

ISSN 1998-0663

№3(25)—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:

**НОВЫЙ АНАЛИЗ
ПАТТЕРНОВ**

**КРОССЛЕКСИКА
КАК УНИВЕРСУМ**

**ИММОД:
ЭВОЛЮЦИОННЫЕ СЕТИ**

**МНОГО-
КРИТЕРИАЛЬНЫЙ
ПОДХОД**

В соответствии с решением президиума ВАК РФ журнал «Бизнес-информатика» с 19.02.2010 включен 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

№3(25)–2013

СОДЕРЖАНИЕ

Анализ данных и интеллектуальные системы

*Ф.Т. Алескеров, В.Ю. Белоусова, Л.Г. Егорова,
Б.Г. Миркин*

Анализ паттернов в статике и динамике, часть 1:
Обзор литературы и уточнение понятия 3

И.А. Большаков

Кросслексика: универсум связей
между русскими словами 19

Ю.В. Таратухина, Д.А. Алдунин

Использование принципов эргономической семиотики
при проектировании пользовательских интерфейсов
в поликультурном контексте 27

Математические методы и алгоритмы решения задач бизнес-информатики

М.А. Хивинцев, А.С. Акопов

Распределенная эволюционная сеть для решения
многокритериальных оптимизационных задач
в системах имитационного моделирования 34

В.В. Подиновский, М.А. Потапов

Метод взвешенной суммы критериев в анализе
многокритериальных решений: Pro et contra 41

В.Н. Никулин, С.А. Палешева, Д.С. Зубарева

Метод эмпирических вероятностей: автоматическая
система для рекомендации следующих десяти лекций
курса после просмотра трех данных лекций 49

Информационные системы и технологии в бизнесе

О.В. Ена, К.В. Нагаев

Автоматизация процессов разработки технологических
дорожных карт. Расчет интегральных показателей
применимости 56

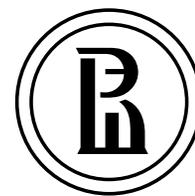
Т.К. Кравченко

Управление требованиями при реализации
ИТ-проектов 63

Б.Б. Славин, И.У. Ямалов

Создание инфраструктуры SMART-региона
на основе развития информационных технологий
и электронного образования 72

Annotations 79



БИЗНЕС-ИНФОРМАТИКА

№3(25)–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 верстка
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 приведенных 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»

АНАЛИЗ ПАТТЕРНОВ В СТАТИКЕ И ДИНАМИКЕ, ЧАСТЬ 1: ОБЗОР ЛИТЕРАТУРЫ И УТОЧНЕНИЕ ПОНЯТИЯ¹

Ф.Т. Алескеров,

доктор технических наук, руководитель Департамента математики факультета экономики Национального исследовательского университета «Высшая школа экономики», заведующий лабораторией Института проблем управления имени В.А. Трапезникова РАН

В.Ю. Белоусова,

кандидат экономических наук, заведующий отделом методологии бюджетного планирования Института статистических исследований и экономики знаний, доцент кафедры банковского дела Департамента финансов факультета экономики, с.н.с. Банковского института Национального исследовательского университета «Высшая школа экономики»

Л.Г. Егорова,

преподаватель Департамента математики факультета экономики Национального исследовательского университета «Высшая школа экономики»

Б.Г. Миркин,

доктор технических наук, профессор кафедры анализа данных и искусственного интеллекта отделения прикладной математики и информатики факультета бизнес-информатики Национального исследовательского университета «Высшая школа экономики»

E-mail: alesk@hse.ru, vbelousova@hse.ru, legorova@hse.ru, bmirkin@hse.ru
Адрес: г. Москва, ул. Мясницкая, 20

Анализ паттернов – это новая область анализа данных, связанная с поиском взаимосвязей исследуемых объектов, построением их классификации и исследованием развития объектов во времени. В первой части статьи вводится понятие «паттерн» и приводится обзор литературы по методам кластерного анализа и анализа паттернов.

Ключевые слова: паттерны данных, динамический анализ паттернов, кластерный анализ.

¹ Работа выполнена при финансовой поддержке Минобрнауки России по государственному контракту от 14.06.2012 г. № 07.514.11.4144 в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 годы».

Исследование осуществлено в рамках программы фундаментальных исследований НИУ ВШЭ в 2012 году. Работа авторов была также поддержана рядом лабораторий НИУ ВШЭ: Лаборатория анализа и выбора решений (Алескеров Ф.Т., Белоусова В.Ю., Егорова Л.Г., Миркин Б.Г.), Лаборатория алгоритмов и технологий анализа сетевых структур (Егорова Л.Г., Миркин Б.Г.). Мы благодарны за эту поддержку.

