010608.htm. — Загл. с экрана.
- Dechter R., Frost D. Backjump-based backtracking for constraint satisfaction problems // Artificial Intelligence. — 2002. — No. 136 (2). — P. 147-188.
- Буханов С.А., Овсянников М.В. Управление конфигурацией с использование метода программирования в ограничениях // Эффективные методы автоматизации подготовки и планирования производства: сборник научных трудов. — М.: Издательский дом «Спектр», 2011. — С. 150-156.
- Смирнов А.В., Пашкин М.П., Шилов Н.Г., Левашова Т.В. Онтологии в системах искусственного интеллекта: способы построения и организации // Новости искусственного интеллекта. — 2002. — №1.

16. Колчин А.Ф., Овсянников М.В., Стрекалов А.Ф., Сумароков С.В. Управление жизненным циклом продукции – М.: Анахарсис, 2002.
17. Яшура А.И. Система технического обслуживания и ремонта общепромышленного оборудования: Справочник – М.: ЭНАС, 2009. (Электронный ресурс: http://fictionbook.ru/author/aleksandr_ignatevich_yashura/sistema_tehnicheskogo_obslyujivaniya_i_re2006_ru/).
18. Taratoukhine V., Bechkoum K. Towards a Consistent Distributed Design: a Multi-Agent Approach. Information Visualisation '99. – London, 1999.
19. Hale M.A., Craig J.I. Preliminary Development of Agent Technologies for a Design Integration Framework // Proceedings of 5th Symposium on Multi-disciplinary Analysis and Optimisation. – Panama City, FL: September 7-9, 1994.



ОСНОВЫ
ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ

А.Н. БИРЮКОВ

ЛЕКЦИИ О ПРОЦЕССАХ
УПРАВЛЕНИЯ
ИНФОРМАЦИОННЫМИ
ТЕХНОЛОГИЯМИ

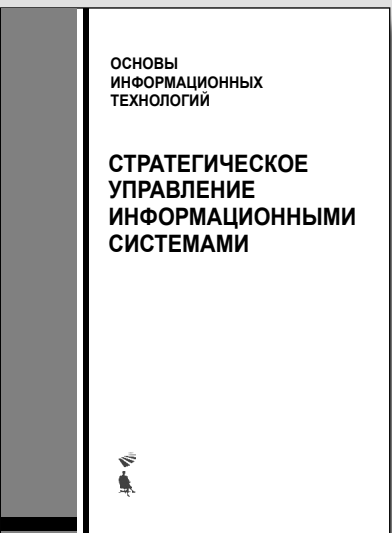
ЛЕКЦИИ О ПРОЦЕССАХ УПРАВЛЕНИЯ
ИНФОРМАЦИОННЫМИ
ТЕХНОЛОГИЯМИ

Учебное пособие

А.Н. Бирюков

Москва: Интуит.РУ, БИНОМ.ЛЗ, 2010.

Рассматриваются основные процессные модели и методики, связанные с управлением ИТ, появившиеся в последние годы. Основное внимание уделяется анализу их взаимосвязей и выявлению общих концепций и подходов. Изложение в большой степени базируется на оригинальных материалах, не переведенных на русский язык.



ОСНОВЫ
ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ

СТРАТЕГИЧЕСКОЕ
УПРАВЛЕНИЕ
ИНФОРМАЦИОННЫМИ
СИСТЕМАМИ

СТРАТЕГИЧЕСКОЕ УПРАВЛЕНИЕ
ИНФОРМАЦИОННЫМИ
СИСТЕМАМИ

Учебник.

**Р.Б. Васильев, Г.Н. Калянв,
Г.А. Левочкина, О.В Лукинова**

Под ред. **Г.Н. Калянв**

Москва: Интернет-университет
Информационных технологий: Бином.
Лаборатория знаний, 2010.

Учебник дает читателю представление о методологической базе и современных подходах и методах стратегического управления информационными системами, обеспечивающего целостный, процессно-ориентированный подход к принятию управленческих решений, направленных на повышение эффективности владения и развития информационных систем для достижения бизнес-целей предприятий и создания новых конкурентных преимуществ.

О РЕШЕНИИ ПРОБЛЕМЫ ВЫБОРКИ МАЛОГО РАЗМЕРА ПРИ ИСПОЛЬЗОВАНИИ ЛИНЕЙНОГО ДИСКРИМИНАНТНОГО АНАЛИЗА В ЗАДАЧАХ РАСПОЗНАВАНИЯ ЛИЦ

В.В. Мокеев,

доктор технических наук, старший научный сотрудник,
заведующий кафедрой информационных систем Южно-Уральского
государственного университета (ЮУрГУ)

С.В. Томилов,

аспирант кафедры информационных систем
Южно-Уральского государственного университета (ЮУрГУ)

E-mail: mokeyev@mail.ru, tomilov_stas@mail.ru

Адрес: г. Челябинск, пр. Ленина, д. 76

Рассматриваются эффективность линейного дискриминантного анализа в задачах распознавания лиц для случая малого размера выборки. Предлагается для вычисления дискриминантных компонент использовать обобщенный метод Якоби, что позволяет не терять дискриминантную информацию. Эффективность предлагаемого подхода демонстрируется в ходе экспериментальных исследований на базе данных ORL.

Ключевые слова: распознавание лиц, выборка малого размера, линейный дискриминантный анализ, база ORL.

1. Введение

Большое внимание в настоящее время уделяется проблемам обработки изображений. Это обусловлено многообразием прикладных задач, в которых используется либо сами изображения, либо результат их обработки. Наиболее актуальной задачей в данной области является распознавание фотоизображений лиц с целью иден-

тификации личности человека. Одной из проблем с которой приходится сталкиваться, решая данную задачу, является отсутствие достаточного количества фотоизображений одного человека, которое бы хорошо описывало всю индивидуальную изменчивость данного субъекта.

Среди различных методов распознавания лиц широкое распространение сегодня нашли мето-

ды, базирующиеся на линейном дискриминантном анализе. Линейный дискриминантный анализ (LDA – Linear Discriminant Analysis) [1] заключается в выборе проекции пространства изображений на пространство признаков таким образом, чтобы было минимизировано внутриклассовое и максимизировано межклассовое расстояние в пространстве признаков.

Несмотря на то, что линейный дискриминантный анализ является полезным инструментом для классификации образов, алгоритмы на основе LDA, страдают от проблемы «малого размера выборки» (SSS – Small Sample Size) [2], которая существует в задачах, где изображения обладают высокой размерностью, в то время как количество изображений, описывающих одну персону, является небольшим. В случае малого размера выборки матрица внутриклассовых различий становится сингулярной. Это значит, что собственные значения матрицы становятся нулевыми. Во многих работах используется понятие нуль-пространства матрицы. Обычно под этим термином понимается пространство, образованное собственными векторами, которым соответствуют нулевые собственные значения. Матрица межклассовых различий также становится сингулярной, в случае если размерность изображений больше числа классов. Так как в традиционном LDA вычисление дискриминантных компонент сопровождается обращением либо матрицы внутриклассовых либо матрицы межклассовых различий, то сингулярность матриц ведет к противоречию, которое и называется проблемой SSS. Чтобы решить данную проблему были разработаны различные разновидности алгоритмов LDA.

Среди различных путей решения этой проблемы, наиболее успешными являются подходы, сочетающие линейный дискриминантный анализ с методом главных компонент [3,4]. Хотя эффективность такого подхода при распознавании лиц очевидна, а теоретические его основы были заложены в [5,6], уже в работе [4] мы находим, что применение PCA не может гарантировать успешное применение LDA, т.е. после преобразования ковариационная матрица внутриклассовых различий может все еще быть сингулярной.

В связи с этим появилось достаточно большое число работ, в которых исследователи ищут более эффективные дискриминантные подпространства [5-9]. В работах [7, 8] были разработаны методы регуляризации, которые устраняют

сингулярность матрицы путем добавления к ней скалярной матрицы. Однако такой подход дает ограниченный эффект и неэффективен в случае, когда вырождены сразу две матрицы. Поиск эффективных дискриминантных пространств показал, что существует ключевая дискриминантная информация в нуль-пространстве матрицы внутриклассовых различий. Такой тип дискриминантной информации называется нерегулярной дискриминантной информацией, в отличие от регулярной дискриминантной информации находящейся вне пределов нуль-пространства. К сожалению, многие из вышеперечисленных методов, используя подход «PCA+LDA» теряют дискриминантную информацию, которая содержится в нуль-пространстве матрицы внутриклассовых различий, при том, что эта дискриминантная информация является очень важной для решения проблемы SSS.

В работе [5] описывается прямой линейный дискриминантный анализ (DLDA – Direct Linear Discriminant Analysis), который удаляет нуль-пространство матрицы межклассовых различий, что позволяет устранить сингулярность этой матрицы. В работе [9] предлагается подход NLDA (null space LDA), который базируется на том, что нуль-пространство матрицы внутриклассовых различий содержит ценную дискриминантную информацию. Основная идея DLDA заключается в том, чтобы исключить нуль-пространство матрицы межклассовых различий, которое не содержит полезной информации и сохранить нуль-пространство матрицы внутриклассовых различий, которое содержит важную дискриминантную информацию. Это может быть достигнуто путем диагонализации сначала матрицы межклассовых различий, и только затем уже матрицы внутриклассовых различий. Процедура NLDA выполняет эти действия в обратном порядке. Общей чертой подходов NLDA и DLDA является то, что все они теряют некоторую дискриминантную информацию.

В статье предлагается алгоритм обобщенного линейного дискриминантного анализа, в котором вычисление дискриминантных компонент осуществляется с помощью обобщенного метода Якоби. Обобщенный метод Якоби использует вращения для одновременной диагонализации матрицы межклассовых и внутриклассовых различий, что позволяет не терять дискриминантную информацию.

2. Обобщенный линейный дискриминантный анализ

Пусть имеется набор изображений, каждое из которых описывается вектором x_i^k , где i – номер изображений ($i = 1, 2, 3, \dots, M_k$), k – номер класса ($k = 1, 2, \dots, K$). Размерность вектора x_i^k равняется числу пикселей образа (N). Таким образом, все изображения можно представить в виде матрицы X , столбцами которой являются векторы x_i . Размерность пространства признаков определяется произведением $N \times M$. Пространство признаков центрируется относительно среднелинейного вектора

$$m = \frac{1}{M} \sum_k \sum_i x_i^k.$$

$$\text{Здесь } M = \sum_k M_k -$$

суммарное число изображений во всех классах. При этом формируется матрица отцентрированных изображений X^0 , строками которого являются векторы

$$\bar{x}_i^k = x_i^k - m.$$

При распознавании изображений целью линейного дискриминантного анализа является поиск проекционной матрицы, которая максимизирует так называемый критерий Фишера. Прежде чем описать этот критерий, необходимо определить две ковариационные матрицы: межклассовых и внутриклассовых различий. Матрица межклассовых различий определяется по формуле:

$$A_b = \frac{1}{M} \sum_k M_k (m_k - m)(m_k - m)^T, \quad (1)$$

где m_k – среднелинейный вектор изображений k -го класса, определяется с помощью выражения

$$m_k = \frac{1}{M_k} \sum_{i=1}^{M_k} x_i^k.$$

Матрица внутриклассовых различий определяется выражением

$$A_\omega = \frac{1}{M} \sum_{k=1}^K \sum_{i=1}^{M_k} (x_i^k - m_k)(x_i^k - m_k)^T. \quad (2)$$

Проекционная матрица V_{lda} выбирается как матрица с ортонормальными столбцами, максимизирующими критерий Фишера

$$V_{lda} = \arg \max_{V \in R^{n \times r}} \frac{|V^T A_b V|}{|V^T A_\omega V|} \quad (3)$$

Матрица дискриминантных компонент V_{lda} определяется из решения обобщенной задачи собственных значений

$$(A_b - \lambda A_\omega) v_0^{lda} = 0. \quad (4)$$

Если существует обратная матрица к матрице A_ω , то решение уравнения (4) сводится к стандартной задаче собственных значений. Вектор новых факторов получается с помощью соотношения

$$y_i^k = V_{lda}^T x_i^k. \quad (5)$$

Однако, в задачах распознавания лиц, из-за большой размерности исходных данных, LDA страдает от двух типов трудностей: сингулярности и высокой размерности матриц внутриклассовых и межклассовых различий.

Метод главных компонент в рамках подхода «PCA + LDA» применяется для уменьшения размерности пространства признаков от N до r ($r \ll N$). В результате формируется матрица главных компонент V_{pca} , которая имеет размерность $N \times r$. В работах [10-12] описаны различные алгоритмы вычисления главных компонент. Далее следует применение линейного дискриминантного анализа для поиска представления свойств объектов в пространстве признаков меньшей размерности. В результате вычисляется матрица дискриминантных компонент V_{lda} . Таким образом, матрица преобразования W_{lda} в рамках подхода «PCA + LDA» определяется следующим образом

$$W_{lda} = \arg \max_w \frac{|W^T V_{pca}^T A_b V_{pca} W|}{|W^T V_{pca}^T A_\omega V_{pca} W|} = \arg \max_w \frac{|W^T A_b' W|}{|W^T A_\omega' W|}. \quad (6)$$

Столбцами матрицы W_{lda} являются собственные векторы w_{0i}^{lda} , которые получаются в результате решения уравнения

$$(A_b' - \lambda A_\omega') w_0^{lda} = 0. \quad (7)$$

Здесь $A_b' = V_{pca}^T A_b V_{pca}$, $A_\omega' = V_{pca}^T A_\omega V_{pca}$. Задача (7) является обобщенной задачей собственных значений. Традиционно уравнение (7) сводится к стандартной задаче собственных значений путем обращения одной из матриц, например, матрицы A_ω' . Естественно, это возможно только в случае если обращаемая матрица не вырождена. Уравнение (7) при этом преобразуется к виду

$$(A_\omega')^{-1} A_b' w_0^{lda} = \lambda I w_0^{lda}. \quad (8)$$

Решение (8) позволяет определить матрицу дискриминантных компонент, столбцами которой являются собственные векторы уравнения (8) с наибольшими собственными значениями. Количество дискриминантных компонент $p < r$.

Для распознавания изображений может использоваться либо классификатор по ближайшему центру класса, либо классификатор по ближайшему соседу [15].

3. Обобщенный метод Якоби

Для решения уравнения (7) предлагается использовать обобщенный метод Якоби, который не требует обращений матриц и не теряет дискриминантную информацию при вычислении дискриминантных компонент.

Обобщенный метод Якоби реализуется в виде многошагового процесса, на каждом шаге которого обнуляются внедиагональные элементы. Подробное описание алгоритма можно найти в литературе по проблеме собственных значений.

Применительно к задаче распознавания изображений в алгоритм обобщенного метода Якоби вычисления собственных значений внесены небольшие дополнения.

1. Если в процесс вращений диагональные коэффициенты матриц принимают значение близкие к машинному нулю, то величина коэффициентов устанавливается на несколько порядков выше машинного нуля. Например, для расчетов с двойной точностью эта величина находится в пределах $10^{-12} \dots 10^{-14}$.

2. Если в процессе диагонализации невозможно найти коэффициенты обнуляющее внедиагональные коэффициенты сразу двух матриц, процесс диагонализации не прерывается. При этом выполняется попытка обнулить внедиагональные коэффициенты только одной матрицы, Если и это не возможно, то осуществляется переход к следующему шагу.

4. Экспериментальные исследования

Экспериментальное исследование эффективности описанного выше обобщенного линейного дискриминантного анализа проводится с использованием базы изображений лиц «ORL» [16]. База ORL содержит 400 фронтальных изображений 40 человек, с различной мимикой, поворотами и на-

клонами головы. Все изображения в базе полутоновые, с 256 градациями яркости. Размер каждого изображения – 92 x 112 пикселей. Рис. 1 показывает изображения нескольких субъектов базы данных ORL.

Исходный набор изображений делится на обучающую и тестовую выборки. Изображения обучающей выборки формируются из L изображений каждого класса. Все оставшиеся изображения составляют тестовую выборку.

В первой серии экспериментов сравниваются два классификатора: классификатор по ближайшему центру класса и классификатор по ближайшему соседу. Для исследования качества распознавания используется процедура кросс-валидации, усредняющая коэффициенты распознавания, полученных при различном делении набора изображений на учебную и тестовую выборки. Эксперименты проводятся для учебных выборок, содержащих 2, 3 и 4 изображений в каждом классе базы ORL, которые выбираются случайно. Все оставшиеся изображения составляют тестовую выборку. Таким образом, обеспечивается не пересечение учебной и тестовой выборок. Учебные выборки при этом содержат 80 (2x40), 120 (3x40) и 160 (4x40) изображений, а тестовые – 320, 280 и 140 изображений, соответственно.

В ходе эксперимента главные компоненты вычисляются по обучающей выборке и далее используются для сокращения размерности обучающих изображений. Редуцированные изображения используются для формирования межклассовых и внутриклассовых матриц и вычисления дискриминантных компонент. Количество дискриминантных компонент определяется как наименьшее число компонент, обеспечивающее максимальную точность распознавания. Варьируя число главных компонент, мы получаем различную точность распознавания. В ходе экспериментов для каждого числа изображений в классе ($L=2, 3, 4$, где L – число изображений в классе) было проведено 15 экспериментов, для каждого из которых учебная выборка формируется путем случайного выбора изображений из базы ORL. Оставшиеся изображения используются для формирования тестовой выборки. Результаты, полученные в ходе экспериментов, обрабатываются и представляются в виде усредненного коэффициента распознавания и среднеквадратического отклонения.

При проведении эксперимента дважды решается задача собственных значений. Первый раз при вычислении главных компонент решается стандартная задача собственных значений. Порядок матриц для различных обучающих выборок варьируется от 80 до 160. Так как порядок матриц небольшой и матрицы хорошо обусловлены, поэтому для вычисления главных компонент можно использоваться метод Хаусхолдера. Второй раз задача собственных значений возникает при вычислении дискриминантных компонент и представляет обобщенную задачу собственных значений. Порядок матриц равен числу главных компонент и не превышает 70. Когда число главных компонент меньше числа классов, матрицы внутриклассовых различий и межклассовых различий хорошо обусловлены. В этом случае все собственные значения уравнения (8) положительные. В случае, когда число главных компонент больше числа классов матрицы внутриклассовых и межклассовых различий становятся полуопределенными. В этом случае с решением уравнения (7) справляется обобщенный метод Якоби.

На рис. 2 — 4 представлены усредненные коэффициенты распознавания (k_{test}) тестовой выборки в зависимости от числа главных компонент (r) для трех значений L (2, 3, 4). Коэффициент распозна-

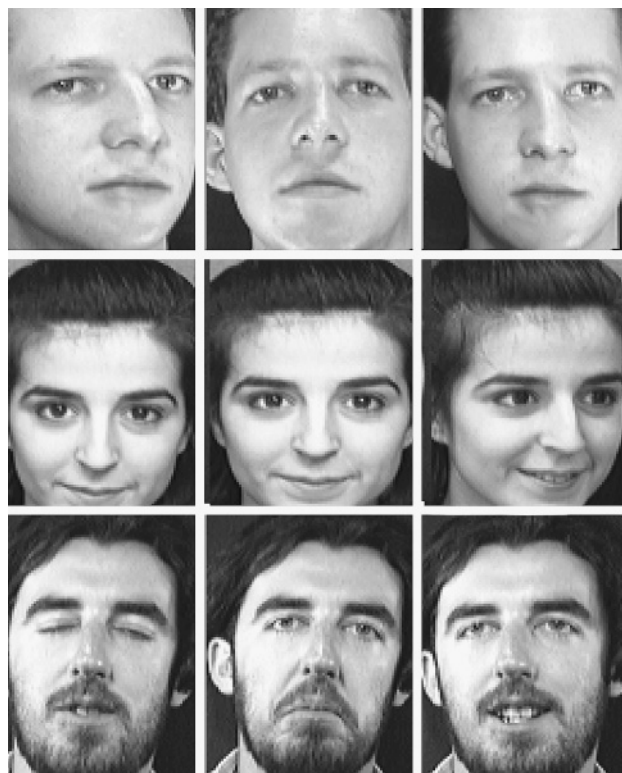


Рис. 1. Примеры лиц базы данных ORL

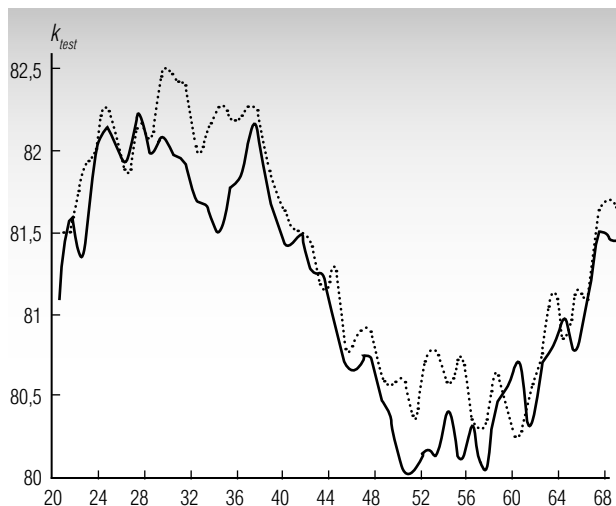


Рис. 2. Точность распознавания тестового набора ($L = 2$)

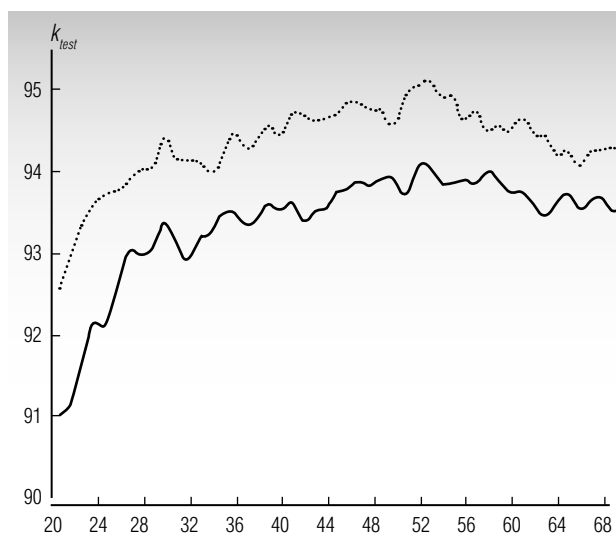


Рис. 3. Точность распознавания тестового набора ($L = 3$)

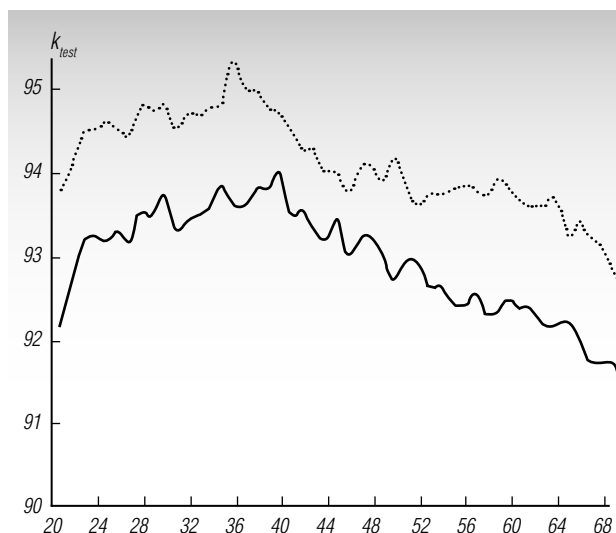


Рис. 4. Точность распознавания тестового набора ($L = 4$)

вания рассчитывается как отношение количества правильно распознанных тестовых изображений к общему числу тестовых изображений. Сплошной линией показаны результаты, полученные при использовании классификатора по ближайшему центру класса, а штриховой линией – результаты, полученные с помощью классификатора по ближайшему соседу.

Анализ полученных результатов показывает, что классификатор по ближайшему соседу дает более высокую точность распознавания.

Качество представленного алгоритма оценивается путем сравнения его с другими алгоритмами, такими как PCA, LDA, NLDA 2DPCA, RLPDA (Regularized Locality Preserving Discriminant Analysis), NDLPP (Null space Discriminant Locality Preserving Projections).

PCA, LDA, NLDA методы были описаны выше. 2DPCA представляет двумерный анализ главных компонент, который позволяет снизить трудоемкость обработки большого числа изображений, и показывает хорошие результаты для SSS проблемы. Описание его можно найти в работе [12]. Методы RLPDA, NDLPP описываются в работах [13, 14] и демонстрируют высокую эффективность при решении SSS проблемы.

Результаты распознавания изображений лиц из базы данных ORL с помощью алгоритма 2DPCA взяты из работы [12], а для алгоритмов PCA, LDA, RLPDA, NDLPP, NLDA – из работы [14]. В табл. 1 представлены результаты сравнения точности распознавания различных методов, где L – это количество образцов каждого класса в учебной выборке, N_{teach} – количество изображений лиц в учебной выборке и N_{test} – количество изображений лиц в тестовой выборке.

Таблица 1.

Результаты сравнения точности распознавания различных алгоритмов

Алгоритм	Вариант ($L/N_{teach}/N_{test}$)		
	2/80/320	3/120/280	4/160/240
PCA [14]	69,5±2,38	78,6±2,59	83,58±1,77
LDA [14]	80,1±2,44	87,9±2,23	91,5±1,85
2DPCA [12]	82,9±3,42	90,2±1,32	92,2±1,53
RLPDA [14]	80,7±2,96	90,4±2,59	94,8±1,79
NDLPP [14]	83±2,31	91,3±1,57	94,7±1,64
NLDA [14]	81,8±2,61	91±2,11	94,4±1,21
Обобщенный LDA	82,5±1,55	91,4±2,59	95,2±1,06

Из табл. 1 видно, что с увеличением количества лиц в классе обучающего набора возрастает точность распознавания для всех алгоритмов. Обобщенный линейный дискриминантный анализ демонстрирует точность распознавания, которая в случае двух изображений в каждом классе обучающей выборки немного уступает алгоритмам 2DPCA и NDLPP, при этом в остальных случаях не уступает другим методам.

5. Заключение

В работе показано применение линейного дискриминантного анализа, позволяющего не терять дискриминантную информацию благодаря использованию обобщенного метода Якоби для вычисления дискриминантных компонент. Экспериментальные результаты на базе данных ORL показывают, что предлагаемый вариант LDA достигает более высокого качества распознавания, чем классический вариант LDA и не уступает другим методам распознавания, что говорит о его высокой эффективности. ■

Литература

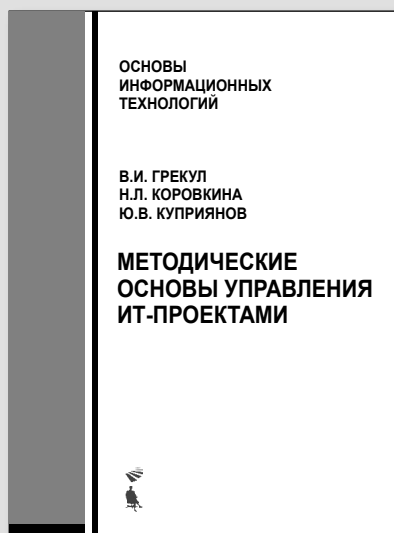
1. Etemad K. Chellappa R. Discriminant Analysis for Recognition of Human Face Images // Journal of the Optical Society of America. – 1997. – Vol. 14. – No. 8. – P. 1724-1733.
2. Raudys S.J., Jain A.K. Small sample size effects in statistical pattern recognition: recommendations for practitioners // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1991. – 13. – P. 252–264.
3. Martinez A.M., Как А.С. PCA versus LDA // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2001. – 23 (2). – P. 228-233.
4. Belhumeur P.N., Hespanha J.P., Kriegman D.J. Eigenfaces vs Fisherfaces: recognition using class specific linear projection // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1997. – 19. – P. 711–720.
5. Yu H., Yang J. A direct LDA algorithm for high-dimensional data with application to face recognition // Pattern Recognition. – 2001. – 34 (10). – P. 2067–2070.
6. Yang, J., Yang J.Y. Why can LDA be performed in PCA transformed space? // Pattern Recognition. – 2003. – 36 (2). – P. 563–566.

7. Dai D.Q., Yuen P.C. Regularized discriminant analysis and its applications to face recognition // Pattern Recognition. – 2003. – 36 (3). – P. 845–847.
8. Pima I., Aladjem M. Regularized discriminant analysis for face recognition // Pattern Recognition. – 2004. – 37 (9). – P. 1945–1948.
9. Chen L.F., H Liao.Y.M., Lin J.C., Kao M.D, Yu G.J. A new LDA based face recognition system which can solve the small sample size problem // Pattern Recognition. – 2000. – 33 (10). – P. 1713–1726.
10. Мокеев В.В. О повышении эффективности вычислений главных компонент в задачах анализа изображений // Цифровая обработка сигналов. – 2011. – №4. – С. 29-36.
11. Мокеев А.В. О точности и быстродействии метода синтеза главных компонент // Бизнес-информатика. – 2010. – № 3 (13). – С. 65-68.
12. Щеголева Н.Л., Кухарев Г.А. Применение алгоритмов двумерного анализа главных компонент для задач распознавания изображений лиц // Бизнес информатика. – 2011. – №4 (18). – С. 31-38
13. Yang L., Gong Wj., Gu X., et al. Null space discriminant locality preserving projections for face recognition // Neurocomputing. – 2008. – 71 (16). – P. 3644-3649.
14. Gu X., Gong W., Yang L. Regularized locality preserving discriminant analysis for face recognition // Neurocomputing. – 2011. – 74 (17). – P. 3036-3042.
15. Форсайт Дж., Понс Ж. Компьютерное зрение современный подход. – М.: Изд. дом «Вильямс», 2004.
16. The ORL Database of Faces – база изображений лиц, подготовленная в научно-исследовательской лаборатории компании Olivetti (Olivetti Research Laboratory), содержит 400 фронтальных изображений лиц 40 человек.

**МЕТОДИЧЕСКИЕ ОСНОВЫ УПРАВЛЕНИЯ
ИТ-ПРОЕКТАМИ**
Учебник

В.И. Грекул, Н.Л. Коровкина, Ю.В. Куприянов

*Интернет-университет информационных технологий: Бином.
Лаборатория знаний, 2011.*



При создании ИТ-решений перед всеми сторонами, вовлеченными в жизненный цикл проекта, возникает целый ряд вопросов, связанных с определением и детальным структурированием необходимых работ, с распределением прав и обязанностей, с управлением и контролем за исполняемыми работами. Одним из действенных инструментов для решения данных вопросов является использование унифицированных подходов, закрепленных в современных международных и российских стандартах и методологиях управления проектами. Представленный учебник содержит детальное описание процедур управления проектами внедрения информационных технологий. Отличительной особенностью данной книги является изложение материала с привязкой к этапам жизненного цикла создаваемого продукта, а не к фазам некоторого абстрактного проекта. Это позволяет читателю сформировать целостное представление о необходимых в ИТ-проекте управленческих процедурах, а также использовать материал последовательно во времени, по мере перехода от одного этапа технологического цикла создания продукта к другому.

АЛГОРИТМЫ ФОРМИРОВАНИЯ СПЕКТРАЛЬНОГО ПРЕДСТАВЛЕНИЯ ЗВУКОВОГО СИГНАЛА НА ОСНОВЕ U-ПРЕОБРАЗОВАНИЯ

В.Е. Гай,

*кандидат технических наук, доцент кафедры вычислительных систем и технологий
Нижегородского технического университета им. Р.Е. Алексеева*

E-mail: vasily.gai@gmail.com

Адрес: г. Нижний Новгород, ул. К. Минина, д. 24

В статье рассматриваются алгоритмы построения многоуровневого (грубо-точного) представления сигнала с помощью фильтров Уолша системы Хармута. Приводятся особенности реализации алгоритмов на ЭВМ. Преимущества предложенных алгоритмов построения разложения: низкая вычислительная сложность и простота реализации.

Ключевые слова: цифровая обработка сигналов, преобразование Уолша, U-преобразование.

1. Введение

Под аналоговым сигналом понимается одномерная функция $F(t)$, где t – временная координата. Под дискретным сигналом понимается одномерный набор отсчетов $f = \{F(nT)\}$ полученный выборкой значений непрерывного сигнала F , взятых периодически с интервалом T , $n = 0, 1, \dots$ [1].

При решении задач обработки сигналов к исходному сигналу могут применяться различные преобразования. Например, преобразование Фурье, вейвлет-преобразование, дискретное косинусное преобразование, преобразование Уолша. Обычно, целью применения преобразования является выделение свойств сигнала, необходимых для решения поставленной задачи.

В настоящей работе рассматриваются вопросы построения многоуровневого (грубо-точного) пред-

ставления сигнала с помощью фильтров Уолша системы Хармута [2, 3]. Необходимо отметить, что существует также и точно-грубая концепция формирования разложения сигнала, которая, реализуется, например, в вейвлет-преобразовании.

2. Классическое преобразование Уолша

2.1. Непрерывное преобразование

Непрерывные функции Уолша являются кусочно-постоянными знакопеременными функциями с интервалами определения $z = [-0,5; 0,5]$ или $z = [0; 1]$. Интервал определения функций Уолша можно представить совокупностью $N = 2n$, $n = 1, 2, \dots$ равных подынтервалов, на каждом из которых функции Уолша принимают значения ± 1 [4].

Записанные совместно и пронумерованные функции Уолша образуют базисную систему

$\{WAL_\alpha(z)\}$, $\alpha = 0, 1, 2, \dots$ (α – номер функции Уолша), в которой возможно разложение произвольного сигнала в ряд Фурье-Уолша.

В практике спектральной обработки сигналов часто применяются три системы фильтров Уолша: Пэли, Адамара и Хармута. Каждую из этих систем можно построить и аналитически описать кусочно-постоянными функциями Радамахера $r_k(z)$, интервал определения которых равен интервалу определения функций Уолша:

$$r_0(z) \equiv 1, r_k(z) = \text{sign}(\sin 2^k \pi z), k = 1, 2, \dots,$$

$$\text{где } \text{sign}(x) = \begin{cases} +1, x \geq 0, \\ -1, x < 0. \end{cases}$$

В настоящей работе рассматриваются только функции Уолша системы Хармута. Функции Уолша в этой системе располагаются в порядке увеличения числа знакоперемен на интервале определения. Систему Хармута аналитически можно записать в виде:

$$WAL_\alpha(z) = \prod_{k=1}^n [r_k(z)]^{\langle \alpha_k \rangle} = (-1)^{\sum_{k=1}^n z_k \langle \alpha_k \rangle},$$

$$\alpha = \sum_{k=1}^n \alpha_k 2^{k-1}, \langle \alpha_k \rangle = \alpha_k \oplus \alpha_{k+1}, \alpha_k = 0; 1,$$

где \oplus – операция поразрядного суммирования по модулю два.

Ряд Фурье-Уолша одномерного сигнала $x(t)$, $t \in [0; T)$ имеет вид:

$$F(t) = \sum_{\alpha=0}^{\infty} c_\alpha WAL_\alpha\left(\frac{t}{T}\right), \text{ где спектр Уолша:}$$

$$c_\alpha = \frac{1}{T} \int_0^T F(t) WAL_\alpha\left(\frac{t}{T}\right) dt.$$

2.2. Дискретное преобразование

Дискретное преобразование Уолша основано на дискретных функциях Уолша и выполняется над решетчатым сигналом f , определённым на конечном числе N равноотстоящих точек. При этом число отсчётов сигнала N (размерность выборки сигнала) должно быть двоично-рациональным, т.е. $N = 2^n, n = 1, 2, \dots$.

Дискретные функции Уолша можно получить дискретизацией их непрерывных аналогов. Выбирать моменты отсчётов функций Уолша следует таким образом, чтобы полученные дискретные системы обладали свойствами полноты и ортогональности. Размерность дискретных сигналов равна числу отсчётов N , с помощью которых они зада-

ются. Полная система базисных функций (СБФ) Уолша должна содержать также N функций.

Выберем первые N непрерывных функций Уолша, из которых отберём функцию с наибольшим числом знакоперемен и расположим отсчёты дискретных функций Уолша в точках изменения знака этой функции, а отсчёты у всех N функций Уолша выберем в этих точках отсчёта, приняв их за значения соответствующих дискретных функций Уолша. Полученная, таким образом СБФ, будет полной, ортогональной и пригодной для разложения решетчатых сигналов произвольной формы с конечным интервалом определения.

Дискретные функции Уолша удобно обозначать как $\{WAL_\alpha(i/N)\}$. Здесь i определяет номер точки дискретного интервала определения.

Дискретный сигнал f , определённый на $[0; N)$ можно разложить в дискретный ряд Уолша:

$$f(i) = \sum_{\alpha=0}^{N-1} c_\alpha WAL_\alpha\left(\frac{i}{N}\right),$$

где спектральные коэффициенты:

$$c_\alpha = \frac{1}{N} \sum_{i=0}^{N-1} f(i) WAL_\alpha\left(\frac{i}{N}\right).$$

Далее, для краткости, вместо понятия «система базисных функций» будет использоваться понятие «набор фильтров».

3. U -преобразование

В [2, 3] рассматривается модификация классического преобразования Уолша для обработки изображений (U -преобразование). U -преобразование заключается в формировании многоуровневого (грубо-точного) представления сигнала с помощью фильтров Уолша системы Хармута, причём:

1) для построения каждого уровня разложения используется фильтры одинаковой длины, которые масштабируются до размера анализируемого участка сигнала;

2) сначала фильтры применяются ко всему сигналу, затем – к его частям;

3) части, на которые разбивается сигнал, должны иметь минимальное число общих отсчётов.

Настоящая работа посвящена разработке алгоритмов, предназначенных для формирования U -преобразования одномерного сигнала.

Введём следующие обозначения:

1) U – оператор U -преобразования, применяемый к дискретному сигналу;

2) D – результат разложения сигнала: $D = U(f)$, $D = \{D_{ij}\}$, где D_{ij} – j -ый спектр, находящийся на i -ом уровне разложения, $i \in [0; J-1]$, J – количество уровней разложения, $j \in [0; M_i-1]$, M_i – количество элементов на i -ом уровне разложения;

3) G – набор фильтров Уолша системы Хармута, используемых для построения разложения: $G = \{G_1, \dots, G_W\}$, где W – количество используемых фильтров, $G_i(j)$ – j -ый отсчёт i -го фильтра;

4) $H_i = \{H_{i1}, \dots, H_{iM_i}\}$ – множество сегментов, полученных из исходного сигнала f с шагом L на i -ом уровне разложения, H_{ij} – j -ый сегмент сигнала на i -ом уровне разложения, $H_{ij}(k)$ – k -ый отсчёт сегмента;

5) L_i – длина анализируемого сегмента сигнала на i -ом уровне разложения.

В настоящей работе рассматриваются три алгоритма построения разложения одномерного сигнала:

1) алгоритм формирования K -ичного дерева разложения сигнала (параметры алгоритма: J – количество уровней разложения, K – количество сегментов на i уровне, на которые разбивается сигнал на $(i-1)$ -ом уровне), в вершине дерева расположено разложение исходного сигнала, $M_i = K^i - 1$;

2) алгоритм построения разложения сигнала на одном уровне с использованием сегмента произвольной длины (параметры алгоритма: L – длина сегмента);

3) алгоритм построения разложения сигнала на одном уровне, при формировании которого сигнал разбивается на заданное количество сегментов (параметры алгоритма: M – количество сегментов, на которые разбивается сигнал).

В ряде случаев может использоваться комбинация первого и второго алгоритмов: на основе второго алгоритма формируется первоначальное разбиение сигнала на сегменты, после чего для каждого из них формируется K -ичное дерево разложения с помощью первого алгоритма.

Каждый из алгоритмов должен учитывать возможность того, что длина сегмента может не быть кратной длине фильтра.

3.1. Генерация многоуровневого представления сигнала

Алгоритм состоит из двух этапов:

1) формирование представление сигнала f в виде дерева H , вершина дерева (нулевой уровень) – исходный сигнал f , на i -ом уровне дерева находится K^i сегментов, каждый из которых представляет собой результат разбиения на K частей сигнала (или

его сегмента) находящегося на $(i-1)$ -ом уровне;

2) вычисление разложения сигнала D с помощью фильтров Уолша системы Хармута по полученному дереву H .

Рассмотрим каждый из шагов алгоритма подробнее:

1.1) формирование разложения исходного сигнала (вершины дерева):

1.1.1) $i = 0$, $H_{i1} = f(t)$, $L_i = |f|$, где $|f|$ – количество отсчётов в сигнале f ;

1.1.2) зная длину анализируемого сегмента L_i , сформируем набор фильтров G_i , длина каждого из которых (C_i) равна L_i и вычисляется следующим образом:

$$C_i = W \cdot [L_i / W],$$

где $[\bullet]$ – операция округления дробного числа в меньшую сторону.