1. Введение

Понятие «паттерн» широко используется в задачах машинного обучения и обработки данных. Впервые (более 50 лет назад) оно упоминалось как составная часть словосочетания «Pattern Recognition», которое было переведено на русский язык и закрепилось в нем как «распознавание образов». При этом не очень чёткое английское слово «pattern» было вполне адекватно представлено в русскоязычной литературе в качестве столь же нечёткого «образа». В машинном обучении этот термин обозначает группу многомерных объектов, указанных «учителем» как в чём-то сходных. Одновременно с ним распространение получило столь же нечёткое на первый взгляд понятие «кластер», означающее группу объектов, просто похожих в признаковом пространстве без участия «учителя». Однако за прошедшее время слово «паттерн» проникло в русский язык в качестве рабочего термина ряда специальных дисциплин, таких как «технический анализ» динамики цен, психология или инженерия, где этот термин получил несколько отличающееся от первоначального значение и начал означать что-то типа «шаблона». Именно это более узкое значение понятия «паттерн» мы и хотим использовать для обозначения определённого класса образов или кластеров, постоянно возникающих в анализе данных и машинном обучении, но, кажется, до сих пор не закреплённого в специально выделенном термине.

Под «паттерном» в данной работе понимается такая комбинация определённых, с точностью до погрешности, значений некоторого подмножества признаков, что объекты с этими значениями достаточно сильно отличаются от других объектов. Это понятие можно считать эмпирическим аналогом концептуально-логического понятия «тип» в той же мере, в которой понятие кластера является эмпирическим аналогом концептуально-логического понятия «класс». Нас это понятие интересует, прежде всего, с точки зрения анализа динамики объектов — они естественным образом распадаются на группы, придерживающиеся единого паттерна на временном интервале, — так называемые устойчивые группы поведения и менее устойчивые, чередующие паттерны с течением времени.

Способность находить и использовать паттерны в данных — одна из движущих сил современной науки и промышленности. Интернет-магазины, такие как Amazon.com, рекомендуют своим покупателям продукты, основанные на закономерностях, обнаруженных в базе данных прошлых транзакций. Биологи могут обнаружить гены во многом таким

же образом, путем автоматического сравнения последовательности генома со всеми известными последовательностями. Google может получать веб-страницы, которые имеют отношение к запросу с использованием аналогичных идей. Этот список можно продолжить.

Ниже приведен обзор работ, связанных с каждым из трёх упомянутых понятий — паттерн, кластер, анализ динамики многомерных объектов, прежде всего, в их отношении к рассматриваемому понятию. Во второй части будет представлено несколько приложений понятия «паттерн» к анализу социально-экономических явлений.

2. Использование понятия «паттерн» в литературе

Понятие паттерн широко используется в финансах, экономике, техническом анализе, прогнозной аналитике, медицине, криминалистике и некоторых других предметных областях. Что характерно, он определяется в этих областях знания по-разному:

- ◆ как сущность явления, имеющего повторяющиеся черты;
- ◆ как свойство повторяющихся компонентов, объединённых общей структурой;
- ◆ как процесс, фиксирующий модель взаимодействия изучаемых объектов, включающего повторения.

Во всех этих предметных областях паттерны данных могут использоваться в смысле выделения групп схожих объектов и изучения их ключевых характеристик с проведением кластерного анализа для разбиения всех объектов выборки на непересекающиеся кластеры для формирования их классификации. В качестве паттерна обозначается некая выявленная закономерность в данных или некая шаблонная структура данных. Например, в [6] термин «паттерн» употребляется для описания некой устойчивой структуры экономических показателей (см. также [1-3]).

Как упоминалось, в англоязычной литературе по искусственному интеллекту и машинному обучению широко используется термин «Pattern recognition» или «Pattern analysis», который переводится на русский язык как распознавание образов. Распознавание образов — это отнесение исходных данных к определённому, не обязательно заранее заданному, классу с помощью выделения существенных признаков, характеризующих эти данные, из общей массы данных. Задачами теории



Рис. 1. Паттерн технического анализа «треугольник» (рис. взят с обучающего портала по техническому анализу «Биржа и мы», <http://exchangeandwe.ru/>)

распознавания образов являются, например, распознавание лиц, речи и изображений, штрих-кодов и автомобильных номеров, классификация документов и проч. Алгоритмы распознавания образов зависят от конкретной задачи и типа исходных данных и могут включать в себя методы классификации и кластеризации, а также нейронные сети, марковские модели и байесовские сети [41].

В техническом анализе паттерном называются устойчивые повторяющиеся изменения сочетания цены, объёма или индикаторов рынка за определенный промежуток времени. Анализ паттернов здесь основывается на одной из аксиом технического анализа — «история повторяется» — считается, что повторная конфигурация данных в динамике приводит к аналогичному результату. В русскоязычной литературе паттерны иногда называют «шаблонами» или «фигурами» технического анализа. Также этот термин может упоминаться в словосочетании «trading pattern».

В этом случае паттерн иллюстрируется линией графика цены или индекса, соединяющей соседние цены (цены на момент закрытия торгов, их максимальное и минимальное значения) или значения индекса за определенный промежуток времени. Анали-

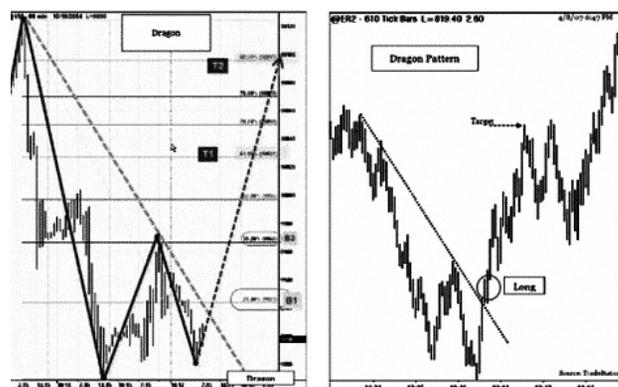


Рис. 2. Паттерн технического анализа «дракон» (рис. взят с обучающего портала по торговле на рынке ценных бумаг ForexTurbo, <http://www.forexturbo.ru/>)

тики рынка пытаются идентифицировать паттерны, чтобы попытаться предугадать ценовые движения рынка в будущем. В техническом анализе выделено много различных видов паттернов: треугольник (рис. 1), клин, флаг, фигура «голова и плечи» и т.д.

Иногда встречаются и совсем экзотичные фигуры, например, дракон (рис. 2).

Прогнозная аналитика использует методы статистики, интеллектуального анализа данных и теории игр для анализа текущих данных и для составления предсказаний о будущих событиях. В бизнесе прогнозные модели используют паттерны, найденные в исторических данных, чтобы идентифицировать риски и возможности. Модели фиксируют связи среди многих факторов, чтобы сделать возможной оценку рисков или потенциала, связанного с конкретным набором условий, при руководстве принятием решений о возможных сделках [36].

Прогнозная аналитика применяется в актуарных расчётах, финансах, страховании, телекоммуникациях, розничной торговле, туризме, здравоохранении, фармацевтике и других областях. Одним из наиболее известных приложений в финансах является кредитный скоринг, модели которого обрабатывают кредитную историю потенциального заемщика, информацию о его текущих займах и финансовом благополучии, потребительские данные и т.д. На основе этих результатов потенциальные заемщики упорядочиваются по вероятности качества обслуживания ими планируемых обязательств по кредиту, включающих выплаты по основному долгу и начисленным процентам за пользование заемными средствами банка в установленные кредитным договором сроки.

Другим прикладным примером является использование паттернов в розничной торговле, когда в результате анализа данных о покупках человека в розничной сети ему могут быть предложены рекламные акции или купоны на скидки по товарам, которые ему вероятнее всего потребуются. Например, если Вы купили купальник в апреле, то в мае Вам предложат крем для загара. Для реализации такого поведенческого «таргетинга» необходимо выявление подобных зависимостей в покупках, т.е. анализ паттернов, называемых в майнинге данных «ассоциациями».

В медицине термин «паттерн» встречается в анализе кардиограмм, энцефалограмм (ЭЭГ) и проч. [9], когда под ним понимают:

- ♦ или последовательность определенных форм колебаний биопотенциалов, повторяющуюся в

одном или нескольких отведениях (то есть парах электродов, с которыми производится регистрация биопотенциалов) при одинаковых состояниях и условиях [27];

- ◆ или же картину, отражающую особенности распределения различных компонентов ЭЭГ по всем отведениям в целом.

Иногда этот термин используется для обозначения последовательности нервных импульсов, имеющей определенное информационное значение [31], например, «паттерны боли при биомеханических нарушениях суставов краниовертебрального перехода и шейного отдела позвоночника» [34] или «паттерны двигательных и чувствительных расстройств при патологии нервных структур в дистальных отделах верхней конечности» [35].

Смысл термина «паттерн» зависит от области знаний, в которой он используется и иногда употребляется для обозначения явлений и процессов, никак не связанных с анализом статистических данных, описанным в разделе 1. К таким областям можно отнести паттерны в программировании, психологии, музыке, педагогике и т.д.

В разработке программного обеспечения шаблон проектирования или паттерн (англ. *design pattern*) представляет собой повторяемую архитектурную конструкцию решения определенной локальной проблемы проектирования в рамках некоторого часто возникающего контекста. Обычно шаблон не является законченным образцом, который может быть прямо преобразован в код; это лишь пример решения задачи, который можно использовать в различных ситуациях. Объектно-ориентированные паттерны показывают отношения и взаимодействия между классами или объектами без определения того, какие конечные классы или объекты приложения будут использоваться [26].