Вследствие выполнения операции округления возможна ситуация, в которой длина фильтра будет больше длины анализируемого сегмента. В этом случае нужно дополнить сегмент до длины фильтра нулями или отсчётами, относящимися к следующему сегменту.

Примечания:

1. на нулевом уровне сегмент H_{01} дополняется нулями, при этом, при генерации первого уровня разложения добавленные нули из сегмента должны быть удалены;

2. если используется методика дополнения сегмента следующими отсчётами, то для последнего сегмента на i -ом уровне не существует отсчётов, которыми его можно дополнить, в этом случае координата начала сегмента в сигнале смещается назад на количество отсчётов, которые надо добавить к сегменту, т.е. отсчёты выбираются из предыдущего сегмента.

Выполним генерацию набора фильтров G'_i на основе G :

$$G'_i = GEN_P(G_i),$$

где $GEN_P(\bullet)$ – оператор, который выполняет повтор P раз каждого отсчёта фильтра G_i ($P = C_i / W$).

Например, пусть $G_0 = [1 \ 1 \ 1 \ 1]$,

$$G_1 = [-1 \ -1 \ 1 \ 1], G_2 = [-1 \ 1 \ 1 \ -1],$$

$$G_3 = [1 \ -1 \ 1 \ -1] \text{ и } P = 2, \text{ тогда}$$

$$G'_0 = [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1],$$

$$G'_1 = [-1 \ -1 \ -1 \ -1 \ 1 \ 1 \ 1 \ 1],$$

$$G'_2 = [-1 \ -1 \ 1 \ 1 \ 1 \ 1 \ -1 \ -1],$$

$$G'_3 = [1 \ 1 \ -1 \ -1 \ 1 \ 1 \ -1 \ -1].$$

1.2) формирование разложения исходного сигнала:

1.2.1) $k=1; i=1;$

1.2.2) $\forall j \in [1; M_{i-1}]:$

$$\{H_{ik}, \dots, H_{ik+K-1}\} = DIV(H_{i-1,j}), k = k + K;$$

где $DIV(\bullet)$ – оператор деления сигнала на K частей.

Оператор DIV выполняет следующие действия:

а) определяет координаты сегментов и выделяет сегменты из сигнала;

б) дополняет сегменты отсчётами так, чтобы длина сегмента была кратна длине фильтра, используемого при построении разложения.

1.2.3) генерация набора фильтров G'_i для сегмента длиной L_i ;

1.2.4) $i=i+1;$

Шаги 1.2.1 – 1.2.4 должны выполняться до тех пор, пока не будут сгенерированы N уровней дерева разложения и соответствующие им фильтры.

Минимальный набор фильтров, который может использоваться для построения разложения включает четыре элемента. В связи с этим, построение дерева разложения сигнала должно быть остановлено, если значение C_i станет равным четырём.

2) Вычисление спектрального представления каждого сегмента полученного дерева H :

$$D_{ij}(n) = \frac{1}{L_i} \cdot \sum_{k=0}^{L_i-1} H_{ij}(k) \cdot G'_m(k), n \in [0; W-1].$$

В случае если вычислять обратное преобразование не требуется, коэффициент перед суммой можно опустить.

Примечания:

1. амплитуда анализируемого сигнала для выполнения операции восстановления должна быть нормирована к отрезку $[0;1]$. Данная операция должна быть выполнена до построения разложения сигнала;

2. при построении разложения набор фильтров генерируется для каждого уровня разложения, вследствие чего снижается скорость построения разложения;

3. полученное разложение избыточно из-за перекрытий сегментов на одном уровне. В результате, один и тот же отсчёт исходного сигнала может входить в два сегмента одновременно ($C_j \cdot W_j > |f|$);

4. результаты тестирования показали, что после выполнения всех операций по дополнению сегментов отсчётами, каждый сегмент на i -ом уровне разложения будет включать одинаковое количество отсчётов;

5. выше предложен вариант дополнения сегмента отсчётами из следующего сегмента, в качестве аль-

тернативы, можно использовать отсчёты не только из следующего, но и одновременно из предыдущего и следующего сегментов.

6. части сигнала не будут пересекаться, только в том случае, если длина анализируемого участка сигнала кратна количеству коэффициентов фильтра.

3.2. Формирование представления сигнала на одном уровне (заданная длина сегмента)

Алгоритм состоит из следующих шагов:

1. Формирование набора фильтров G' с учётом длины анализируемого сегмента L (см. шаг 1.2 предыдущего алгоритма);

2. Выборка из сигнала f очередного сегмента H_i ;

3. Сравнение длины фильтра C и длины сегмента L (их размеры могут быть различны вследствие применения операции округления на предыдущем шаге):

3.1 если длина сегмента меньше длины фильтра, то сегмент дополняется отсчётами;

3.2 если длина сегмента больше длины фильтра, сегмент обрезается до длины фильтра.

4. Формирование спектрального представления сегмента H_i :

$$D_j(n) = \frac{1}{L+E} \cdot \sum_{k=0}^{L+E-1} H_j(k) \cdot G'_n(k),$$

$$n \in [0; W-1], E = C - L;$$

5. Шаги 3 – 4 повторяются до тех пор, пока не будут обработаны все сегменты сигнала f .

Примечания:

1. При тестировании алгоритма оказалось, что число добавляемых (удаляемых) отсчётов не превышает половины длины фильтра из набора G ;

2. Если используется методика дополнения сегмента следующими отсчётами, то для последнего сегмента не существует отсчётов, которыми его можно дополнить, в этом случае координата начала сегмента в сигнале смещается назад на количество отсчётов, которые надо добавить к сегменту;

3. В предлагаемом алгоритме, также как и в предыдущем, можно использовать обе методики расширения сегмента отсчётами;

4. При использовании данного алгоритма перекрытия сегментов можно избежать, если использовать значение длины сегмента, кратное количеству используемых фильтров.

Один из вариантов алгоритма формирования раз-

ложения сигнала на одном уровне состоит в разбиении сигнала на заданное число сегментов и вычисления U -преобразования каждого из сегментов. Реализация этого алгоритма описана в шагах 1.2.1-1.2.4 пункта 2.1.

3.3. Обратное U -преобразование

Обратное U -преобразование записывается следующим образом:

$$S' = U^{-1}[D],$$

где S' – сигнал, полученный после выполнения обратного преобразования, U^{-1} – оператор обратного преобразования. Вычислить обратное преобразование можно следующим образом:

$$H'_{ij}(k) = \sum_{k=0}^{W-1} D_{ij}(k) \cdot G(k).$$

Примечания:

1. При выполнении обратного преобразования используется набор фильтров G , а не G' ;

2. Вне зависимости от уровня разложения с помощью обратного преобразования вычисляется W коэффициентов, следовательно, для точного восстановления сигнала необходимо выполнение условия $L_i = W$;

3. Если при выполнении обратного преобразования $L_i \neq W$, то сигнал, являющийся результатом преобразования, будет представлять собой «огрублённую» версию исходного сигнала, диапазон амплитуд исходного и полученного сигналов будет отличаться: у «огрублённого» сигнала максимальное значение амплитуды будет меньше, чем у исходного;

4. Для точного восстановления сигнала, необходимо, чтобы сегменты, по которым построено разложение сигнала не пересекались и полностью покрывали весь анализируемый сигнал;

5. Если при выполнении обратного преобразования $L_i \neq W$, то размер сигнала, полученного в результате обратного преобразования будет меньше размера исходного сигнала, для того чтобы получить сигнал по длине равный исходному, можно применить интерполяцию;

6. Не рекомендуется применять обратное преобразование к разложению сигнала, полученному при построении дерева разложения, в этом случае, при обратном восстановлении сигнала, вследствие пересечения сегментов, возможно дублирование отсчётов, в результате полученный сигнал будет отличаться от исходного.

4. Исследование способов дополнения сегментов

В каждом из разработанных алгоритмов рассматривались две возможные методики дополнения сегментов до нужного размера:

- ◆ дополнение нулями;
- ◆ дополнение отсчётами, относящимися к следующему сегменту.

Рассмотрим особенности каждой методики на примере первого алгоритма построения разложения.

На *рис. 1* приведены два разложения сигнала (см. *рис. 1.а*) на четвёртом уровне (длина сегмента – 24 отсчёта). Первое разложение построено с расширением сегментов до нужного размера нулями (см. *рис. 1.б*), второе – следующими отсчётами сигнала (см. *рис. 1.в*).

Если сравнить первые спектры обоих разложений на четвёртом уровне, то в разложении на *рис. 1.в*

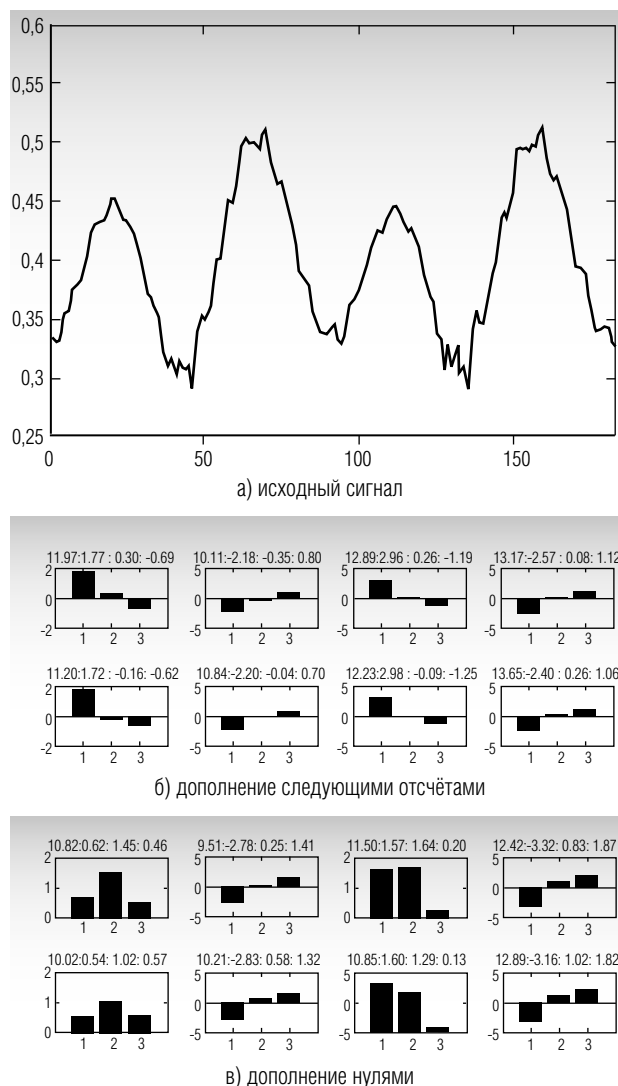


Рис. 1. Исследование методик дополнения отсчётов

структура спектра указывает на то, что в соответствующем спектре сегменте сигнала ровно посередине находится максимум. Это следует из того, что отклик второго фильтра максимален. В действительности, максимум находится в конце сегмента. Следовательно, рассмотренный спектр не отражает реальную картину размещения максимумов в сигнале и дополнение сегмента нулями искажает получаемое разложение.

В разложении на *рис. 1.6* первый спектр указывает на то, что максимум находится левее центра сегмента, второй спектр – на то, что максимум находится правее. Следовательно, данный максимум находится на границе сегментов. Полученный результат достоверен.

Таким образом, при построении разложения сигнала (в не зависимости от используемого алгоритма) будет использоваться дополнение сегмента только следующими отсчётами.

5. Заключение

В работе предлагаются два алгоритма построения *U*-представления одномерного сигнала с помощью фильтров Уолша. Рассматриваются специфические задачи, связанные с построением разложения сигнала.

К достоинствам предложенного разложения можно отнести:

- ♦ простоту реализации: при вычислении разложения используются только операции сложения и сравнения;
- ♦ скорость работы: по сравнению с преобразованием Фурье преобразование Уолша выполняется быстрее, так как в нём выполняется работа с действительными, а не с комплексными числами;
- ♦ инвариантность к амплитуде и частоте сравниваемых сигналов: если сравниваемые сигналы обладают разной частотой, но одинаковой формой, то их разложения будут подобны.

Разработанные алгоритмы использовались автором при решении различных задач обработки сигналов, связанных с исследованием речевых сигналов в [5-10].

Предложенный подход к формированию разложения одномерных данных представляет несомненный интерес не только при анализе звуковых сигналов, но и в анализе экономических временных рядов. Разработанные алгоритмы могут найти своё применение при решении задачи сравнения качественного поведения экономических временных рядов, обладающих различной длиной. ■

Литература

1. Рабинер Л.Р., Шафер Р.В. Цифровая обработка речевых сигналов: Пер. с англ. / Под ред. М.В.Назарова и Ю.Н.Прохорова. – М.: Радио и связь, 1981.
2. Утробин В.А. Информационные модели системы зрительного восприятия для задач компьютерной обработки изображений. – Н. Новгород: НГТУ, 2001.
3. Утробин В.А. Компьютерная обработка изображений. Принятие решений в пространстве эталонов. – Н. Новгород: НГТУ, 2004.
4. Смирнов Ю.М. Проектирование специализированных информационно-вычислительных систем. – М.: Высшая школа, 1984.
5. Гай В.Е. Алгоритм выделения особенностей в речевом сигнале // Информационные системы и технологии: материалы международной научно-технической конференции (г. Нижний Новгород, 23 апреля 2010 г.). – Нижний Новгород, 2010. – С. 361.
6. Гай В.Е. Алгоритм классификации сигналов в системе речевого управления мобильным роботом // Современные проблемы информатизации в анализе и синтезе технологических и программно-телекоммуникационных систем: Сб. трудов. – Вып. 16. – Воронеж, 2011. – С. 316-318.
7. Гай В.Е. Алгоритм сегментации речевого сигнала // Нейрокомпьютеры и их применение: Тезисы докладов Всероссийской конференции. – Москва: 2011. – С. 31.
8. Гай В.Е. Алгоритм классификации сигналов // Труды Нижегородского государственного технического университет им. Р.Е. Алексева. – 2011. – № 1 (86). – С. 29-34.
9. Гай В.Е. Алгоритм оценки степени искажения сигнала // Информационные системы и технологии: Материалы международной научно-технической конференции (г. Нижний Новгород, 22 апреля 2011 г.). – Нижний Новгород, 2011. – С. 394.
10. Гай В.Е. Исследование влияния шума на речевой сигнал // Вестник Нижегородского университета им. Н.И. Лобачевского. – 2011. – № 2 (1). – С. 197-200.

НАДЕЖНОЕ ОБНАРУЖЕНИЕ ПЛАГИАТА МАЛЫМ ЧИСЛОМ ПОИСКОВЫХ ЗАПРОСОВ

В.В. Дягилев,

*аспирант кафедры высшей математики и математического моделирования
Алтайского государственного технического университета им. И.И. Ползунова*

А.А. Цхай,

*доктор технических наук, профессор, заведующий кафедрой математики
и прикладной информатики в экономике Алтайской академии экономики и права*

С.В. Бутаков,

*кандидат технических наук, доцент кафедры математики и прикладной
информатики в экономике Алтайской академии экономики и права*

E-mail: dyagilev@mail.ru, taa1956@mail.ru, sergey.butakov@gmail.com

Адрес: г. Барнаул, проспект Ленина, д. 46

Представлены усовершенствованная архитектура системы обнаружения плагиата и эксперимент, показывающий эффективность предложенного подхода. В предложенном варианте поисковый Интернет-сервис выделен в отдельный модуль, размещаемый на стороне организации-контролера. В отличие от традиционных архитектур предполагается, что контролер не получает проверяемый документ целиком. Проведенный эксперимент демонстрирует преимущества предложенного подхода.

Ключевые слова: определение плагиата, архитектура сервиса.

1. Введение

В данной статье плагиат определен как использование чужих работ без прямого указания источника. Повсеместное развитие сети Интернет способствует распространению плагиата во многих сферах деятельности. Плагиат является острой проблемой, как в образовании, так и в других отраслях. Так в результатах исследований, проведенных в Национальном исследовательском университете «Высшая школа экономики»,

утверждается, что около 50% студентов российских вузов используют рефераты и курсовые работы из сети Интернет [1]. Росфиннадзор отметил проблему плагиата при аудите использования бюджетных средств выделенных на НИОКР в 2009 году [2]. Не менее актуально эта проблема стоит при рассмотрении конкурсных проектов в государственные инновационные программы и фонды, а также при регистрации заявок в Роспатенте. Примитивный плагиат заключается в прямом копировании мате-

риалов источника, в то время как более сложные формы стараются скрыть связь с последним, например, при помощи перефразирования. Само понятие «сходство» может быть также формализовано различными способами [3], включая, например, сходство индексов [4] или семантическое сходство. Само понятие «Сходство документов» не является предметом данной работы и поэтому в данном случае мы ограничимся рассмотрением плагиата как прямого копирования текста из ранее опубликованных документов с минимальными вариациями.

При определении возможных источников плагиата необходимо определиться с границами поиска, а также с исполнителем, чьими силами будет осуществляться этот поиск. Если поиск необходимо производить по коллекции документов какой-нибудь библиотеки, то это один случай. Например, сотрудник университета сможет выполнить такой поиск, используя локальную информационную систему. Другой вариант – это поиск схожих документов, опубликованных на web-страницах в сети Интернет. Для данного вида поиска, чтобы сравнить миллиарды документов, необходимо иметь собственную поисковую машину, периодически сканирующую всю сеть Интернет, либо воспользоваться услугами сторонней организации.

При передаче текстов в стороннюю организацию на проверку неизбежно возникают вопросы, связанные с защитой интеллектуальной собственности. Как поступить в случаях, когда необходимо осуществить проверку текста на наличие плагиата, но при этом желательно ограничить доступ третьей стороны к информации, содержащейся в документе? Подобные ситуации могут возникнуть, например, в случаях обработки заявок на изобретения, полезные модели, промышленные образцы, проверки отчетов по научно-исследовательской работе или при рассмотрении конкурсных проектов в государственные инновационные программы и фонды и т.п. Ограничение при помощи лицензирования, используемое в настоящее время большинством подобных сервисов, должно быть дополнено применением методологии создания средств ограничения доступа.

Использование подобных средств, таких как, например, хэширование и передача хэшей вместо оригинального текста, с одной стороны, гарантирует невозможность восстановления документа, но, с другой стороны, накладывает существенные ограничения на практическое применение данного подхода. Ограничения возникнут в силу того, что проверяющая

организация будет вынуждена поддерживать хэш поискового индекса на своей стороне. При этом основная проблема защиты содержания проверяемого текста не решится, так как поисковый индекс будет в любом случае содержать открытый текст из потенциального источника плагиата. Кроме того, применение хэширования сделает невозможным использовать для целей поиска обычные поисковые системы, такие как Яндекс, Google или Bing.

Возможным вариантом решения проблемы является существенное изменение проверяемого текста таким образом, чтобы его восстановление потребовало бы значительных усилий. Кроме того, такие действия по восстановлению могут быть квалифицированы как неправомерный доступ к интеллектуальной собственности.

Далее предлагается оригинальный способ построения системы обнаружения плагиата, при котором требуемое качество поиска сохраняется, но существенно затрудняется возможность полного восстановления проверяемого текста из переданных данных. Предлагаемая архитектура системы основана на использовании общедоступных поисковых машин Интернета и базового постулата о том, что обнаружение местонахождения потенциальных источников плагиата в сети Интернет обеспечивается использованием ограниченного объема информации из проверяемого документа.

Оставшаяся часть статьи организована следующим образом. Во втором разделе характеризуются типовые схемы системы обнаружения плагиата. В третьем разделе детально описывается предлагаемая клиент-серверная архитектура системы обнаружения плагиата. В заключении приводятся результаты эксперимента, показывающие зависимость качества обнаружения источников плагиата от объема переданной информации и числа запросов, а также делаются выводы о применимости предложенного подхода.