В психологии это слово чаще всего используется в контексте паттернов поведения или гипнотических паттернов [42]. Паттерн поведения — это набор стереотипных реакций или последовательность действий индивида. Каждый человек имеет типовые способы взаимодействия с окружающим миром — устойчивые модели поведения, которые он предпочитает использовать при общении с другими людьми. Кто-то чаще выбирает паттерны уверенного поведения, кто-то — саморазрушающие или манипулятивные паттерны. Чтобы с высокой точностью предсказать, как поведет себя человек в той или иной ситуации, иногда достаточно определить основные паттерны поведения, которых он придерживается в своей

жизни в целом и в подобных ситуациях в частности. Гипнотический паттерн — это текст, который использует гипнолог для наведения транса и последующей работы с индивидом, находящимся в трансе. Некоторые гипнологи могут оказывать подобное влияние без введения в глубокий транс, когда человек искренне считает, что находится в сознании, но его выбор в какой-то мере заведомо определен используемым гипнотическим паттерном.

В трекерной, т.е. создаваемой на компьютере, музыке паттерн — это таблица, определяющая порядок и режим воспроизведения семплов на нескольких каналах за некоторый промежуток времени или группа одновременно воспроизводимых каналов, представляющая полноценную часть музыкальной композиции. Внутри композиции паттерны могут повторяться; это делает возможным относительно быстрое заполнение общей структуры произведения.

В зарубежной литературе по педагогике используют педагогические паттерны (*Pedagogical Patterns*) [16], которые формируются для передачи передового опыта в конкретной области. С их помощью пытаются осуществить передачу экспертных знаний о практике преподавания и обучения. Цель состоит в обучении каждого нового преподавателя стандартному набору педагогических знаний и навыков. Паттерн включает, например, мотивацию студентов, выбор материалов и последовательность их донесения, правила оценки студентов и тому подобное.

Используется этот термин также в криминалистике, в основном, как синоним слов «шаблон» и «структура», например, «паттерн образца ДНК подозреваемого», «паттерн роста опухоли», «паттерн кровоснабжения папиллярной карциномы» или «паттерн сосудистой системы сетчатки», а словосочетание «*rapillary patterns*» означает папиллярные узоры в дактилоскопии.

Как видно, понятие «паттерн» широко используется в самых разных областях науки и практики. Несмотря на кажущуюся специфичность, все приведенные примеры его использования в общем и целом соответствуют тому пониманию, которое было нами сформулировано во введении.

3. Методы кластер-анализа в литературе

В соответствии с [33], методы кластер-анализа разделим по типу формируемых кластерных структур:

1) Методы, выделяющие отдельные кластеры:

- Кластеры по Апресьяну.
- Метод ФОРЭЛЬ.

- Аппроксимационные кластеры.
- Монотонные кластеры.
- Логические таксоны.

2) Методы построения разбиений:

- Метод кластеризации К-средних (K-means).
- Имитирующие природу алгоритмы: генетические, эволюционные алгоритмы и алгоритм роя частиц.

- Методы нечеткой кластеризации.
- Метод аномальных кластеров.
- Модернизированные методы К-средних: использование весов для переменных, вероятностная формулировка и EM-алгоритм, самоорганизующиеся карты Кохонена.

● Методы построения разбиений по матрицам связи, включая так называемые иерархические, спектральные, локальные и графо-теоретические подходы.

3) Методы построения иерархий:

- агломеративные методы,
- дивизимные методы.

4) Бикластерный анализ.

Рассмотрим перечисленные подходы подробнее.

3.1. Методы, выделяющие отдельные кластеры

3.1.1. Кластеры по Апресью [4]

Подмножество объектов выборки S называется А-кластером, если для любых различающихся объектов $x_i, x_j, x_k \in X$ таких, что объекты x_i, x_j принадлежат А-кластеру S , а объект x_k не принадлежит S , расстояние между объектами кластера x_i и x_j меньше расстояния между первым объектом x_i и объектом x_k , не принадлежащим А-кластеру ($d_{ij} \leq d_{ik}$, где $D=(d_{ij})$ – это заданная матрица коэффициентов различия или расстояния между объектами). Нетрудно видеть, что множество всех А-кластеров для данной матрицы D образует иерархию, т.е. если два А-кластера пересекаются, то один из них является частью второго.

3.1.2. Метод Форэль [28]

Этот метод кластеризации основан на идее последовательного объединения элементов в кластер в областях их наибольшего сгущения. Идея метода заключается в следующем: на первом шаге выбирается «центральный» объект g из выборки, далее кластером объявляется множество всех объектов, находящихся от выбранного объекта g на расстоянии меньше заданного заранее расстояния R (то есть все объекты, находящиеся внутри шара радиу-

са R с центром в точке, соответствующей выбранному объекту g), и находится центр тяжести S этого кластера. Найденный центр тяжести объявляется новым центром и кластер переопределяется уже вокруг нового центра. Так продолжается до тех пор, пока центр не стабилизируется. При необходимости можно продолжить процедуру кластеризации, предварительно удалив уже найденный кластер.

На реальных данных этот алгоритм производит несколько крупных кластеров и множество мелких, что в контексте «частичного» кластер-анализа можно отнести к числу преимуществ.

3.1.3. Аппроксимационные кластеры [20]

В этом методе искомым кластер S представляется бинарным вектором принадлежности $s = (s_i)$, где $s_i = 1$, если объект $x_i \in S$, и $s_i = 0$ – в противном случае. Затем вводится положительный уровень интенсивности λ , так чтобы суммарная разница квадратов $L(S) = \sum_{i,j \in X} (b_{ij} - \lambda s_i s_j)^2$ была минимальной, где $B = (b_{ij})$ заданная матрица связи между объектами. Кластер S формируется последовательным присоединением «оптимальных» объектов до тех пор, пока критерий $L(S)$ не перестает уменьшаться.

К преимуществам метода относятся:

- ◆ математически доказанное свойство «тесноты» получаемого кластера S : для всякого объекта из кластера S средняя связь с S превышает $\lambda/2$, а для всякого объекта $x_i \notin S$ средняя связь с S меньше, чем $\lambda/2$;
- ◆ декомпозиция разброса связей (т.е. суммы их квадратов) на части, объясненные и не объясненные кластерной структурой;
- ◆ возможность получения пересекающихся решений.

3.1.4. Монотонные кластеры [29]

Рассмотрим монотонную функцию сходства $f(i, S)$ между всеми объектами $x_i \in X$ и подмножествами $S \subset X$. Монотонность подразумевает, что либо $f(i, S) \geq f(i, S \cup T)$ для всех S и T (сходство между объектом и кластером S больше, чем между объектом и объединенным кластером $S \cup T$), либо $f(i, S) \leq f(i, S \cup T)$ для всех S и T . Например, при заданной неотрицательной матрице связи $A = (a_{ij})$ можно определить $f(i, S)$ как $\min_{j \in S} a_{ij}$ или $\sum_{j \in S} a_{ij}$, так чтобы f была монотонна по S в нужную сторону. Для определенности рассмотрим монотонно возрастающую $f(i, S)$ и определим функцию множеств $F(S) = \min_{i \in S} f(i, S)$, характеризующую «наислабейшее звено» в S . Такие функции $F(S)$ можно максимизировать, подбирая

на каждом шаге наилучшего кандидата, что ведет к «оптимальной» последовательности объектов, которая определяет не только «ядро» – кластер S , максимизирующий $F(S)$, как фрагмент этой последовательности, но и совокупность его «оболочек» – объемлющих фрагментов. Монотонные кластеры использовались для анализа организационных систем.

3.1.5. Логические таксоны [30]

Любой предикат (математическое высказывание, в котором есть хотя бы одна переменная), сформированный из признаков, например, “ $y_1=3$ и $y_2/y_3 > 5$ ”, где y_i – один из признаков, описывающих объекты выборки, определяет множество объектов S , удовлетворяющих ему. Доля f множества S в I – это наблюдаемая частота истинности предиката. С другой стороны, каждая составляющая предиката, связанная с отдельным признаком (в нашем примере – “ $y_1=3$ ” и “ $y_2/y_3 > 5$ ”), имеет свою частоту: f_1 и f_2 . Эти индивидуальные частоты легко скомбинировать так, чтобы определить ожидаемую частоту предиката ef по правилам вероятности для логических операций – например, $ef=f_1 * f_2$. Оптимизационный критерий – это разница между f и ef . Чем больше эта разница, тем лучше. Вероятно, этот подход можно применить и для непосредственного построения паттернов как логических таксонов путем формирования предикатов, имеющих вид конъюнкций значений одних и тех же признаков.

3.2. Методы построения разбиений

3.2.1. Метод кластеризации K-средних (k-means)

Это наиболее популярный метод кластеризации, разбивающий все объекты выборки на заранее заданное число кластеров k , представленных их центральными точками (центроидами) и списками попавших в них объектов. В качестве критерия эффективности выбирается суммарное квадратичное отклонение точек кластеров от центров этих кластеров. Начальные центры кластеров выбираются случайным образом, затем осуществляется последовательность итераций, состоящая из двух шагов. На первом шаге происходит обновление кластера: каждый объект приписывается одному из центров по правилу минимального расстояния, на втором – обновление центра кластера: для каждого кластера вычисляется центр тяжести, который объявляется его новым центром. Все повторяется до тех пор, пока на какой-то итерации центры масс не останутся прежними. Среди недостатков этого мето-

да самыми серьезным является то, что он обычно сходится к локальному, а не глобальному минимуму. Это приводит к зависимости результата от выбора начальных центров кластеров, поэтому существует множество модернизированных методов K-средних; основные рассмотрены ниже.