2. Существующие решения

Архитектура сервиса обнаружения плагиата (СОП), приведенная далее, типична для систем, применяемых в таких областях как журналистика, научные исследования и образование. Именно на последнюю область мы будем опираться в качестве примера, но выводы, предложенные в работе, могут быть расширены на другие смежные сферы деятельности. Анализ декларируемых принципов работы и детальное рассмотрение открытых СОП

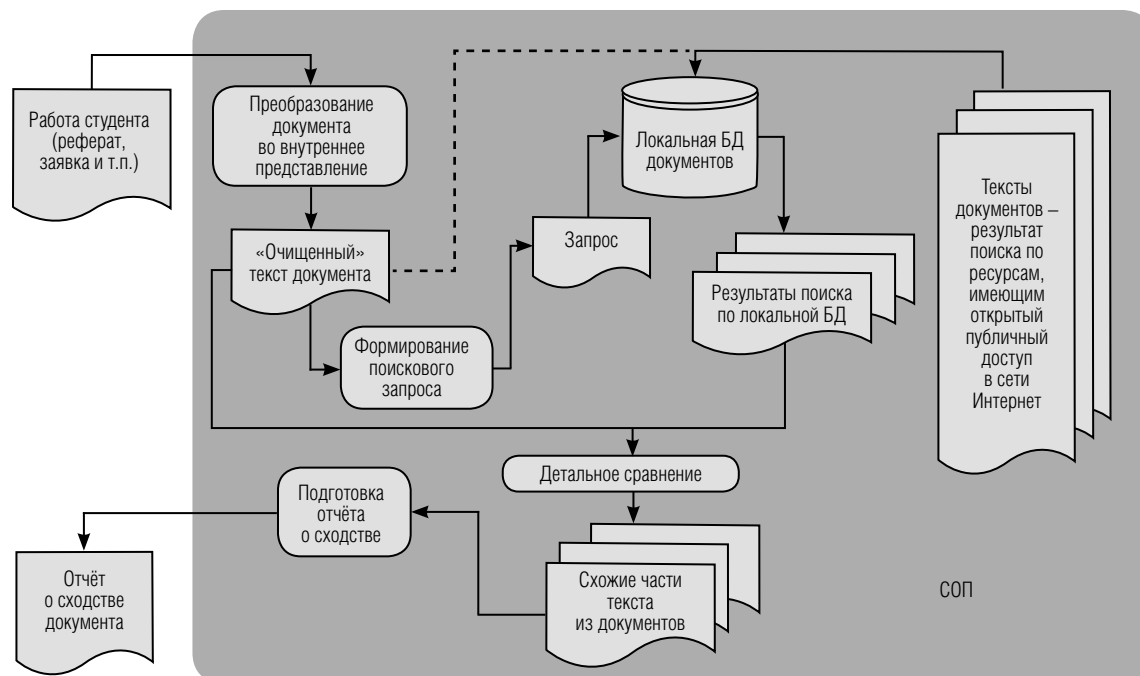


Рис. 1. Типовая архитектура СОП

позволяют выявить их типовую структуру. Последняя может быть охарактеризована описанием продуктов – лидеров рынка, таких как системы «Антиплагиат» (лидер русскоязычного рынка) и «Turnitin» (лидер рынка в англоговорящих странах). Упомянутая структура СОП отображена на схеме (рис. 1).

Пользователь (студент или преподаватель) передает на проверку документ в СОП через информационную систему своего университета, либо напрямую через web-интерфейс СОП. Затем содержимое документа преобразуется системой, с целью выделения «чистого» текста, т.е. избавления от форматирования документа, присущего текстовым процессорам. На основе полученного текста строится запрос к базе данных документов СОП, результатом которого является набор документов, потенциально возможных источников плагиата.

Детальное сравнение текстов документов позволяет определить подобные части и сформировать отчет о найденных совпадениях. В результате, пользователь получает отчет о проведенной проверке, с указанием частей текста и источников «заимствования», если таковое обнаружено СОП. При этом база данных документов СОП может содержать индексы открытых сегментов сети Интернет (как в случае с продуктом «Turnitin»), так и доступ к некоторым библиотекам с ограниченным доступом.

Например, продукт «Антиплагиат» имеет доступ к базе данных диссертаций ВАК РФ.

Рассматривая процесс обнаружения плагиата, следует отметить ключевой момент – содержимое проверяемого документа передается в СОП, то есть третьей стороне. Ограничение при помощи лицензирования [5] не исключает технической возможности неправомерного использования передаваемой информации.

Если исключить применение СОП из процесса, то один из возможных способов поиска плагиата – это использование общедоступных поисковых машин, таких как, например, Яндекс или Google (т.е. использования ограниченного объема информации из текста документа, для того, чтобы получить потенциальные источники плагиата из сети Интернет). Общепринятый способ – это, когда часть текста помещается в кавычки и осуществляется глобальный поиск схожих документов, т.е. производится попытка найти точные совпадения текста в документах, опубликованных в открытом доступе в сети Интернет. Исследования показали, что такая техника поиска, с использованием публичных поисковых машин по критерию точного совпадения фраз, может быть эффективной [6], но такой поиск – медленный, так как осуществляется вручную. При этом остается открытым вопрос определения ключевых фраз в тексте, по которым следует вести поиск.

Для данной цели возможно использование открытого программного обеспечения, например, системы «CROT» [7]. Данная система имеет типовую архитектуру СОП, за тем исключением, что локальная база данных документов состоит только из внутренних ресурсов организации пользователя, а нахождение внешних потенциальных источников плагиата, выполняется путем отправки запросов к поисковой машине Интернета.

Система «CROT» выполняет исчерпывающий поиск, посылая запросы, сформированные «скользящим окном». Алгоритм «скользящего окна» осуществляет прямой перебор фраз. Например, для шекспировской фразы «to be, or not to be: that is the question» при длине окна $X=4$ алгоритм сформирует 7 следующих запросов: «to be or not», «be or not to», «or not to be», «not to be that», «to be that is», «be that is the», «that is the question». Авторы системы «CROT» указывают, что, если значительная часть текста документа была взята из какого-либо источника в Интернете, то нет необходимости посылать все возможные запросы, а достаточно только 10% от этих числа, чтобы определить местонахождение этого источника [7].

3. Предложенная архитектура

На схеме (рис. 2) отображена основная концепция предлагаемой архитектуры СОП. Сам сервис разделен на внутреннюю (клиентскую) часть, работающую на инфраструктуре пользователя, и на внешнюю (серверную) часть, использующую инфраструктуру сторонней организации. Внутренняя часть выполняет функции сервера для обращений со стороны пользователей и одновременно с этим является клиентом, выполняющим запросы к внешней части сервиса.

Предполагается, что разделенная структура сервиса будет выполнять работы в следующем порядке.

1. Клиентская часть системы, получив документ, переданный пользователем для проверки, преобразует его в «чистый» текст.
2. Клиентская часть создает специальный запрос, путем случайного выбора определенного количества запросов, сформированных методом «скользящего» окна.
3. Клиентская часть отправляет специальный запрос в серверную часть СОП.

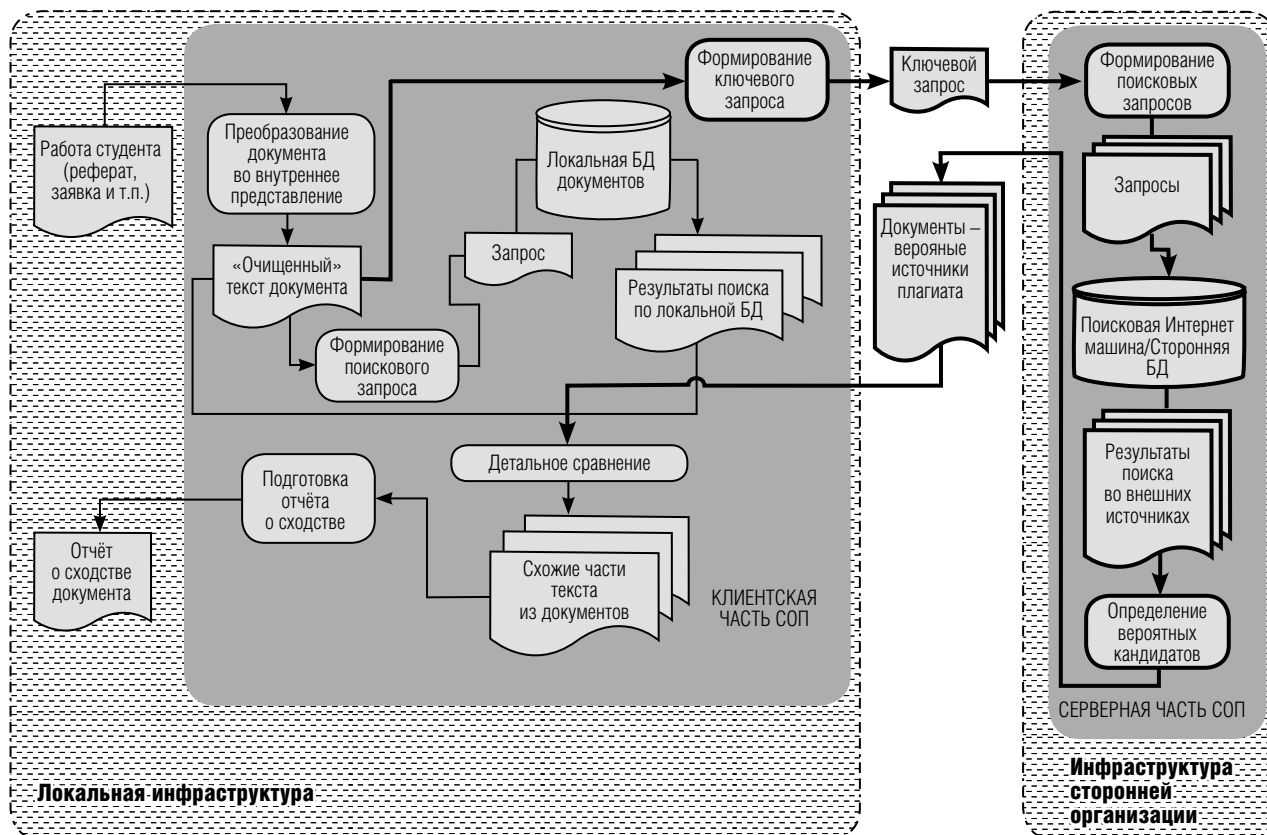


Рис. 2. Схема предлагаемой архитектуры СОП

4. Серверная часть, получив специальный запрос, формирует к поисковой машине Интернет запросы на нахождение документов, опубликованных в открытом доступе в сети Интернет.

5. Серверная часть, как результат выполнения запросов, получает множество ссылок на документы и выбирает те ссылки, которые встречаются чаще всего.

6. Серверная часть по выбранным ссылкам загружает найденные документы, и отправляет последние в клиентскую часть.

7. Клиентская часть производит детальное сравнение исходного текста и полученных документов.

8. Клиентская часть определяет схожие части текста и направляет отчет о сходстве исходного документа пользователю.

Большая часть работы сервиса определения плагиата происходит на оборудовании организации-пользователя. Сторонняя же организация выполняет только глобальный поиск возможных источников плагиата из сети Интернет.

Для составления специального запроса в клиентской части СОП можно использовать алгоритм «скользящего окна», реализованный в системе «CROT». Данный алгоритм осуществляет полный перебор всех фраз длины X , доступных в проверяемом документе. С учетом того, что большинство поисковых машин допускает в обрабатываемых запросах не более 10 слов, то можно ограничить $X \leq 10$.

Очевидно, что при длине текста Y слов, общее количество запросов N , определяется формулой $N = Y - X + 1$, где $Q = \{q_1, q_2, \dots, q_n\}$ – массив запросов. С учетом того, что Y существенно больше X , можно утверждать, что подобный алгоритм сформирует число запросов, близкое к числу слов в документе: $N \approx |Q|$.

Планируя выполнение большого числа запросов к поисковой машине, необходимо рассмотреть возможность появления следующих затруднений:

- ◆ увеличения времени поиска, что обуславливается повышенными требованиями к Интернет – каналу в случае распараллеливания выполнения данных запросов;

- ◆ возможности восстановления документа из фраз, переданных поисковой машине.

Исследования показывают, что порядок передачи поисковых фраз Q не влияет на результаты поиска [7]. Иными словами, если фразы массива Q будут

перемешаны в случайном порядке, то результат поиска совпадет с результатом, полученным при последовательной передаче. Случайное перемешивание не решает проблемы возможного восстановления документа – случайно перемешанная мозаика может быть восстановлена простым перебором, так как в полном массиве соседние элементы q_i и q_{i+1} содержат $X - 1$ совпадающих слов, что делает восстановление тривиальной задачей сбора мозаики фраз по пересечениям из $X - 1$ слов.

Чтобы определить местонахождение источника «заимствования», нет необходимости посылать все возможные запросы Q . Достаточно использовать только небольшой процент случайно выбранных элементов данного массива [7]. Насколько возможно использование этих свойств в случае передачи текста сторонней организации?

Пусть Q_1 – массив случайно выбранных элементов из полного массива запросов $Q: |Q_1| < |Q|$. Общее число слов в запросах, передаваемых в поисковую машину, будет равно $Y_s = |Q_1| * X$. Таким образом, если $|Q_1|$ удовлетворяет неравенству $Y_s < Y$, где Y – общее число слов в документе, то можно гарантировать, что полное восстановление исходного текста на стороне поисковой машины из запросов, переданных ей, становится невозможным.

То есть, если

$$|Q_1| < \frac{Y}{X},$$

то исходный документ – гарантированно невозстановим в первоначальном виде без использования специальных методов лингвистического анализа, активно развиваемых в последнее десятилетие (см., например, [8]). Это ограничение рекомендуется для определения доли запросов при формировании специального запроса, направляемого в серверную часть СОП.

Очевидно, что приведенная архитектура увеличивает требования к мощности вычислительных ресурсов, используемых на стороне клиентской части СОП. Данное требование относится как к повышению вычислительной мощности (требуются вычислительные затраты для детального сравнения документов), так и увеличению дискового пространства для хранения данных документов.

В части дискового пространства в случае использования алгоритмов хеширования, схожих с алгоритмом Винновинг [9], для надежного обнаружения плагиата требуется хранить около 5% хешей, в расчете от числа символов в документе. При исполь-

зовании 128- или 256-битных хешей и однобайтной кодировки текста можно говорить о том, что объем хешей будет примерно равен объему чистого текста в документе. Данное увеличение дискового пространства не представляется сколько-нибудь значимым с учетом постоянно снижающейся стоимости дисковой памяти.

В части вычислительных ресурсов хеширование одного документа не требует существенных вычислительных затрат при использовании локальных алгоритмов, подобных Винновинг [9], так как вычислительные затраты на сравнение документов линейны по отношению к длине текста. Повышенные требования могут предъявляться к СУБД на клиентской части СОП. Фактически именно затраты на приобретение лицензии и обслуживание СУБД будут увеличивать стоимость клиентской части СОП в предложенной архитектуре. Однако большинство организаций уже имеют и поддерживают СУБД, поэтому стоимость дополнительного лицензирования может быть незначительной.

Потенциальным недостатком предложенной архитектуры является то, что она не сможет найти «заимствованный» документ, если он не публиковался в сети Интернет, и поисковые машины Интернета не получали к нему доступ. Однако поисковые машины, такие как Яндекс, Bing или Google, постоянно расширяют свои возможности, получая доступ к научным библиотекам и журналам. Также следует отметить, что по данным специального исследования – ресурсы сети Интернет являются главным источником плагиата [10]. Поэтому авторы склонны считать, что указанный недостаток не является критичным по сравнению с преимуществами, возникающими при использовании предложенной архитектуры СОП.

4. Эксперимент

Главная цель эксперимента была направлена на проверку возможности применения и оценку практического значения предложенного подхода. Основные цели эксперимента:

- ◆ исследовать связь качества поиска с объемом используемой информации в различных случаях;
- ◆ определить объем информации, необходимой для гарантируемого обнаружения плагиата.

Для эксперимента был построен набор документов, имитирующий плагиат. В качестве основы набора был использован оригинальный текст, длиной

около 3000 слов, который никогда не был опубликован в Интернете. Источником плагиата служили тексты 50 документов, сопоставимых по длине и опубликованных в составе общедоступного ресурса Wikipedia.org. Часть текста оригинального документа заменялась текстом из публичного источника. Место вставки «публичного текста» выбиралось случайным образом. Выборка имитировала десять различных уровней плагиата. Размер вставляемых блоков текста изменялся от 5% до 50% общего объема документа. Таким образом, были сгенерированы 500 документов с различной степенью содержания плагиата в них. Далее, для каждого документа были сформированы поисковые запросы методом «скользящего окна» длиной 6 слов. По случайно отобранному запросу осуществлялся поиск в сети Интернет. Объем запросов составлял 5%, 10%, ..., 50% от общего их числа.

На рис. 3 представлена поверхность, отображающая процент обнаружения источников плагиата в сети Интернет в зависимости от количества заимствованного текста и числа запросов, отправленных поисковой машине Интернета.

Светлая часть графика указывает на лучшие результаты. Как видно из параметров графика надежное обнаружение плагиата может быть достигнуто для низкого числа запросов, если существенная часть текста была заимствована из источников сети Интернет. А именно, если больше 50% текста в документе заимствовано, то для того, чтобы достоверно обнаружить плагиат, до-

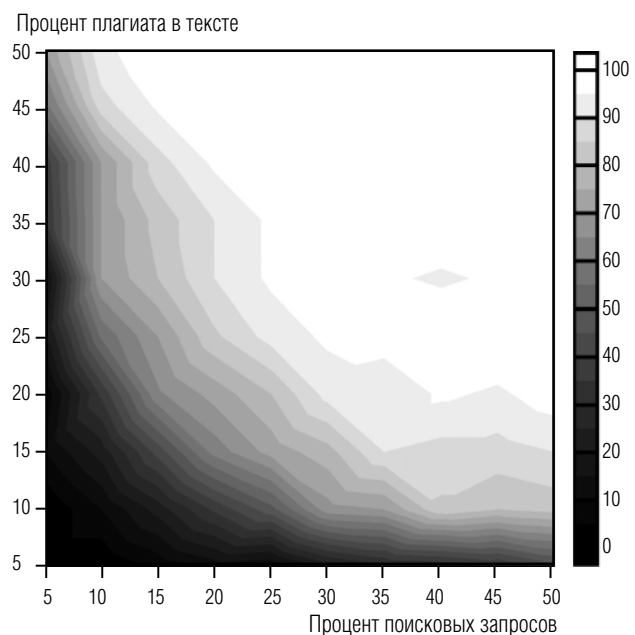


Рис. 3. Результаты обнаружения источников плагиата

статочно будет отправить только 15% запросов, сформированных «скользящим окном». При узко-специфической тематике документа, при отправке тех же 15% запросов достаточно будет уже 25% заимствованного текста для стопроцентного обнаружения плагиата. Это может объясняться тем, что запросы с более специфическими ключевыми словами в большинстве случаев получают более релевантные ссылки на возможный источник плагиата.

Худшие результаты поиска для документов с низким уровнем плагиата могут быть объяснены тем, что фрагменты текста с плагиатом не попали в случайно выбранное число запросов. Можно предположить, что для увеличения качества поиска следует предварительно проанализировать документ, с целью определения сегментов текста, которые наверняка должны участвовать в поисковых запросах. Один из возможных способов реализации – это использование стилеметрии (определения изменений стиля текста) [11]. Использование стилеметрии планируется авторами в качестве одного из направлений развития исследований в данной области.

5. Направления дальнейших исследований

Одно из возможных направлений исследований – это включение стилеметрии [11] (определения стиля текста) во внутреннюю часть СОП. Определение изменений стиля текста позволит улучшить процесс формирования поисковых запросов, что в свою очередь должно положительно отразиться на качестве поиска документов в сети Интернет при малом объеме информации, передаваемой в стороннюю организацию.

Другим возможным направлением продолжения исследований является использование языковой семантики при проведении поиска плагиата. Известно, что избыточность естественных языковых конструкций позволяет восстанавливать тексты по их фрагментам. Известен ряд проектов (например, [12]), использующих семантику для поиска плагиата. Для затруднения восстановления текстов необходимо рассмотреть возможность исключения семантических индикаторов из поисковых запросов. С этой целью будут использованы модели и методы математической лингвистики и математической информатики, в частности, излагаемые в монографии [8]. Такое исключение должно существенно снизить вероятность восстановления исходного текста документа из содержания запросов. Однако нужно исследовать, как это повлияет на качество поиска документов в сети Интернет.

6. Заключение

Предложена новая архитектура СОП, позволяющая определять в проверяемом документе наличие присвоенного материала из текстов, опубликованных в сети Интернет. При этом сторонняя организация, осуществляющая поиск похожих документов, получает не содержимое документа целиком, а только набор поисковых запросов. Качество поиска документов при таком подходе остается на прежнем, высоком уровне.