3.2.2. Имитирующие природу алгоритмы

Имитирующие природу алгоритмы: генетические [17], эволюционные алгоритмы [5], алгоритм роя частиц, муравьиной колонии и пр.

Генетическими и эволюционными называются алгоритмы кластерного анализа, использующие механизмы, напоминающие биологическую эволюцию. Они решают оптимизационные задачи с использованием моделей естественной эволюции, таких как наследование, мутации, отбор и кроссовер. Несколько другой биологический процесс, отыскания пищи роем пчел, имитируется в методе роя частиц. Рой движется в случайном направлении, при этом запоминая наилучшие места из уже посещенных.

Среди недостатков этих методов можно выделить невозможность оценки качества получаемых решений и, главное, плохую масштабируемость для задач с большой размерностью данных, поскольку методы, имитирующие природу, требуют довольно много вычислительного времени.

3.2.3. Методы нечеткой кластеризации [25]

Нечеткий кластер задается функцией принадлежности $z_i = (z_i)$, $x_i \in X$, так что величины принадлежности z_i ($0 \leq z_i \leq 1$) интерпретируются как степени принадлежности объектов кластеру (для четких кластеров значения z_i могут быть только 1 или 0).

Нечеткое разбиение объектов на кластеры – это разбиение по кластерам с центрами $c_k = (c_{k1}, \dots, c_{kv}, \dots, c_{kv})$ в пространстве признаков и K векторов принадлежности $z_k = (z_{1k}, \dots, z_{ik}, \dots, z_{Nk})$, $0 \leq z_{ik} \leq 1$, таких что $\sum_k z_{ik} = 1$ для всех $x_i \in X$.

3.2.4. Метод аномальных кластеров [19]

Этот метод заключается в применении стратегии последовательного исчерпания данных, в данном случае – по одному кластеру. Начальный шаг состоит в том, что точка, наиболее удаленная от центра тяжести всех точек, берется в качестве центра аномального кластера, затем формируется сам кластер, в который входят все точки, расстояние от которых до центра тяжести меньше, чем расстоя-

ние до центра аномального кластера. Далее центр аномального кластера заменяется на центр тяжести кластера и происходит следующая итерация.

3.2.5. Модернизированные методы K -средних

Модернизированные методы K -средних: введение весов для переменных [12], EM-алгоритм [7] и пр.

Модернизированный метод K -средних с введением весов для переменных похож на метод K -средних нечеткого кластер-анализа, только веса здесь отражают не в функции принадлежности объектов, а в признаках в кластере. EM-алгоритм – это метод кластерного анализа, требующий априорного знания модели порождения данных. Согласно модели, наблюдаемые данные – независимая случайная выборка из распределения с функцией плотности вида смеси распределений, параметры которой неизвестны. Каждая итерация алгоритма состоит из двух шагов: на первом (E -шаге) вычисляется ожидаемое значение функции правдоподобия, на втором (M -шаге) вычисляется оценка максимального правдоподобия; таким образом увеличивается величина ожидаемого правдоподобия, вычисляемая на E -шаге. Полученные значения параметров используются для E -шага на следующей итерации. Особенно удобно это осуществлять при нормальных распределениях, моделирующих отдельные кластеры. EM-алгоритм или его модификации используются практически во всех разработках, основанных на модели распределения. К его недостаткам относятся:

1. трудность инициализации,
2. большое число переменных, оценка которых возможна только при большом числе объектов в кластере;
3. сходимости к локальному оптимуму.

3.2.6. Методы построения разбиений по матрицам связи

При заданной матрице сходства $A=(a_{ij})$ между объектами множества X попробуем разбить X на две части – S_1 and S_2 – таким образом, чтобы сходство между S_1 и S_2 было минимально, тогда как сходство внутри – максимально. Методы получения оптимальных разбиений бывают: иерархические, спектральные, локальные и графо-теоретические.

3.2.7. Иерархические алгоритмы [15]

Такие методы работают либо «агломеративно», склеивая на каждом шаге два наиболее близких кла-

стера, начиная с тривиального разбиения на одноэлементные кластеры, либо «дивизивно» – разбивая на каждом шаге какой-либо кластер, начиная с универсального кластера, состоящего из всех объектов.

3.2.8. Спектральный подход [18]

Этот метод решает задачу кластерного анализа в терминах так называемого Лапласиана $L(A)$ как задачу минимизации отношения Рэлея, известного в теории собственных чисел и векторов, при котором кластеры получаются путем огрубления собственных векторов Лапласовой матрицы.

Принцип спектрального кластер-анализа в данной версии: найти собственный вектор, соответствующий минимальному ненулевому собственному значению, после чего определить кластеры в соответствии со знаком компонент: индексы положительных элементов – в один класс, а индексы отрицательных элементов – в другой. Иными словами, этот собственный вектор аппроксимируется вектором из 1 и -1, так что положительные компоненты заменяются на 1, а отрицательные – на -1, после чего определяется разбиение $\{S_1, S_2\}$ следующим образом: S_1 как множество объектов, соответствующих 1, а S_2 – соответствующих -1.