Проанализированы результаты экспериментального поиска по различным вариантам смоделированного плагиата. Наилучшие результаты получены для поиска в случае узкоспецифических текстов. В дальнейшем планируется совершенствование программной реализации предложенной архитектуры СОП. ■

Литература


1. Ивойлова И. Украденные мысли: Половина студенческих рефератов и курсовых скачивается из Интернета // Российская газета. Федеральный выпуск. – № 4830. – 20 января 2009 г. (<http://www.rg.ru/2009/01/20/referaty.html>).
2. Информационная справка о результатах проверки использования министерствами, ведомствами, внебюджетными фондами и подведомственными им организациями бюджетных средств, выделенных в 2009 году на научно-исследовательские и опытно-конструкторские работы. – Росфиннадзор, 2010. (<http://www.rosfinnadzor.ru>).
3. Федотов А.М., Барахнин В.Б. К вопросу о поиске документов «по аналогии» // Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2009. – Т. 7. – Вып. 4. – С. 5–7.
4. Козлов А.В., Мальцева С.В. Методы повышения эффективности автоматического индексирования документов // Автоматизация и современные технологии, 2004, № 6. – а С. 22-27.

5. Жарова А.К. Правовая защита интеллектуальной собственности / Отв. ред. С.В.Мальцева. – М.: Юрайт, 2011.
6. Culwin F., Child M. Optimizing and Automating the Choice of Search Strings when Investigating Possible Plagiarism // Proceedings of 4th International Plagiarism Conference, Newcastle, June 2010. – 2010.
7. Butakov S., Shcherbinin V. On the Number of Search Queries Required for Internet Plagiarism Detection // Proceedings of the 2009 Ninth IEEE international Conference on Advanced Learning Technologies (July 15 - 17, 2009). – Washington, DC: ICAIT. IEEE Computer Society, 2009. – P. 482-483.
8. Fomichov V.A. Semantics-Oriented Natural Language Processing: Mathematical Models and Algorithms / Series: IFSR International Series on Systems Science and Engineering, Vol. 27. – New York, Dordrecht, Heidelberg, London: Springer, 2010.
9. Schleimer S., Wilkerson D., Aiken A. Winnowing: Local Algorithms for Document Fingerprinting // Proceedings of the ACM SIGMOD International Conference on Management of Data, June 2003. – 2003. – P. 76-85.
10. Jones S., Johnson-Yale C., Millermaier S., Pérez F.S. Academic work, the Internet and U.S. college students // The Internet and Higher Education. – 2008. – Vol. 11. – Issues 3-4. – P. 165-177.
11. Stamatatos E. A survey of modern authorship attribution methods // Journal of the American Society for Information Science. – 2009. – 60. – P. 538–556.
12. Lin W.Y., Peng N., Yen C.C., Lin S.D. Online plagiarism detection through exploiting lexical, syntactic, and semantic information // Proceedings of the ACL 2012 System Demonstrations (ACL '12). – Stroudsburg, PA: Association for Computational Linguistics, 2012. – P. 145-150.

ОСНОВЫ
ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ

В.И. ГРЕКУЛ
Н.Л. КОРОВКИНА
Д.А. БОГОСЛОВЦЕВ
Н.Н. СИНАЙСКАЯ

**АВТОМАТИЗАЦИЯ
ДЕЯТЕЛЬНОСТИ
ПРЕДПРИЯТИЯ
РОЗНИЧНОЙ ТОРГОВЛИ
С ИСПОЛЬЗОВАНИЕМ
ИНФОРМАЦИОННОЙ
СИСТЕМЫ MICROSOFT
DYNAMICS NAV**



**АВТОМАТИЗАЦИЯ ДЕЯТЕЛЬНОСТИ
ПРЕДПРИЯТИЯ РОЗНИЧНОЙ ТОРГОВЛИ
С ИСПОЛЬЗОВАНИЕМ ИНФОРМАЦИОННОЙ
СИСТЕМЫ MICROSOFT DYNAMICS NAV**

**В.И. Грекул, Н.Л. Коровкина,
Д.А. Богословцев, Н.Н. Синайская**

Москва: Бином. Лаборатория знаний, 2009.

В книге рассмотрены процедуры настройки информационной системы Microsoft Dynamics Navision 4.0 и работа в системе при ее использовании в качестве инструмента автоматизации управления деятельностью торговой компании. Книга предназначена для подготовки пользователей системы, а также может использоваться в качестве пособия при проектировании информационных систем автоматизации торговли. В отличие от традиционных инструкций по применению программных продуктов, материал излагается в контексте выполнения задач производственной деятельности, который задается моделями типичных для отрасли бизнес-процессов.

ИНФОРМАТИЗАЦИЯ КОРПОРАТИВНОГО ПЛАНИРОВАНИЯ И БЮДЖЕТИРОВАНИЯ

Д.В. Исаев,

кандидат экономических наук, доцент кафедры бизнес-аналитики

Национального исследовательского университета «Высшая школа экономики»

E-mail: disaev@hse.ru

Адрес: г. Москва, ул. Кирпичная, д. 33/5

В статье рассматриваются вопросы построения комплексных информационных систем, способных в полной мере обеспечить решение задач корпоративного планирования и бюджетирования. Выделены четыре группы задач, выходящих за рамки типовой функциональности специализированных информационных систем планирования и бюджетирования. В качестве комплементарных решений, способных обеспечить полноту решения задач корпоративного планирования и бюджетирования, рассмотрены системы имитационного моделирования, системы статистического анализа данных, системы финансового анализа и моделирования, системы принятия решений на основе экспертных оценок и системы бизнес-интеллекта.

Ключевые слова: планирование, бюджетирование, стратегический менеджмент, система управления эффективностью, информационная система, аналитическое приложение, имитационное моделирование, статистический анализ, финансовый анализ и моделирование, принятие решений, системы бизнес-интеллекта.

1. Введение

Планирование, включая планирование в финансовых терминах (бюджетирование), является одним из основных элементов систем управления предприятиями и организациями. При этом следует отметить объективные различия между стратегически ориентированным планированием корпоративного уровня и планированием текущей деятельности организации.

Планирование корпоративного уровня носит глобальный характер: такие планы формируются для длительных (более года) временных периодов, охватывают всю организацию (или группу) и используют

обобщенные, агрегированные показатели. Что касается планов текущей деятельности, то они являются локальными и краткосрочными, хотя и более детальными по составу показателей. Обе категории планов объективно необходимы, поскольку в совокупности они обеспечивают преимущество системы планирования — от стратегического уровня до уровня отдельных операций. Далее, говоря о системах планирования и бюджетирования, будем иметь в виду стратегически ориентированные планы корпоративного уровня.

В идеале, корпоративные планы и бюджеты должны являться эффективным средством реализации стратегии организации, поскольку именно они при-

званы выразить стратегические цели в виде конкретных плановых показателей и бюджетных статей, с привязкой к соответствующим временным периодам. По сути, планы и бюджеты корпоративного уровня являются связующим звеном между стратегией и повседневной деятельностью организации, тем самым обеспечивая целостность системы управления и способствуя преодолению проблемы «стратегического разрыва» (strategy gap), свойственной многим современным компаниям [7]. Для этого разработан целый ряд подходов к формированию планов и бюджетов, включая гибкие бюджеты, скользящее планирование, функционально-стоимостное бюджетирование, планирование «с нуля» и многие другие [3; 5; 8].

В современной экономической среде, весьма изменчивой и труднопредсказуемой, к корпоративному планированию и бюджетированию предъявляются новые требования, прежде всего, в части гибкости и адаптивности. Это означает многовариантность планирования, основанного на разных сценариях развития компании и внешней среды и, соответственно, на разных наборах планово-бюджетных предпосылок, а также возможность оперативной корректировки планов и бюджетов по мере изменения внешней и внутренней среды организации.

Однако, как показывает практика, системы корпоративного планирования и бюджетирования крайне редко обладают желаемыми свойствами. Это объясняется сложностью планово-бюджетных моделей, высокими трудозатратами на их поддержку, а также недостаточно эффективной организацией взаимодействия специалистов, задействованных в бюджетном процессе. В результате все чаще возникают вопросы о несостоятельности «традиционного» планирования и бюджетирования и целесообразности применения более гибких (хотя и менее детализированных) методов управления, основанных на формировании целевых показателей и делегировании полномочий [10].

Решение или, по крайней мере, значительное смягчение существующих проблем корпоративного планирования и бюджетирования становится возможным при помощи современных информационных систем, позволяющих существенно повысить скорость обработки информации, гарантировать точность расчетов и обеспечить эффективное взаимодействие участников бюджетного процесса. Разумеется, речь не идет о «легких» решениях (к которым, прежде всего, относятся электронные таблицы): учитывая сложность решаемых задач и требования к оперативности обработки управлен-

ческой информации, для корпоративного планирования и бюджетирования следует применять более функциональные программные средства.

В связи с этим приобретают актуальность вопросы о том, какие информационные системы могут применяться для решения задач корпоративного планирования и бюджетирования, как эти системы должны взаимодействовать между собой и как они должны быть интегрированы с внешними информационными объектами – источниками и потребителями управленческой информации. Для ответа на эти вопросы следует рассмотреть роль корпоративного планирования и бюджетирования в системе информационной поддержки корпоративного управления и стратегического менеджмента, определить требования к информатизации планирования и бюджетирования и затем, на основе этих требований, – оценить применимость тех или иных классов информационных систем.

2. Роль корпоративного планирования и бюджетирования в системе управления

С точки зрения информатизации управления, система корпоративного планирования и бюджетирования является частью более общей системы – системы информационной поддержки корпоративного управления и стратегического менеджмента (СИП КУСМ). В соответствии с авторским определением, СИП КУСМ представляет собой «комплекс средств, направленных на решение задач сбора, хранения, аналитической обработки и представления информации, являющейся ключевой для обеспечения информационной прозрачности организации и поддержки принятия стратегических управленческих решений внешними и внутренними заинтересованными лицами» [2, с.103].

Концептуальная информационно-логическая модель СИП КУСМ включает четыре крупных функциональных блока, одним из которых является блок корпоративного планирования и бюджетирования (к другим блокам относятся: блок стратегического анализа и стратегического выбора, блок целевого управления и блок корпоративной отчетности) [2; 11]. Роль блока корпоративного планирования и бюджетирования в СИП КУСМ заключается в том, что он обеспечивает преемственность управления и непротиворечивость управленческой информации, связывая стратегические цели и ключевые показатели организации с более детальными финансовыми и операционными планами.

В состав блока корпоративного планирования и бюджетирования входят два функциональных модуля: модуль формирования планов и бюджетов и модуль анализа выполнения планов.

Модуль формирования планов и бюджетов предназначен для формирования корпоративных планов (в операционных терминах) и бюджетов (в финансовых терминах), которые впоследствии становятся основой для текущего планирования. Таким образом, в качестве исходящего информационного потока модуля выступают сформированные корпоративные планы и бюджеты.

В качестве исходных данных для формирования планов и бюджетов выступают целевые значения ключевых показателей, принятые к реализации сценария развития, а также прогнозы состояния внешней и внутренней среды организации. В случае формирования планов на основе фактических данных предшествующих периодов («от достигнутого») также используется учетная информация. Среди перечисленных входящих информационных потоков блока особую роль играют сценарии развития, поскольку именно они являются основой для формирования бюджетных предпосылок. Множественность рассматриваемых сценариев обеспечивает гибкость планирования — возможность составления разных вариантов планов для разных наборов предпосылок.

Модуль анализа выполнения планов позволяет фиксировать отклонения фактических значений статей корпоративных планов и бюджетов от соответствующих плановых значений, а также выявлять причины этих отклонений. Результаты такого анализа представляют собой исходящий информационный поток модуля. Что касается входящей информации модуля, то она представлена плановыми значениями показателей, сформированными модулем формирования планов и бюджетов, а также фактическими значениями тех же показателей, поступающими из учетных систем организации.

Таким образом, блок корпоративного планирования и бюджетирования связан входящими и исходящими информационными потоками с другими функциональными блоками СИП КУСМ, а также с другими системами управления предприятия. В частности, блок стратегического анализа и стратегического выбора является поставщиком прогнозных данных и потребителем результатов анализа выполнения планов, а блок целевого управления предоставляет информацию о сценариях развития и целевых значениях ключевых показателей. Кроме

того, в группах компаний может быть задействован блок корпоративной отчетности: это происходит в случае децентрализованной схемы корпоративного планирования и бюджетирования, когда бюджетные модели создаются на уровне отдельных бизнес-единиц и впоследствии консолидируются на уровне группы в целом.

Что касается других систем управления организации, то основными потребителями сформированных корпоративных планов и бюджетов являются системы текущего планирования, а в роли поставщиков входящей управленческой информации выступают учетные системы.

Приведенная информационно-логическая модель может служить основой для обсуждения конкретных информационных систем, способных обеспечить практическую реализацию функций корпоративного планирования и бюджетирования.

3. Специализированные информационные системы планирования и бюджетирования

В качестве первого (и наиболее важного) класса информационных систем, применяемых для решения задач планирования на корпоративном уровне, рассмотрим специализированные системы планирования и бюджетирования. В настоящее время такие системы присутствуют в продуктовых линейках многих мировых разработчиков программного обеспечения — Oracle, SAP, IBM, SAS и других. Более того, часто такие системы позиционируются в качестве составных частей комплексных систем управления эффективностью бизнеса (Business Performance Management, BPM), называемых также системами управления корпоративной эффективностью (Corporate Performance Management, CPM) или системами управления эффективностью предприятия (Enterprise Performance Management, EPM) [1; 4; 6; 7].

Отметим, что специализированные системы планирования и бюджетирования существенно отличаются от модулей планирования, присутствующих в составе систем управления ресурсами предприятий (ERP-систем). Причина различий состоит в том, что если ERP-системы обеспечивают текущее (краткосрочное) планирование в рамках отдельных подразделений с применением детальных показателей, то планирование и бюджетирование корпоративного уровня охватывает всю компанию (группу) и оперирует агрегированными показателями в рамках достаточно длительных плановых периодов.

Как правило, специализированные системы корпоративного планирования и бюджетирования основаны на принципах многомерного анализа управленческой информации и разрабатываются с применением технологий аналитической обработки данных в режиме реального времени (On-Line Analytical Processing, OLAP). Аналитические измерения многомерных баз данных («кубов»), в которых хранится управленческая информация, имеют вполне определенную интерпретацию, отражающую логику формирования планов и бюджетов. В частности, все полнофункциональные системы корпоративного планирования и бюджетирования способны поддерживать такие измерения, как перечень планово-бюджетных статей (план счетов), годы и периоды планирования (календарь), валюты и их курсы, элементы организационной структуры (центры финансовой ответственности), сценарии планирования, версии планов и бюджетов, виды продукции, географические сегменты и т.п. При этом некоторые из аналитических измерений (например, план счетов и календарь планирования) считаются обязательными, а другие могут определяться пользователями, в зависимости от специфики задач конкретной организации.

Многие из систем корпоративного планирования и бюджетирования также способны поддерживать функции управления процессами формирования планов и бюджетов, включая многоступенчатые процедуры согласования и утверждения планово-бюджетных показателей.

Еще одним важным свойством систем этого класса является возможность их интеграции с другими системами управления, в частности, с системами управления по ключевым показателям, системами консолидации финансовой отчетности (эти два класса систем также позиционируются как компоненты систем управления эффективностью бизнеса), а также учетными системами и системами текущего планирования, которые часто представляют собой модули систем управления ресурсами предприятий (ERP-систем).

Таким образом, специализированные информационные системы планирования и бюджетирования в целом соответствуют требованиям соответствующего функционального блока СИП КУСМ. В качестве информационно-технологических решений, которые в той или иной мере могут считаться альтернативой специализированным системам планирования и бюджетирования, могут рассматриваться электронные таблицы, надстройки над

ERP-системами и заказные разработки. Однако все они по разным причинам значительно уступают специализированным системам планирования и бюджетирования по своей функциональности, масштабируемости и эксплуатационной надежности. Поэтому именно специализированные системы могут считаться основным классом инструментальных средств, применяемых для корпоративного планирования и бюджетирования.

В то же время можно выделить четыре группы задач, которые представляются достаточно важными, но которые не покрываются (или покрываются в недостаточной степени) функциональностью специализированных информационных систем корпоративного планирования и бюджетирования.

Задачи первой группы, так или иначе, связаны с проблемой формирования планово-бюджетных предпосылок — исходных данных, на основании которых формируются планы и бюджеты. К числу таких задач, в частности, относятся задачи прогнозирования, финансового моделирования (например, построения схем распределения косвенных затрат), а также описания сложных причинно-следственных взаимосвязей между планово-бюджетными статьями (такие взаимосвязи могут обладать нелинейностью, носить вероятностный характер и обладать временными лагами).

К числу задач второй группы относятся задачи, связанные с принятием решений, т.е. с выбором из двух или более альтернатив. Такие задачи могут возникать при оценке правдоподобности разработанных сценариев и сформированных наборов предпосылок, а также при определении политики организации в части распространения планово-бюджетных статей на низовые звенья организации. В таких случаях последствия принимаемых решений далеко не всегда могут быть «просчитаны» экономически, часто требуется принимать во внимание трудноформализуемые факторы. Поэтому решение задач выбора может основываться на экспертных оценках и их последующей обработке, что тоже выходит за пределы функциональности специализированных информационных систем планирования и бюджетирования.

К третьей группе относятся задачи анализа исполнения планов и бюджетов. Разумеется, системы планирования и бюджетирования позволяют фиксировать отклонения «факта» от «плана», как в абсолютном выражении, так и в процентах. В то же время их функциональность оказывается недо-

статочной для углубленного анализа причин возникновения таких отклонений.

Наконец, к четвертой группе задач следует отнести задачи, связанные с визуализацией управленческой информации, необходимой для представления планово-бюджетной информации в наглядном виде, удобном для осмысления и принятия соответствующих решений.

Сказанное позволяет сделать вывод о том, что специализированные системы планирования и бюджетирования, занимая центральное место в комплексе решений для стратегически ориентированного планирования на корпоративном уровне, все же не могут считаться самодостаточными. Поэтому они должны дополняться некоторыми другими комPLEMENTАРНЫМИ аналитическими приложениями, способными обеспечить информационную поддержку ряда вспомогательных, но, тем не менее, достаточно важных функций.

4. Комплементарные информационные системы

Рассмотрим основные классы аналитических информационных систем прикладного характера (аналитических приложений), которые могут быть полезными при решении задач, не охватываемых стандартной функциональностью специализированных систем планирования и бюджетирования.

Для решения задач формирования планово-бюджетных предпосылок весьма полезными могут оказаться три класса аналитических приложений: системы имитационного моделирования, системы статистического анализа данных и системы финансового анализа и моделирования.

Системы имитационного моделирования являются универсальными, они позволяют строить модели сложных систем (в том числе социально-экономических) с учетом сложных взаимосвязей между их элементами. Возможность моделирования внешней и внутренней среды организации дает возможность применять такие системы для прогнозирования и, соответственно, для формирования предпосылок для последующего формирования корпоративных планов и бюджетов. Системы имитационного моделирования довольно разнообразны и могут основываться на разных экономико-математических методах и моделях. В частности, такие системы могут применять методы системной динамики [9] (примером может служить программный продукт Powersim Studio) или методы Монте-Карло (пример – Oracle Crystal Ball).

Системы статистического анализа данных также являются универсальным решением, применимым в самых разных областях знаний. С точки зрения задач планирования и бюджетирования такие системы также позволяют строить прогнозы (на основе исторических данных) и, соответственно, тоже могут считаться важным инструментом для формирования планово-бюджетных предпосылок (примером систем этого класса является IBM SPSS).

Системы финансового анализа и моделирования, в отличие от систем имитационного моделирования и статистического анализа, являются не универсальными, а предметно-ориентированными. Такие системы весьма разнообразны, но их объединяет то, что все они, так или иначе, предусматривают анализ и моделирование финансовых показателей организации. Такие системы могут использоваться для построения финансовых прогнозов и формирования сценариев (пример – Oracle Strategic Finance), а также для функционально-стоимостного моделирования, которое, в условиях значительного объема косвенных затрат, является важным источником информации для анализа себестоимости и бюджетирования расходов (пример – Oracle Hyperion Profitability and Cost Management).

Для решения задач принятия решений часто применяются методы и модели, позволяющие систематизировать трудноформализуемую информацию. Такие методы предусматривают сбор и обработку экспертных оценок, тем самым позволяя принять во внимание качественные характеристики объектов исследования, опираясь на опыт экспертов и их интуицию. На этой основе можно оценивать параметры планово-бюджетных моделей и производить выбор планово-бюджетного сценария, который будет принят за основу для более детального планирования на нижестоящих уровнях управленческой иерархии. Методы принятия решений на основе экспертных оценок довольно разнообразны, примером может служить метод анализа иерархий [12], реализуемый информационной системой Expert Choice.