3.2.9. Локальные алгоритмы [20]

Такие методы на каждом шаге рассматривают некоторое разбиение $\{S\}$ и производят локальное его преобразование в сторону улучшения значения критерия. Из всех возможных локальных преобразований чаще всего рассматривается только перенос одного объекта из класса в класс. Для этого отыскивается объект, на котором приращение критерия максимизируется, и если оно положительно – перенос производится. Если нет – разбиение объявляется окончательным результатом.

3.2.10. Графо-теоретический подход [20]

Это алгоритмы кластерного анализа данных, визуализация которых обеспечивается с помощью графов. Объекты представляются как вершины, или узлы графа, а связи между объектами – как дуги или рёбра. Для разных методов виды графов могут различаться направленностью, ограничениями на количество связей и дополнительными данными о вершинах или рёбрах.

Как и большинство визуальных способов представления зависимостей, графы быстро теряют наглядность при увеличении числа объектов.

3.3. Построение иерархий

Здесь существует два вида методов — агломеративные и дивизимные методы.

3.3.1. Агломеративные методы [15,24]

Агломеративными называются методы, в которых вычисления начинаются с тривиальных — одиночных — кластеров, и продолжаются итерациями, каждая из которых состоит в объединении двух ближайших кластеров с последующим определением расстояния между вновь построенным кластером и остальными. В качестве примеров агломеративных методов можно назвать метод ближнего (дальнего) соседа и метод Уорда.

Агломеративные методы обычно используют все попарные расстояния для отыскания минимума, что может существенно осложнить вычисления на больших данных.

3.3.2. Дивизимные методы [20]

Эти методы создают кластеры путем разбиения больших кластеров на меньшие части, начиная со всего множества. Удобство таких методов состоит в том, что процесс деления можно в любой момент остановить. Три наиболее популярных дивизимных метода — бисекция k -средних, бисекция главной компоненты и концептуальный кластер-анализ. Каждый шаг концептуального кластер-анализа состоит в разбиении какого-либо кластера на две части по одному из имеющихся признаков так, что в одну часть включаются объекты, на которых значение этого признака меньше, чем некоторое оптимально подбираемое значение a , а в другую — больше a . Поэтому каждый из финальных кластеров характеризует некоторый паттерн — тот, который определяется конъюнкцией предикатов, ведущих к этому кластеру от корня построенной иерархии.

3.4. Бикластерный анализ [8, 32]

При заданной матрице признаков описаний объектов $Y = (y_{iv})$, где $i \in I$ обозначает объект и $v \in V$ — признак, I — множество строк, V — множество столбцов, бикластером называется пара множеств (S, T) , где $S \subset I$ — подмножество строк, а $T \subset V$ — подмножество столбцов, такая что подматрица $Y(S, T) = (y_{iv})$, где $i \in S$ и $v \in T$, имеет какую-либо особенность, обычно — одинаковые строчки или даже одинаковые значения, свидетельствующие об определенной связи между S и T . Однако иногда рассматривают и более сложные зависимости — например, рост значения, пропорциональный номеру столбца.

Из перечисленных методов кластер-анализа наиболее соответствуют задаче построения паттернов следующие три: (а) построение логического таксона, (б) метод K -средних и (в) концептуальный кластер-анализ. Центроиды кластеров по методу K -средних, как комбинации определённых значений всех признаков, играют роль паттернов. К сожалению, такие решения не всегда удовлетворительны, так как используют все имеющиеся признаки и недостаточно заботятся о том, чтобы паттерны были действительно различны. Последнее определяется требованием, чтобы кластеры образовывали разбиение, т.е. покрывали все объекты, какими бы неадекватными они не были. Напротив, каждый кластер, формируемый методом концептуального кластер-анализа, определяется комбинацией небольшого числа признаков. Однако эти комбинации имеют тот недостаток, что могут включать совершенно различные признаки, тем самым исключая возможность сравнения паттернов друг с другом. В этом плане наибольшие перспективы может иметь метод логического таксона, но применимость этого метода к реальным данным остается неясной, поскольку за 30 лет своего существования, метод никем, кроме авторов [30], не использовался.

4. Исследование многомерных динамических процессов

Существует большое количество исследований по анализу динамики развития социально-экономических объектов. Как правило, такие работы связаны с использованием методов кластерного анализа и распознавания сигналов и изображений. В качестве примеров таких работ, «близких по духу» динамическому анализу паттернов, в данном разделе приводятся краткие описания зарубежных научных теоретических и практических работ, в том числе несколько широкомасштабных проектов Европейской комиссии.