Для решения задач анализа исполнения планов и бюджетов могут применяться уже рассмотренные выше системы статистического анализа данных и системы финансового анализа и моделирования. Используя разные методы и модели, эти системы могут применяться для описательного анализа, а также для выявления тенденций развития на основе исторических данных, включая анализ причин отклонений фактических значений показателей от плановых.

Наконец, для задач визуализации управленческой информации в качестве основного инструментального средства могут рассматриваться системы бизнес-интеллекта (Business Intelligence, BI). Будучи универсальными средствами сбора, хранения, и представления данных, такие системы позволяют представить в наглядном виде всю необходимую информацию, включая планово-бюджетные предписания, плановые и фактические значения показателей, а также отклонения и результаты их анализа.

Таким образом, в качестве комплементарных информационно-технологических решений, способных расширить и дополнить функциональность специализированных систем планирования и бюджетирования, можно выделить следующие классы аналитических приложений:

- ◆ системы имитационного моделирования;
- ◆ системы статистического анализа данных;
- ◆ системы финансового анализа и моделирования;
- ◆ системы принятия решений на основе экспертных оценок;
- ◆ системы бизнес-интеллекта.

Для формирования полнофункциональной информационной системы все перечисленные элементы должны быть интегрированы как между собой, так и с другими информационными системами, являющимися поставщиками или потребителями соответствующей управленческой информации.

5. Выводы

В качестве главного вывода можно отметить, что традиционные методы и подходы к корпоративному планированию и бюджетированию не исчерпали своих возможностей, даже с учетом новых требований, предъявляемых к системам управления. Однако необходимым условием практической реализации этих методов и подходов является применение соответствующих информационных систем.

Как показывает анализ функциональных возможностей основных классов информационных систем, ни один из них не способен в полной мере обеспечить решение всех задач корпоративного планирования и бюджетирования. Это означает, что необходимая функциональность может быть достигнута только путем построения комплексной системы, включающей информационные системы разных классов. Основная роль в таких комплексах отводится специализированным информационным системам планирования и бюджетирования, а в качестве комплементарных решений выступают системы имитационного моделирования, системы статистического анализа данных, системы финансового анализа и моделирования, системы принятия решений на основе экспертных оценок, а также системы бизнес-интеллекта. ■

Литература

1. Брускин С.Н. Разработка и внедрение систем управления финансовой эффективностью // Бизнес-информатика. – 2010. – №2 (12). – С. 50-53.
2. Исаев Д.В. Корпоративное управление и стратегический менеджмент: информационный аспект. – М.: Изд. дом ГУ-ВШЭ, 2010.
3. Немировский И.Б., Старожуков И.А. Бюджетирование. От стратегии до бюджета – пошаговое руководство. – М.: Изд. дом Вильямс, 2008.
4. Управление эффективностью бизнеса: Концепция Business Performance Management / Е.Ю.Духонин, Д.В. Исаев, Е.Л. Мостовой и др., под ред. Г.В.Генса. – М.: Альпина Бизнес Букс, 2005.
5. Хруцкий В.Е., Сизова Т.В., Гамаюнов В.В. Внутрифирменное бюджетирование: Настольная книга по постановке финансового планирования. – М.: Финансы и статистика, 2005.
6. Cokins G. Performance Management: Finding the Missing Pieces (to Close the Intelligence Gap). – New York: John Wiley & Sons, Inc., 2004.
7. Coveney M., Ganster D., Hartlen B., King D. The Strategy Gap: Leveraging Technology to Execute Winning Strategies. – New York: John Wiley & Sons, Inc., 2003.
8. Drury C. Management and Cost Accounting. 7th edition. – London: Cengage Learning EMEA, 2008.
9. Forrester J.W. Industrial Dynamics. – Cambridge, MA: MIT Press, 1961.
10. Hope J., Fraser R. Beyond Budgeting: How Managers Can Break Free from the Annual Performance Trap. – Boston: Harvard Business School Press, 2003.
11. Isaev D. Performance Management Information Support System: A Conceptual Model // European Journal of Economics, Finance and Administrative Sciences. – 2012. – Issue 52. – P. 6-20.
12. Saaty T.L., Peniwati K. Group Decision Making: Drawing out and Reconciling Differences. – Pittsburgh: RWS Publications, 2008.

ИЗМЕРЕНИЕ УРОВНЯ ЗРЕЛОСТИ ПРОФЕССИОНАЛЬНОГО СООБЩЕСТВА

Ф.В. Краснов,

кандидат технических наук, директор проектного офиса «Виртуальное Сколково»,
Фонд развития центра разработки и коммерциализации новых технологий – ФРЦ
РИКНТ (Фонд Сколково)

Р.Э. Яворский,

кандидат физико-математических наук, менеджер по связям с партнерами,
Фонд развития центра разработки и коммерциализации новых технологий –
ФРЦ РИКНТ (Фонд Сколково)

E-mail: fk@sk.ru, ryavorsky@sk.ru

Адрес: г. Москва, Краснопресненская наб., д. 12

В статье описана формальная модель онлайн сообщества в терминах комбинации социальной и семантической сетей. Предложен подход для оценивания зрелости профессиональных онлайн сообществ, основанный на измерении двух групп параметров, характеризующих уровень компетенции и плотность сети контактов в сообществе. Приведены результаты пилотной апробации предлагаемого подхода для оценивания нескольких городских ИТ сообществ центральной части России. Авторы используют свой опыт и наработки в рамках программы «Виртуальное Сколково» по развитию онлайн сообщества портала <http://sk.ru>.

Ключевые слова: профессиональное сообщество, онлайн сообщество, социальная сеть, семантическая сеть, параметры онлайн сообществ, рейтинг сообществ.

1. Необходимость развития сети профессиональных сообществ в рамках программы «Виртуальное Сколково»

«Виртуальное Сколково» создается как проект национального уровня, опирающийся на федеральные ресурсы и поддержку «экстерриториального» бизнеса [1]. При этом в международном информационном пространстве проект тесно ассоциируется с Москвой и может стать ключевой точкой городско-

го развития — как интерфейс взаимодействия с исследовательскими подразделениями мировых университетов, «фабриками мысли», крупным и средним инновационным бизнесом.

Сообщество — это сложный организм, который состоит из людей разных профессий, разных компетенций, с разными целями и интересами, из компаний (больших и маленьких), университетов, научных центров, технопарков, бизнес-инкубаторов, инвестиционных фондов, и проводимых ими мероприятий.

Одной из функций сообщества является создание:

- ◆ новых специалистов и проектных команд;
- ◆ новых компаний и проектов;
- ◆ новых мероприятий.

Понятие «виртуальное» (в смысле представленное в интернет) не следует понимать, как выстроенное непременно по последнему слову IT-технологий. Виртуальное сообщество Сколково может быть таковым, но не обязано. Более существенная характеристика – активность в виртуальном пространстве – должна по структуре и по назначению соответствовать характеру оффлайн-новых коллективных взаимодействий участников инновационного сообщества. Главное, чтобы исполнялись актуальные задачи и насущные потребности участников, чтобы в виртуальной среде присутствовали процессы, важные сами по себе, но более красивые и операциональные в Интернет. Приоритет при создании виртуального пространства должен быть отдан построению удобной среды социальных отношений и коллективных взаимодействий, удобной для пользователей, как агентов инновационных процессов. И лишь во вторую очередь – новшествам ИКТ.

Для российских условий вопрос профессиональных сообществ – не вопрос ресурсов, и не вопрос законов, а вопрос специализации и компетентности. Основной вопрос будущему участнику профессионального сообщества – кто ты? С кем ты объединяешься? Кто твой начальник? Каково твоё окружение?

Говоря научным языком – это субъектность. В российских условиях для выстраивания отношений важно идентифицировать субъекта с профессиональной стороны, а потом уже с «инвестиционно-деятельностной». Важно определить его вхождение в те или иные «кластеры». Другими словами, для профессионального сообщества первична компетентностная прибыльность от участника.

Сетевой принцип построения в условиях высокой территориальной разделённости скорее необходимость, чем условие. Игнорирование того факта, что участники продуктивно существуют в различных часовых поясах может пагубно сказаться на продуктивности любой онлайн-обсуждения или мероприятия, поэтому безконтактность взаимодействий является нормой для профессиональных сообществ.

С другой стороны, организация в виде сети под-

разумеет наличие узлов, семантика которых должна быть определена при построении сети. Узел профессионального сообщества – это не индивидуум, как в социальных сетях и не группа по интересам. Наиболее уместным, на наш взгляд, стал территориально-кластерный принцип создания узлов. Поиску подходов к построению формальной модели сообщества посвящена данная статья.

2. Формальная модель онлайн сообщества

Для формального описания мы используем следующую математическую модель профессионального сообщества [2]. Социо-семантической сетью будем называть граф, устроенный следующим образом. Множество вершин делится на два типа – участники сети и контент. Таким образом, в можно явно выделить социальный подграф (вершины «участники» и соединяющие их рёбра) и семантический подграф (вершины «контента» и связывающие их отношения). К каждой вершине приписан набор атрибутов в соответствии с её типом. Рёбра в социо-семантической сети делятся на три категории:

- ◆ рёбра в социальном подграфе (обозначают различные отношения между членами сообщества);
- ◆ рёбра в семантическом подграфе (обозначают взаимные ссылки и другие связи между элементами контента);
- ◆ рёбра, соединяющие участников и контент. В первую очередь это отношение авторства, а также оценки, выставленные участниками элементам контента.

Поскольку нас интересуют активные и динамически развивающиеся сообщества, граф социо-семантической сети, описывающий такое сообщество также будет изменяться со временем за счёт появления новых участников, новых элементов контента и связей между ними. На практике социальный подграф описывает личные контакты между участниками сообщества и его можно аппроксимировать по открытым данным социальных сетей. При этом важно иметь ввиду, что для IT-сообществ эта аппроксимация будет достаточно качественной, в то время как для профессиональных сообществ в других предметных областях, например, в биомедицине или ядерной физике, по различным экспертным оценкам, до 50% членов сообщества никак не представлены ни в одной из социальных сетей.

Атрибуты вершин описывают структуру компетенций участников сообщества. Как правило, эти данные в явном виде не доступны, хотя набор основных компетенций можно определить на основе открытых данных о публикациях, патентах, участии в отраслевых конференциях. Вершины семантического графа соответствуют значимым для сообщества документам в сети. Помимо статей, комментариев и дискуссий по вопросам актуальным для данной предметной области, это могут быть также веб-страницы компаний, календари отраслевых событий, сайты мероприятий и другие ресурсы с уникальным URL. Атрибуты вершин – это ключевые слова и тэги, характеризующие содержание этих материалов.

Возвращаясь к задаче измерения основных характеристик сообщества, данная модель позволяет явно выделить две группы характеристик:

1. «Контакты». Характеристики множества связей в графе (их общее количество, плотность, распределение, динамика изменения и т.п.) При этом термин «контакты» следует понимать достаточно широко, в частности, связи между участниками через взаимные ссылки в создаваемом ими контенте очень важны, независимо от наличия прямых социальных контактов между ними.

2. «Компетенции». Характеристики атрибутов вершин, агрегирующие индивидуальные характеристики участников сообщества и элементов контента.

3. Обзор методик измерения сложных социально-экономических систем

Задача оценки зрелости профессиональных сообществ во многом близка задаче оценки эффективности работы вузов, которую принято решать посредством вычисления рейтингов университетов. Для сравнения мы рассмотрим здесь один из трёх наиболее признанных рейтингов, а именно, Academic Ranking of World Universities (более известный как Шанхайский рейтинг университетов, [3]). Этот рейтинг рассчитывается на основе следующих характеристик:

1. Выпускники-Нобелевские лауреаты (10 процентов).
2. Сотрудники-Нобелевские лауреаты (20 процентов).
3. «Часто цитируемые исследователи» (20 процентов).
4. Статьи в Nature или Science (20 процентов).
5. Индексы цитирования (20 процентов).

6. Размер университета (10 процентов).

Основная часть характеристик нацелена на оценивание именно уровня компетенции сотрудников университета. Возможно, это связано с тем, что уровень коммуникаций внутри университетской среды трудно измерить и во многом он зависит только от размера университета (пункт 6).

Другим примером оценивания характеристик является система сбалансированных показателей, см. например [4]. В рамках этого подхода организация рассматривается и оценивается в четырех измерениях:

1. Финансовое состояние (общепринятые показатели).
2. Позиция организации на рынке (число клиентов, доля рынка и т.д.).
3. Внутренние бизнес-процессы (насколько они настроены, и эффективны).
4. Мотивация, развитие и обучение персонала (управление знаниями).

Если адаптировать этот подход к оценке профессиональных сообществ, то первые два пункта можно заменить оценкой размера и статуса отраслевых компаний, сотрудники которых входят в сообщество. Третий и четвертый пункт в переложении на сообщество являются характеристиками структуры контактов в сообществе и их динамики (дискуссии, мероприятия, совместные проекты).

4. Предлагаемый рейтинг зрелости сообществ «Компетенции + Контакты»

Как уже было отмечено выше, ключевыми компонентами, определяющими уровень зрелости профессионального сообщества, являются компетенции и контакты. В каждой из этих двух групп мы выделяем четыре показателя с равными весами. Последний дополнительный показатель дает возможность учесть факторы, позитивно характеризующие сообщество, но не помещающиеся в рамки восьми основных параметров.

◆ Компетенции (max = 40 баллов)

- ◇ Развитость ИТ бизнеса (10 баллов)
- ◇ Уровень ИТ-образования (10 баллов)
- ◇ Уровень бизнес-образования (10 баллов)
- ◇ Компьютерные науки (10 баллов)

◆ Контакты (max = 40 баллов)

- ◇ Конференции, семинары (10 баллов)

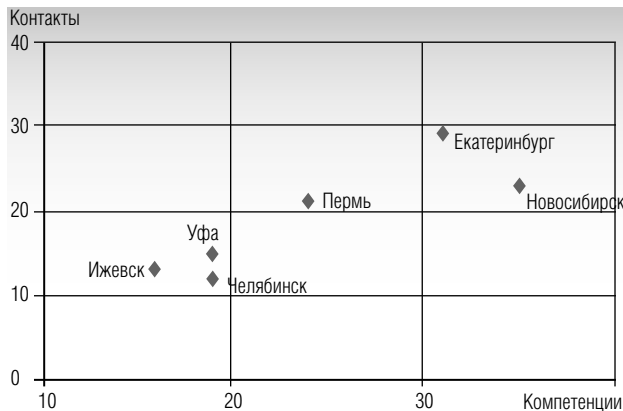


Рис. 1. Экспертные оценки городских профессиональных сообществ в сфере ИТ

- ◆ Специализированные порталы, блоги и форумы (10 баллов)
- ◆ Группы в социальных сетях (10 баллов)
- ◆ Наличие специализированных СМИ (10 баллов)

◆ **Активность и другие бонусы (max = 20 баллов)**

Такое определение рейтинга позволяет системно подойти к задаче оценивания текущего состояния профессиональных сообществ. Тем не менее, предлагаемый подход обладает очевидными недостатками. Во-первых, можно уточнять список параметров и их веса, в частности, учесть финансовые показатели региона, социально-демографическую информацию о регионе, данные о зарегистрированных юридических лицах и т.п. Также требует отдельной проработки методика оценивания каждого параметра. В настоящее время оценки выставляются экспертами, но в будущем хотелось бы избежать, или, по крайней мере, минимизировать субъективность.

5. Апробация рейтинга на примере городских ИТ-сообществ

Пилотная апробация предлагаемого подхода была проведена в октябре-ноябре 2012 года. Мы попросили экспертов, знакомых с ситуацией в разных регионах оценить уровень развития городских ИТ сообществ центральной части России (мы сознательно не включали в это исследование сообщества Москвы и Санкт-Петербурга ввиду их столичной специфики). Агрегированные результаты приведены на графике.

6. Заключение

Построение формальной модели онлайн сообщества представляет еще малоизученную область социально-экономических систем. Предложенный нами подход к построению модели нуждается в более тщательном экспериментальном подтверждении и дальнейшем изучении.

Модели зрелости – это распространенный подход к институцированию творческих деятельностей. Например, такие модели зрелости как СММ и СММИ[5], разработанные по инициативе Software Engineering Institute, широко применяются на практике, поэтому перспективы дальнейшей работы над моделированием и измерением уровня зрелости профессиональных сообществ представляются нам многообещающими.

Мы планируем развивать предложенную методику рейтингования сообществ, двигаясь в сторону замены субъективных экспертных оценок системой формальных показателей, рассчитываемых на основе предлагаемой модели. ■

Литература

1. Краснов Ф.В. Развитие через общение // Intelligent Enterprise. – 2012. – №9. – С. 18-21.
2. Yavorskiy R. Research Challenges of Dynamic Socio-Semantic Networks // CDUD'11–Concept Discovery in Unstructured Data. – Moscow, 2011.
3. Academic Ranking of World Universities (<http://www.shanghairanking.com/>).
4. Громов А.И. Использование системы сбалансированных показателей в стратегическом управлении информационными технологиями корпорации // Бизнес-информатика. – 2010. – №3 (13). – С. 34-40.
5. Ахен Д.М., Клауз А., Тернер Р. СММИ: Комплексный подход к совершенствованию процессов. Практическое введение в модель. – М.: МФК, 2005.

ПРИМЕНЕНИЕ СОВРЕМЕННЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ДЛЯ АНАЛИЗА ПРОСТРАНСТВЕННО РАСПРЕДЕЛЕННЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ДАННЫХ

К.В. Латышев,

аспирант кафедры прикладной математики и аналитической поддержки принятия решений Московского городского университета управления Правительства Москвы

В.Н. Сидоренко,

кандидат экономических наук, кандидат физико-математических наук, кандидат юридических наук, доцент экономического факультета МГУ им. М.В. Ломоносова, доцент кафедры прикладной математики и аналитической поддержки принятия решений Московского городского университета управления Правительства Москвы

E-mail: konstantin73@inbox.ru, v_sidorenko@mail.ru

Адрес: г. Москва, ул. Сретенка, д. 28

В статье описываются современные методы представления и анализа социально-экономических данных с использованием «пространственных» (spatial) OLAP-систем. Для комплексного анализа в режиме «реального времени» социо-экономических данных, привязанных к географическим объектам, предложено современное инновационное, недорогое в своем классе решение на основе широко распространённых, проверенных программных продуктов крупных изготовителей программного обеспечения. В настоящее время в академическом и бизнес-сообществе подобных решений пока нет.

Ключевые слова: пространственная OLAP-система, анализ социо-экономических данных.

1. Введение

В настоящее время крупными государственными структурами и коммерческими компаниями достаточно активно используются современные системы поддержки принятия решений (постановка целей, планирование и под-

держка принятия управленческих решений), обладающие, как правило, средствами предоставления пользователю агрегатных данных для различных пространственно-временных выборок из исходного набора в удобном для восприятия и анализа виде. Какие же современные технологии лежат в основе таких систем? К таким технологиям в настоящее

время относятся OLAP, BI, GIS и ряд других технологий, обзор которых проводится в данной работе.

2. OLAP-технологии

Под OLAP-технологиями (On-Line Analytical Processing) понимают технологии комплексного многомерного анализа данных. OLAP — это ключевой компонент организации хранилищ данных. Концепция OLAP была описана в 1993 году Эдгаром Коддом, известным исследователем баз данных и автором реляционной модели данных (см. E.F. Codd, S.B. Codd, and C.T. Salley, Providing OLAP (on-line analytical processing) to user-analysts: An IT mandate. Technical report, 1993). В 1995 году на основе требований, изложенных Коддом, был сформулирован так называемый тест FASMI (Fast Analysis of Shared Multidimensional Information — быстрый анализ разделяемой многомерной информации), включающий следующие требования к приложениям для многомерного анализа:

- ◆ предоставление пользователю результатов анализа за приемлемое время (обычно не более 5 с), пусть даже ценой менее детального анализа;
- ◆ возможность осуществления любого логического и статистического анализа, характерного для данного приложения, и его сохранения в доступном для конечного пользователя виде;
- ◆ многопользовательский доступ к данным с поддержкой соответствующих механизмов блокировок и средств авторизованного доступа;
- ◆ многомерное концептуальное представление данных, включая полную поддержку для иерархий и множественных иерархий (это — ключевое требование OLAP);
- ◆ возможность обращаться к любой нужной информации независимо от ее объема и места хранения.
- ◆ OLAP-функциональность может быть реализована различными способами, начиная с простейших средств анализа данных в офисных приложениях и заканчивая распределенными аналитическими системами, основанными на серверных продуктах [1].

Как правило, агрегатные функции в OLAP-системах образуют многомерный (и, следовательно, нереляционный) набор данных (называемый гиперкубом или метакубом), оси которого содержат параметры, а ячейки — зависящие от них агрегатные данные. Вдоль каждой оси данные могут быть организованы в виде иерархии, представляющей различные уровни их детализации. Благодаря такой модели данных пользователи могут формулировать сложные запросы, генерировать отчеты, получать подмножества данных [1].

В ячейках OLAP-куба могут содержаться результаты выполнения иных агрегатных функций языка SQL, таких как SUMM, MIN, MAX, AVG, COUNT, а в некоторых случаях — и других (дисперсии, среднеквадратичного отклонения и т.д.). Для описания значений данных в ячейках используется термин *summary* (в общем случае в одном кубе их может быть несколько). С целью обозначения исходных данных, на основе которых они вычисляются, — термин *measure*. Для обозначения параметров запросов применяется термин *dimension* (переводимый на русский язык обычно как «измерение», когда речь идет об OLAP-кубах, и как «размерность», когда речь идет о хранилищах данных). Значения, откладываемые на осях, называются членами измерений (*members*) [2].

Говоря об измерениях, следует упомянуть о том, что значения, наносимые на оси, могут иметь различные уровни детализации. Возможность получения агрегатных данных с различной степенью детализации соответствует одному из требований, предъявляемых к хранилищам данных, — требованию доступности различных срезов данных для сравнения и анализа.

Современные системы OLAP поддерживают «историчность» любых измерений, кроме календарного, т.е. совмещают функции OLAP и хранилищ данных (Data Warehouse). Новейшая реализация OLAP от Microsoft: MS SQL Server 2012 Analysis Services.

3. BI-технологии

Под BI-технологиями (Business Intelligence, бизнес аналитика) чаще всего понимают информационные технологии, созданные для помощи менеджеру в анализе информации о своей компании и её окружении. Термин Business Intelligence впервые появился в 1958 году в статье исследователя из IBM Ханса Питера Луна (англ. Hans Peter Luhn). Он определил этот термин как: «Возможность понимания связей между представленными фактами.» [3] Говард Дреснер (аналитик Gartner) определил «Business intelligence» как общий термин, описывающий «концепции и методы для улучшения принятия бизнес-решений с использованием систем на основе бизнес-данных».

В настоящее время существует несколько вариантов понимания этого термина:

- *Бизнес-аналитика* (в узком смысле) — это методы и инструменты для построения информативных отчетов о текущей ситуации (например, консолидированная отчетность, возникающая в задачах общего менеджмента). Эта информация может оказать

ся жизненно необходимой для принятия текущих управленческих решений.

● *Бизнес-аналитика* (в широком смысле) — это инструменты преобразования, хранения, анализа, моделирования, доставки и трассировки информации в ходе работы над задачами, связанными с принятием решений на основе фактических данных. С помощью этих средств лица, принимающие решения, должны получать нужные сведения в нужное время [4].

Таким образом, BI в узком смысле является лишь одним из секторов бизнес-аналитики в широком смысле. Помимо отчётности туда входят инструменты интеграции и очистки данных (ETL), DWH или OLAP и средства Data Mining [5]. Согласно второму варианту понимания этого термина, BI-технологии позволяют анализировать большие объёмы информации, заостряя внимание пользователей лишь на ключевых факторах эффективности, моделируя исход различных вариантов действий, отслеживая результаты принятия тех или иных решений.

В 2011 г. появились методы анализа данных типа BIG Data на основе технологии Hadoop, которые можно отнести к BI в широком смысле [6], а первый стабильный релиз такой системы (Apache Software Foundation Hadoop) появился в конце 2011 г. В анализе данных типа BIG Data широко применяются методы Data Mining (Интеллектуальный анализ данных), причем модель данных не определяется точно.

4. GIS-технологии [7]

Под GIS-технологиями (GIS или геоинформационные системы) часто понимаются технологии компьютерного кодирования (ввода), сохранения, трансформации, анализа и отображения пространственно-распределенной информации об объектах и событиях. Первые GIS были разработаны в 1950-х и 60-х годах, первоначально в гражданском секторе. В 1970-х и 80-х годах развилась сильная и активная GIS-индустрия.

В настоящее время GIS-технологии объединяют традиционные операции работы с базами данных, такими как запрос и статистический анализ, с преимуществами полноценной визуализации и географического (пространственного) анализа, которые предоставляет карта. GIS применяется в широком спектре задач, связанных с анализом и прогнозом явлений и событий окружающего мира, с осмыслением и выделением главных факторов и причин, а также их возможных последствий, с планированием стратегических решений и текущих последствий предпринимаемых действий.

Современная GIS включает в себя пять ключевых составляющих:

- ◆ аппаратные средства,
- ◆ программное обеспечение,
- ◆ данные,
- ◆ методы обработки данных,
- ◆ исполнители.

Программное обеспечение GIS содержит функции и инструменты, необходимые для хранения, анализа и визуализации географической (пространственной) информации. Ключевыми компонентами программных продуктов являются: инструменты для ввода и оперирования географической информацией; система управления базой данных (DBMS или СУБД); инструменты поддержки пространственных запросов, анализа и визуализации (отображения); графический пользовательский интерфейс (GUI или ГИП) для легкого доступа к инструментам.

Данные о пространственном положении (географические данные) и связанные с ними (геокодированные) табличные данные могут собираться и подготавливаться как самим пользователем, так и приобретаться у поставщиков на коммерческой или другой основе. В процессе управления пространственными данными GIS интегрирует пространственные данные с другими типами и источниками данных, а также может использовать СУБД, применяемые многими организациями для упорядочивания и поддержки имеющихся в их распоряжении данных. GIS хранит информацию в виде набора тематических слоев, которые объединены на основе географического положения.

Современные GIS могут работать с двумя существенно отличающимися типами данных — векторными и растровыми. В векторной модели данных информация о точках, линиях и полигонах кодируется и хранится в виде набора координат X,Y. Местоположение точки (точечного объекта) описывается парой координат (X,Y). Линейные объекты, такие как дороги, реки или трубопроводы, сохраняются как наборы координат X,Y. Полигональные объекты, такие как земельные участки и др., хранятся в виде замкнутого набора координат. Векторная модель особенно удобна для описания дискретных объектов и меньше подходит для описания непрерывно меняющихся свойств, таких как типы почв или доступность объектов. В растровой модели данных изображение представляет собой набор значений для отдельных элементарных составляющих

(ячеек), оно подобно отсканированной карте или картинке. Растровая модель оптимальна для работы с непрерывными свойствами. Обе модели имеют свои преимущества и недостатки.

Современные ГИС подразделяются на ГИС общего назначения и специализированные ГИС. ГИС общего назначения обычно выполняет пять процедур (задач) с данными: ввод, манипулирование, управление, запрос и анализ, визуализацию. Остановимся более детально на двух последних задачах, имеющих отношение к рассмотренным выше OLAP и BI-технологиям.

Запрос и анализ. При наличии ГИС и географической информации можно получать ответы на простые так и более сложные вопросы. Примеры простых вопросов: «Кто владелец данного земельного участка?», «На каком расстоянии друг от друга расположены эти объекты?», «Где расположена данная промзона?». Более сложными вопросами, требующими дополнительного анализа, являются «Где есть места для строительства нового дома?», «Каков основной тип почв под еловыми лесами?», «Как повлияет на движение транспорта строительство новой дороги?». Ответы на поставленные вопросы выполняются с помощью запросов, которые можно задавать как путем выделения тех или иных объектов, так и посредством развитых аналитических средств. С помощью ГИС можно выявлять и задавать шаблоны для поиска, проигрывать сценарии по типу «что будет, если...». Современные ГИС имеют множество мощных инструментов для анализа, среди которых наиболее значимы два: анализ близости и анализ наложения. Для проведения анализа близости объектов относительно друг друга в ГИС применяется процесс, называемый буферизацией. Он помогает ответить на вопросы типа: Сколько домов находится в пределах 100 м от этого водоема? Сколько покупателей живет не далее 1 км от данного магазина? Процесс наложения включает интеграцию данных, расположенных в разных тематических слоях. В простейшем случае это операция отображения, но при ряде аналитических операций данные из разных слоев объединяются физически.

Визуализация. Для многих типов пространственных операций конечным результатом является представление данных в виде карты или графика. Карта — это эффективный и информативный способ хранения, представления и передачи географической (имеющей пространственную привязку) информации.

ГИС тесно связаны с рядом других типов информационных систем. Хотя и не существует единой общепринятой классификации информационных

систем, необходимо отличать ГИС от настольных картографических систем (Desktop Mapping), систем САПР (CAD), дистанционного зондирования (Remote Sensing), систем управления базами данных (СУБД или DBMS) и технологии глобального позиционирования (GPS).

Примером наиболее распространенной ГИС, ставшей мировым промышленным стандартом, является ESRI ArcGIS, использующая технологии управления реляционными базами данных ESRI SDE (Spatial Database Engine), основанными на MS SQL Server.

5. SOLAP-технологии

Естественным развитием OLAP, BI и ГИС-технологий для анализа и планирования «сверху-вниз» стала технология Spatial OLAP data engine, а рамках которой геоданные организованы в гиперкубы по нескольким иерархиям-измерениям (основная иерархия — географическая вложенность, затем следуют: организационное деление, иерархия временных отрезков, классификаторы и т.д.) [8]. Данные технологии активно развиваются в последнее десятилетие. Среди ведущих производителей программного обеспечения, таких как ESRI, Microsoft, IBM, Oracle подобные решения только разрабатываются, хотя, например у GeoMondrian [9] уже реализованы технологии построения SOLAP-кубов.

Какие же задачи могут быть решены при помощи SOLAP-технологий?

Задача 1: Создание визуального интерактивного представления срезов данных аналитического куба в виде матриц (Matrix) или сводных таблиц (Pivot Table) и синхронизированного с ним фрагмента ГеоКарты для аналитического куба, у которого одно из измерений может быть отображено на ГеоКарте в виде вложенных друг в друга слоев (Shapes). Осуществление на фрагменте Геокарты интерактивного выбора уровня иерархии по принципу «сверху-вниз» и выбирать нужные слои карты (Shapes), задавая тем самым фильтр для сопряженной матрицы (Matrix). Изменение при интерактивной работе с Matrix-компонентом позиционирования и масштаба отображения Геокарты, выделение цветом выбранных слоев (Shapes), соответствующих активным в данный момент элементам матрицы (Matrix). Данная задача в настоящий момент решена в продукте SAS Enterprise BI Server [10], а также в продукте SpotOn Vantage Maps [11]. Кроме того, данная задача может быть решена с использованием платформы и технологий Microsoft SQL Server 2012 SSAS и ESRI ArcMaps.

Задача 2: Создание многопользовательской системы комплексного визуального анализа геоданных с функциями прокрутки от целого к частному и обратно (Drill-up и Drill-down) и визуальным отображением результатов многомерных (MDX) запросов к ГеоГиперкубу на ГеоКарте и (или) в сводных таблицах (Pivot Table) (в Matrix). В настоящее время существует решение данной задачи на базе некоммерческой SOLAP-системы, которая является первой попыткой коллектива энтузиастов создать полнофункциональную OLAP систему для работы с Геоданными. Geomondrian SOLAP Server ориентирован на комплексный анализ Геоданных в режиме реального времени. В качестве набора мер ГеоГиперкуба системы GeoMondrian могут быть Геоданные (Geometry и Geography). Для отображения срезов и результатов MDX-запросов используется система PostGIS. Аналогичные решения также могут быть построены на платформе Microsoft SQL Server 2012 и ESRI ArcMaps.

Задача 3: Вычленение из массива данных, входящих в аналитические кубы, мастер-данных (согласно методике использования аналитических кубов часть измерений (часто — это классификаторы) трактуются как мастер-данные¹), управление мастер-данными комплексной системы визуального анализа геоданных на основе «бизнес-правил» и строгих процессов согласования изменений». Данная задача еще не решалась для SOLAP-систем, но может быть решена с помощью технологии Microsoft SQL 2012 server MDS, появившейся в 2012 г. и позиционируемой как бесплатный компонент к Microsoft SQL 2012 Server.

6. Программные средства для решения SOLAP-задач

6.1. Microsoft SQL 2012 Server Analysis Services (SSAS) [12]

Службы Microsoft SQL 2012 Server Analysis Services позволяют проектировать, создавать и управлять многомерными структурами, которые содержат де-

тализирующие и статистические данные из нескольких источников данных, таких как реляционные базы данных, в одной унифицированной логической модели, поддерживаемой встроенными средствами проведения вычислений.

Службы Analysis Services обеспечивают возможность быстрого, доступного для понимания пользователя, нисходящего анализа крупных количеств данных, основанного на этой унифицированной модели данных, результаты которого могут доставляться пользователям на нескольких национальных языках, с пересчетом на несколько разных валют, с представлением дат в национальном формате. Они применяются для работы с хранилищами данных (DataWarehouse), витринами данных (DataMart), производственными базами данных OLTP и хранилищами оперативных данных, поддерживая анализ данных с предысторией и данных, поступающих в реальном времени. Они также позволяют анализировать большие объемы данных. С их помощью можно проектировать, создавать и управлять многомерными структурами, которые содержат подробные и статистические данные из нескольких источников данных.

Для управления кубами OLAP и работы с ними используется среда SQL Server Management Studio. Для создания новых кубов OLAP и моделей интеллектуального анализа данных в службах SQL Server Analysis Services используется среда Business Intelligence Development Studio (BIDS), представляющая собой среду Microsoft Visual Studio 2008 с расширениями, специфичными для решений бизнес-аналитики. Данный инструмент предоставляет ряд уникальных функций, позволяющих работать с проектами служб Analysis Services и интегрировать проекты служб Analysis Services со службами Reporting Services и Integration Services. К наиболее важным элементам BIDS следует отнести наличие различных конструкторов:

◆ *конструктор представлений источников данных* обеспечивает возможность добавления или удале-

¹ Мастер-данные (Master Data) или НСИ — совокупность постоянной и условно-постоянной информации, необходимой в ИС автоматизации бизнес-процесса. Бизнес-процессом, в рамках которого определена данная НСИ, может быть любой процесс обработки пространственно-распределенной социально-экономической информации. Одна и та же информация может являться или не являться НСИ в зависимости от контекста ее использования. В самом процессе создания НСИ информация, являющаяся нормативно-справочной (например, классификатор ОКАТО т др.) для других процессов, собственно ни нормативной ни справочной не является. Система НСИ обеспечивает хранение, обработку и предоставление постоянной и условно-постоянной информации пользователям системы и предназначена для поддержания (больших) массивов данных в актуальном состоянии, обеспечению полноты, устранению ошибок, контролю целостности и непротиворечивости данных. Модификация хранимых в системе НСИ данных и их структуры допускается только экспертами системы. Все действия по модификации данных строго регламентируются. Пользователями информации выступают прочие ИС предприятия, получающие данные через заранее определенные интерфейсы. Исторически, системы НСИ появились в корпоративных хранилищах данных (DataWarehouse), являющихся источниками данных для OLAP систем, для поддержания корректности корпоративных справочников и классификаторов.

ния объектов, назначения логических первичных столбцов, определения отношений между таблицами, замены таблиц именованными запросами или другими таблицами и добавления именованных вычислений к существующим таблицам в представлении источника данных.

◆ *конструктор кубов* предоставляет среду для настройки куба и содержащихся в нем объектов. Для приложений на других языках можно добавлять переводы для объектов Analysis Services. При обработке кубов можно просматривать их структуру и данные.

◆ *конструктор измерений* предоставляет среду для настройки измерения и содержащихся в нем объектов. Для локализуемых приложений можно добавлять переводы для метаданных измерения. При обработке измерений можно просматривать их структуру и данные.

В SQL Server 2012 предусмотрены новые возможности бизнес-аналитики, обеспечивающие гибкую, но мощную аналитику в режиме самообслуживания, которая дополняет традиционный OLAP и решения интеллектуального анализа данных. В состав средств самостоятельно проводимой бизнес-аналитики (BI) входят следующие продукты:

◆ *PowerPivot для Sharepoint*. Новая версия служб Analysis Services, которая может храниться на ферме SharePoint. Экземпляр, который хранится в Sharepoint, является глубокой модификацией служб Analysis Services, в которой реализован новый метод хранения в памяти, позволяющий загружать данные по запросу и управлять использованием и потреблением ресурсов без настройки. Хранимый экземпляр имеет также веб-службу среднего уровня, управляющую обновлением данных, доступом к данным и мониторингом.

◆ *PowerPivot для Excel*. Клиентская надстройка для Excel 2010, которая может устанавливаться вместе с Excel 2010. Эта надстройка предоставляет средства для построения многомерных наборов данных в Excel. Надстройка сочетает в себе хранилище сжатых данных, снабженное хранимым в памяти экземпляром служб Analysis Services и существующими средствами визуализации данных Excel, например сводные таблицы и сводные диаграммы.

◆ *DAX*. В состав PowerPivot для Excel входит новый язык выражений (выражения анализа данных, DAX), упрощающий построение сложных расчетов, использование логики операций со временем и выполнение поиска.

6.2. Microsoft SQL 2012 Server Master Data Services (MDS) [12]

Службы SQL Server Master Data Services являются источником основных данных организации. Интеграция разрозненных операционных и аналитических систем со службами Master Data Services (MDS) обеспечивает централизованный, точный источник данных для всех приложений организации. С помощью служб MDS создается единый источник основных данных, а также поддерживается доступная для аудита запись изменений данных со временем. MDS поддерживает гибкие настраиваемые иерархии, которые можно использовать для группировки и обобщения основных данных. При изменении потребностей бизнеса можно соответствующим образом обновить эти иерархии, меняя структуру формируемых отчетов и включая новые аспекты бизнеса. Эти изменения не приводят к потере данных или появлению дубликатов, что может произойти, если данные управляются несколькими системами, не сообщающимися друг с другом. Бизнес-правила позволяют обеспечить точность данных и уведомлять пользователей о ситуациях, когда данные не удовлетворяют правилам.

MDS включает следующие компоненты и инструментальные средства:

- *диспетчер конфигурации служб MDS* – средство, используемое для создания и настройки баз данных MDS и веб-приложений.

- *диспетчер основных данных* – веб-приложение, используемое для управления основными данными организации и выполнения других административных задач и позволяющее постоянно поддерживать точное представление об основных данных компании. (Любые изменения могут быть точно отражены во множестве иерархий, каждую из которых несложно обновить. Кроме того, имеется возможность сохранять версии данных и их структуру, позволяя пользователям работать с самой свежей информацией.). Диспетчер обеспечивает пользователя единым, точным источником для управления основными данными компании; гранулярным, безопасным доступом к основным данным за счет интеграции с Active Directory; гибким моделированием данных, позволяющим определить модель, точно отражающую структуру организации и гибко модифицируемую при изменении потребностей бизнеса; бизнес-правилами, обеспечивающими высокое качество данных; уведомлениями по электронной почте, оповещающими выбранных пользователей или группы об обнаружении несоответствия данных бизнес-правилам; поддержкой версий данных для соз-

дания журнальных моментальных снимков данных и модели; управлением иерархиями с поддержкой как уровневых иерархий, так и менее совершенных иерархий «родители-потомки»; созданием пакетов, применяемых для переноса моделей, бизнес-правил и данных из тестовых систем в производственные.

- *веб-служба MDS*, с помощью которой разработчик может расширять или создавать пользовательские решения MDS для конкретной среды.

- *база данных MDS*, содержащая все сведения, необходимые для системы MDS. Она является основой для развертывания MDS, в ней хранятся параметры, объекты базы данных и данные, необходимые системе MDS, она содержит промежуточные таблицы, используемые для промежуточного хранения данных из исходных систем, и предоставляет схему и объекты базы данных для хранения основных данных из исходных систем. База данных MDS поддерживает управление версиями, в том числе проверку бизнес-правил и уведомления по электронной почте. База данных MDS предоставляет представления для систем-подписчиков, которым надо извлекать данные из базы данных.

На верхнем уровне MDS обладает следующими функциями:

- ◆ Управление сущностями в стиле «один ответственный пользователь за один справочник или классификатор». Ответственный пользователь может добавлять, редактировать или удалять записи НСИ.

- ◆ Построение сложных моделей данных справочников и классификаторов НСИ с помощью «фирменного» визуального интерфейса MDS.

- ◆ Управление иерархиями трех возможных типов: сбалансированные иерархии на базе нескольких сущностей, произвольные конечные иерархии на основе одной сущности, рекурсивные иерархии – возможно загружать иерархии либо строить их в пользовательском интерфейсе.

- ◆ Управление версиями на уровне моделей данных. Сущности, иерархии и наборы группируются в «модели». Различные версии моделей независимы друг от друга.

- ◆ Бизнес-правила и «рабочие процессы согласования изменений НСИ» становятся доступными при сопряжении MDS с SharePoint Server 2010. При наступлении определенных условий автоматически запускается «рабочий процесс согласования изменений». Рабочая версия «модели данных НСИ» обновляется только после согласования «ответственными пользователями».

6.3. ESRI ArcGIS Server [13]

ArcGIS Server предназначен для совместного использования географической информации неограниченным числом пользователей. ArcGIS Server используется на небольших, средних и крупных предприятиях для того, чтобы предоставлять географические информационные ресурсы в виде сервисов по интранет/интернет сетям, оптимизировать внутренние рабочие процессы, разрешать производственные проблемы, координировать деятельность различных служб.

ArcGIS Server предоставляет удобную платформу для создания корпоративной геоинформационной системы, позволяющей:

- ◆ Управлять всеми пространственными данными и картографическими службами централизованно.

- ◆ Увеличить производительность существующих картографических web-приложений и создать новые web-приложения, которые представляют собой не просто карты с изменяемым масштабом и получением информации об объектах на карте, но и дают возможность обработки и обновления представляемой информации.

- ◆ Создать web-приложения, обладающие функциональностью настольных ГИС ArcGIS Desktop.

Внедрить геоинформационную систему в существующую информационную структуру предприятия, объединяя ГИС сервер и пространственные данные с другими информационными системами предприятия, среди которых системы управления отношений с клиентами (CRM) или системы планирования и управления ресурсами предприятия (ERP).

- ◆ Быстро решать специализированные задачи, создавая приложения, объединяющие географическое содержание с функциональными возможностями ГИС.

- ◆ Создать корпоративный геопортал и сформировать инфраструктуру пространственных данных.

ArcGIS Server используется, как «коробочное» решение, не требующее специального программирования и серьезной настройки. ArcGIS Server поставляется с рядом готовых к использованию клиентских приложений и служб (программы-просмотрщики, основанные на web-браузере редакторы баз геоданных и т.д.) для картографирования, анализа, сбора, редактирования, распространения и администрирования пространственной информации. Для решения узкоспециализированных задач существуют развитые и хорошо документированные средства разработки

(платформы .NET и JAVA) с помощью которых можно создавать пользовательские приложения.

В рамках корпоративных ГИС, созданных на базе ArcGIS Server, обеспечивается следующая функциональность:

◆ *на уровне ГИС-сервер администратора(ов)* – управление и конфигурирование серверных объектов, используемых настольными и веб-приложениями, разработанными с использованием технологии ArcGIS Server. Данные функции администратор выполняет через свою локальную сеть с помощью ArcCatalog, настольного приложения ArcGIS. Администратор может управлять выходными каталогами сервера, просматривать статистические и выходные лог-файлы для поиска неисправностей в случае возникновения ошибок. Кроме того, администратор должен использовать инструменты операционной системы для контроля доступа пользователей к ГИС-серверу.

◆ *на уровне разработчики приложений и веб-служб* – создание и тиражирование .Net и Java веб-приложений, веб-служб и настольных приложений, включающих дополнительную ГИС-функциональность, используя при соединении с ГИС-сервером работающие на нем ArcObjects. Компонент ESRI ArcGIS Mapping for SharePoint позволяет работать с сервером ArcGIS сервер с помощью открытой в браузере страницы портала Microsoft Sharepoint 2010.

◆ *на уровне пользователи ArcGIS Desktop* – использование ArcMap и ArcCatalog для соединения

с ГИС-сервером через локальную сеть или через Internet. В обоих случаях пользователи могут пользоваться картами и службами геокодирования, опубликованными как объекты map server objects и geocode server objects.

◆ *на уровне пользователи веб-приложений* – использование соответствующей ГИС-функциональности через интернет-браузер, необходимый для связи с веб-приложениями, опубликованными разработчиком.

7. Заключение

Таким образом, в данной статье были описаны современные методы представления и анализа социально-экономических данных с использованием «пространственных» (spatial) OLAP-систем (SOLAP). Показано, что хотя в последние годы данное направление информационных технологий активно развивается, в настоящее время мало технологических решений, позволяющих решать задачи SOLAP-анализа. Для комплексного анализа в режиме «реального времени» социо-экономических данных, привязанных к географическим объектам, в работе предложено современное инновационное, недорогое в своем классе решение на основе широко распространённых, проверенных программных продуктов крупных изготовителей программного обеспечения, причем в настоящее время в академическом и бизнес-сообществе подобных решений пока не отмечено. ■

Литература

1. Фёдоров А., Елманова Н. Введение в OLAP-технологии Майкрософт. – М.: Диалог-МИФИ, 2002. – 272 с.
2. Books Online for SQL Server 2012 – MSDN – Microsoft. – <http://msdn.microsoft.com/en-us/library/ms130214.aspx> (режим доступа: декабрь 2012).
3. Luhn H.P. A Business Intelligence System. // IBM Journal, Vol 2., 1958, pp. 314-319.
4. Колесов А. На смену Business Intelligence приходит Business Analytics? // PC Week/RE, № 41 (599), 6 ноября – 12 ноября, 2007. – <http://www.pcweek.ru/idea/article/detail.php?ID=103694>.
5. PC Week Review: бизнес-аналитика, октябрь 2009. – <http://www.pcweek.ru/business/article/detail.php?ID=120748>.
6. Winckler M. Apache Hadoop takes top prize at Media Guardian Innovation Awards. // The Guardian, 25 March, 2011. – <http://www.guardian.co.uk/technology/2011/mar/25/media-guardian-innovation-awards-apache-hadoop>.
7. http://www.dataplus.ru/Industries/100_GIS/GIS.htm.
8. Bédard Y., T. Merrett & J. Han, Fundamentals of Spatial Data Warehousing for Geographic Knowledge Discovery / Geographic Data Mining and Knowledge Discovery, Taylor & Francis, Vol. Research Monographs in GIS, 2011, No. Chap. 3, pp. 53-73.
9. <http://www.spatialytics.org/projects/geomondrian>.
10. http://www.sas.com/technologies/bi/touext/olapviewer_itour_flash.html.
11. <http://www.spotonsystems.com/products/spoton-vantage-maps>.
12. Books Online for SQL Server 2012 – MSDN – Microsoft. – <http://msdn.microsoft.com/en-us/library/ms130214.aspx> (режим доступа: декабрь 2012).
13. <http://www.dataplus.ru/Soft/ESRI/ArcGIS/ArcGISServer/Index.html>.

◆
**THE ROLE OF THE SUBJECTIVITY
IN BUSINESS PROCESSES**

V. Chebotarev, A. Gromov

Annotation

The article is devoted to the research of the subjectivity, the subjective structure and the reflectiveness in the business-process management of the company. This research was made with the financial support of the government of the Russian Federation (Ministry of education and science of Russia) in frames of the contract №13.G25.31.0096 «The creation of high technological production of cross-platform systems for processing the non-structured information on the base of freeware for increasing the efficiency of company innovation activity management in modern Russia».

Key words: subjective structure, business process, systems thinking, interactive modeling, business architecture.

◆
**MATHEMATICAL BASES
OF THE PRODUCTION-CONSUMPTION THEORY:
DEFINITION OF A TYPE, STRUCTURE
AND PARAMETERS OF MODELS**

M. Ivlev

Annotation

Within the production-consumption theory of goods in open social-economic systems the mathematical models describing the goods conformity to inquiries of the consumer are offered. Basis of models is the finite planar focused graph with flows. The way of the minimum splitting of the graph on sets of independent arches covering it is founded, rules of tops are formulated. On the basis of the analysis of components of the focused graph the criterion functions and necessary restrictions for the solution of a number of optimizing tasks are defined.

Key words: theory of production-consumption, problem of optimization, finite planar focused graph with flows, criterion functions and restrictions.

◆
**SIMULATION OF MASS EVENTS
(ILLUSTRATED BY THE XXII WINTER OLYMPIC GAMES 2014)**

V. Abrosimov, V. Lebidko

Annotation

The basic principles of modeling the organization of events using a multi-agent technology developed. Identified characteristics of events, shows the practical importance of modeling. Visitors events and support staff are the agents that exhibit in meeting the goals of the collective property of the behavior based on their own autonomy. On the model of access to the sports events at the Olympic Stadium in Sochi the possibilities simulate the effects of different critical situations.

Key words: simulation, modeling, mass action, multi-agent approach, Olympic Games, Sochi, queue, crowd.

◆

APPLYING METHODS OF CONSTRAINTS SATISFACTION PROBLEMS SUITABLE FOR SOLVING THE SCHEDULING PERIODIC MAINTENANCE PROCESSES

A. Fedotova, M. Ovsyannikov, V. Taratukhin

Annotation

The article describes the methods of planning maintenance analysis, problem statement, maintenance, formation constraints, methods of constraints satisfaction problems suitable for solving the scheduling periodic maintenance processes. Methods are chosen and the technique of the problem decision, based on a combination of search methods and preliminary restriction of search space is formulated. Algorithms of methods realization are offered.

Key words: maintenance and repair of equipment, methods of the scheduling periodic maintenance processes, methods of constraints satisfaction problems, periodic maintenance processes, combinatorial problem, methods to find solutions, backtracking, axioms, domain ontology.

◆

ON SOLUTION OF SMALL SAMPLE SIZE PROBLEM WITH LINEAR DISCRIMINANT ANALYSIS IN FACE RECOGNITION

V. Mokeyev, S. Tomilov

Annotation

Effectiveness of linear discriminant analysis of face recognition in case of small sample size is described. For computing of discriminant components is offered to use the generalized method of Jacobi that allows not to lose discriminant information. The efficiency of proposed approach is demonstrated by experimental researches with ORL database.

Key words: face recognition, linear discriminant analysis, small sample size, database ORL.

◆

ALGORITHMS OF FORMING SPECTRAL REPRESENTATIONS OF SOUND SIGNAL BASED ON U-TRANSFORM

V. Gai

Annotation

This article discusses algorithms of multilevel (roughly-accurate) signal representation using filters Walsh system Hartmut. Given the characteristics of the implementation of computer algorithms. Advantages of the proposed algorithms build decomposition: low computational complexity and simplicity of implementation.

Key words: digital signal processing, Walsh transform, U-transform.

◆

RELIABLE PLAGIARISM DETECTION SERVICE WITH LIMITED NUMBER OF SEARCH QUERIES

V. Dyagilev, A. Tskhay, S. Butakov

Annotation

The paper describes improved architecture for plagiarism detection service. It also shows experimental results which support the advantages of the proposed approach. The proposed novel architecture assumes moving Internet search out of core system to the infrastructure of plagiarism detection service provider. In contrast to traditional outsourcing this approach guarantees that provider never gets the original document. In this case the proposed architecture never violates copyrights protection requirements for the documents submitted for the checkup. The experiment shows applicability of the proposed approach on the documents with different level of plagiarism.

Key words: plagiarism detection, service architectures.

◆
**INFORMATION SYSTEMS
FOR CORPORATE PLANNING
AND BUDGETING**

D. Isaev

Annotation

In the paper integrated information systems for corporate planning and budgeting are considered. Four groups of practical tasks exceeding the bounds of typical functionality of special-purpose planning and budgeting information systems are allocated. Several classes of information systems (simulation, statistical analysis, financial analysis and modeling, group decision making, business intelligence), which may provide the completeness of corporate planning and budgeting are denoted as solutions complementary to special-purpose planning and budgeting systems.

Key words: planning, budgeting, strategic management, performance management system, information system, analytical application, simulation, statistical analysis, financial analysis and modeling, decision making, business intelligence systems.

◆
**MEASUREMENT OF MATURITY LEVEL
OF A PROFESSIONAL COMMUNITY**

F. Krasnov, R. Yavorskiy

Annotation

In this paper we consider formal model of an online community as a combination of social and semantic networks. An approach is suggested to measurement of maturity level of a professional community based on two groups of parameters, which characterize competency level and density of contacts correspondingly. We present preliminary results of the pilot testing of the proposed approach for evaluating several city IT-communities in Central Russia. The paper is based on the authors' experience and achievements obtained in their work to develop Virtual Skolkovo online community.

Key words: professional community, online community, social network, semantic network, parameters of online communities, rating of communities.

◆
**APPLICATION OF ADVANCED
INFORMATION TECHNOLOGIES
FOR OPERATIONAL ANALYSIS
OF SOCIO-ECONOMIC DATA**

K. Latyshev, V. Sidorenko

Annotation

The paper is devoted to modern methods of presentation and analysis of socio-economic data by means of Spatial-OLAP-systems. The modern innovative cost-effective solution, based on well-known software of leading software vendors, is proposed for complex analysis of «real time» socio-economic geo-referenced data. There are no analogous solutions in the academic and business communities at the present moment.

Key words: Spatial-OLAP-system, analysis of socio-economic data.

**ЖУРНАЛ «БИЗНЕС-ИНФОРМАТИКА»
ОСУЩЕСТВЛЯЕТ РАЗМЕЩЕНИЕ РЕКЛАМНЫХ
И РЕКЛАМНО-ИНФОРМАЦИОННЫХ МАТЕРИАЛОВ**

Расценки:

Обложка: 2, 3, 4 страница обложки, полноцветная печать, полоса 210×290 мм (А4) – 40 тыс. руб.

Текстовый блок, чёрно-белая печать:

- ◆ полоса – 20 тыс. руб.;
- ◆ 1/2 полосы – 15 тыс. руб.;
- ◆ 1/4 полосы – 10 тыс. руб.;
- ◆ меньший объём – 7 тыс. руб.

Вставка (4 полосы, полноцветная печать – 60 тыс. руб.).

Рекламно-информационный блок (8 полос, полноцветная печать) – 80 тыс. руб.

Рекламно-информационный блок (16 полос, полноцветная печать) – 90 тыс. руб.

Корпоративный специальный выпуск – по договоренности.

Материалы принимаются с учётом следующих параметров:

- ◆ дообрезной формат – 215×300 мм;
- ◆ обрезной формат – 210×290 мм;
- ◆ поле набора полосной рекламы – 190×270 мм – с отступом от границ обрезного формата по 10 мм с каждой стороны;
- ◆ файл TIF, EPS, PDF – разрешение не менее 300 dpi.

Тематические рубрики журнала
«БИЗНЕС-ИНФОРМАТИКА»

№	Рубрика
1	Математические модели социальных и экономических систем
2	Программная инженерия
3	Анализ данных и интеллектуальные системы
4	Математические методы и алгоритмы решения задач бизнес-информатики
5	Моделирование и анализ бизнес-процессов
6	Информационные системы и технологии в бизнесе
7	Электронный бизнес
8	Интернет-технологии
Дополнительные рубрики вне номенклатуры	
9	Тематические обзоры
10	Правовые вопросы бизнес-информатики
11	Стандартизация, сертификация, качество, инновации
12	Дискуссионный клуб / Опыт бизнеса

К СВЕДЕНИЮ АВТОРОВ

Редакция просит авторов при оформлении статей и тематических обзоров придерживаться следующих правил и рекомендаций:

1. Предоставляемый авторами материал должен соответствовать рекомендуемой структуре статей журнала.

2. Статья направляется в редакцию в электронном виде (в формате MS WORD версия 2003) и в виде бумажной копии, распечатанной на одной стороне листов А4. Первая страница оригинала подписывается всеми авторами статьи.

3. Ориентировочный объём статьи, предлагаемой к публикации, – 20–25 тыс. знаков (с пробелами) или 30–35 тыс. знаков – для обзорных статей по направлениям.

4. Кегль набора – 12 пунктов с полуторным интервалом. Нумерация страниц – сверху по центру. Поля: левое – 2,5 см, верхнее, нижнее и правое – по 1,5 см.

5. При наборе выключных и строчных формул должен быть использован редактор формул MS Equation. В формульных и символических записях греческие (русские) символы, а также математические функции записываются прямыми шрифтами, переменные аргументы функций в виде английских (латинских) букв записываются наклонным начертанием (курсивом), например, « $\cos a$ », « $\sin b$ », « \min », « \max ».

6. Формулы, таблицы и сноски (не концевые) оформляются стандартными средствами редактора MS WORD. Нумерация формул, рисунков и таблиц – сквозная, по желанию авторов допускается двойная нумерация формул с указанием структурного номера раздела статьи и – через точку – номера формулы в разделе.

7. Рисунки (графики, диаграммы и т.п.) оформляются средствами Word, Excel, Illustrator. Ссылки на рисунки в тексте обязательны и должны предшествовать позиции размещения рисунка. Допускается использование графического векторного файла в формате wmf/emf или cdr v.10. Фотографические материалы предоставляются в формате TIF или JPEG с разрешением не менее 300 dpi.

8. Библиографический список составляется в соответствии с требованиями ГОСТ. Нумерация библиографических источников – в порядке цитирования. Ссылки на иностранную литературу – на языке оригинала без сокращений.

Структура статей строится по правилам, рекомендованным журналом «Бизнес-информатика».

**Плата с аспирантов
за публикацию рукописей не взимается.**

Журнал «БИЗНЕС-ИНФОРМАТИКА»
Рекомендуемая структура статей

Журнал публикует исследовательские научные статьи, размещаемые в рубриках журнала, тематические обзоры, отражающие современное состояние проблем в области бизнес-информатики и сообщения, размещаемые в рубриках «Дискуссионный клуб» и «Опыт бизнеса».

Титульный лист рукописи начинается с указания Ф.И.О. авторов публикации с обязательным указанием учёной степени, учёного звания, должности, основного места работы и e-mail. Титульный лист должен быть подписан всеми авторами статьи.

**I. Исследовательские научные статьи
(для размещения в тематических рубриках)**

Редколлегия рекомендует авторам после названия статьи приводить **аннотацию**, в которой излагается краткое содержание статьи, её основные результаты и область применения. Авторам рекомендуется структурировать статью, выделяя **введение**, содержащее описание проблемы или задачи, обзор существующих подходов или методов решения, их недостатки, и основную цель статьи; **постановку задачи**, включающую допущения и ограничения; **содержательную часть** статьи, в которой предлагаемые решения должны быть аргументированы и сравниваться с существующими подходами или решениями; **заключение**, содержащее краткое изложение новых результатов, полученных в статье и область их применения; **библиографический список**, оформленный в соответствии с ГОСТ. Текст статьи должен содержать нумерованные ссылки на все указанные библиографические источники. Структурирование статьи и нумерация её разделов проводится по усмотрению авторов.

Возможный вариант структуры статьи:

- ◆ Ф.И.О.;
- ◆ учёная степень, учёное звание, должность, основное место работы, e-mail;
- ◆ название статьи.
- ◆ аннотация;
- ◆ 1. Введение.
- ◆ 2. Постановка задачи.
- ◆ 3. Основная содержательная часть статьи.
- ◆ 4. Экспериментальные результаты (опционально).
- ◆ 5. Заключение.
- ◆ 6. Библиографический список.

II. Тематические обзоры по направлениям

Редколлегия рекомендует авторам структурировать обзор, выделяя аннотацию, содержащую тематику, краткое содержание обзора и область применения; **введение**, в котором даётся краткий исторический обзор тематики; **содержательную часть** обзора с критическим анализом существующих направлений; **заключение**, в котором отражаются перспективы развития в рамках обзора тематики и наиболее интересные направления с точки зрения научных и практических разработок и методов; **библиографический список**, оформленный в соответствии с ГОСТ.

Текст обзора должен содержать нумерованные ссылки на все указанные библиографические источники. Структурирование обзора и нумерация его разделов проводится по усмотрению авторов.

Возможный вариант структуры обзора:

- ◆ Ф.И.О.;
- ◆ учёная степень, учёное звание, должность, основное место работы, e-mail;
- ◆ название обзора;
- ◆ аннотация;
- ◆ 1. Введение.
- ◆ 2. Основная содержательная часть обзора.
- ◆ 3. Заключение.
- ◆ 4. Библиографический список.

Редколлегия журнала проводит обязательное рецензирование рукописей. Статья принимается к публикации только после получения положительного заключения рецензента и одобрения на заседании редакционной коллегии журнала.