Исследования в области динамического анализа паттернов социально-экономических объектов, проведенные отечественными авторами, рассмотрены во второй части обзора.

В [14] проводится эмпирический анализ закономерностей в эволюции кривых процентных ставок, или кривых доходности — interest rate curves (IRC). Кривые IRC рассматриваются как объекты (кривые) в многомерном пространстве, и ставится проблема изучения сходств и различий между ними. Это типичная проблема кластеризации и классификации в машинном обучении. В качестве примера рассмотрены данные по ежедневной доходности швей-

царского франка (CHF). Информация о динамике доходности была доступна для разных интервалов времени (ежедневная, еженедельная, ежемесячная) с разными сроками. В этом исследовании анализировались данные по дневным ставкам лондонского межбанковского рынка (LIBOR) со сроком 1 неделя, 1, 2, 3, 6, 9 месяцев и ставки по свопам (SWAP) со сроком 1, 2, 3, 4, 5, 7 и 10 лет (рис. 3).

Было получено 4 основных кластера, паттерны которых представлены на рис. 4. Эти паттерны мало различаются – например, паттерн кластера 4 очень похож на паттерны кластеров 1 и 3.

Однако, на самом деле, соответствующие паттерны наблюдаются в разные моменты функционирования системы. Учет этого обстоятельства делает паттерны хорошо различимыми, что показано на рис. 5, где выделены временные периоды действия этих паттернов.

Интересной находкой является выявление нескольких типичных видов поведения кривых, отраженных кластером с низким уровнем ставок, кластером с высоким показателем уровня ставок и переходными между ними кластерами. Информация о динамике процентной ставки используется для экономических и финансовых решений и управления рисками. Такой анализ может помочь в прогнозировании кривых доходности, оценке стоимости финансовых инструментов и управлении финансовыми рисками.

В [23], как и во множестве других подобных работ, задача состоит в выявлении кластеров генов, которые демонстрируют одновременную экспрессию, то есть процесс, в ходе которого наследственная информация от гена преобразуется в функциональный продукт — РНК или белок. Модель, представленная в [23], позволяет учитывать предварительную инфор-

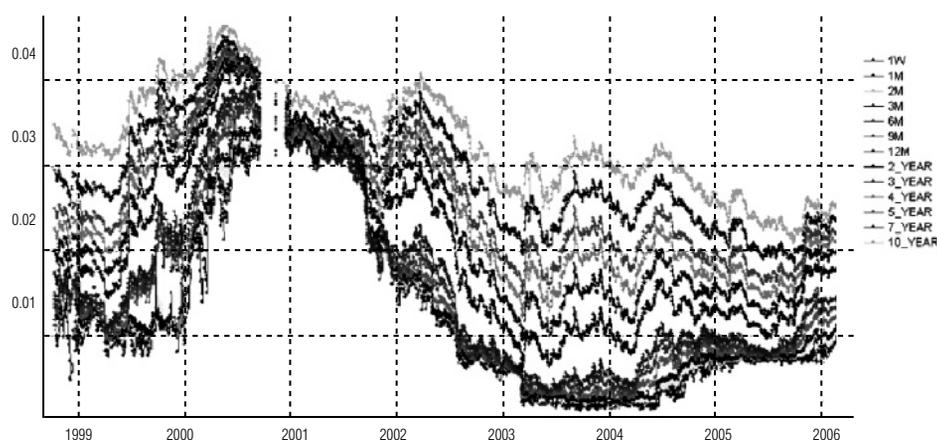


Рис. 3. Эволюция доходности швейцарского франка

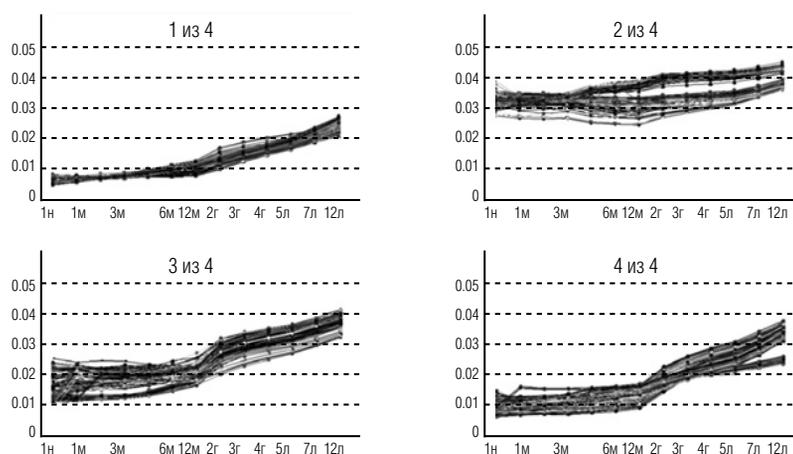


Рис. 4. Примеры IR кривых для каждого из 4 кластеров (на координатных осях – срочность и доходность)

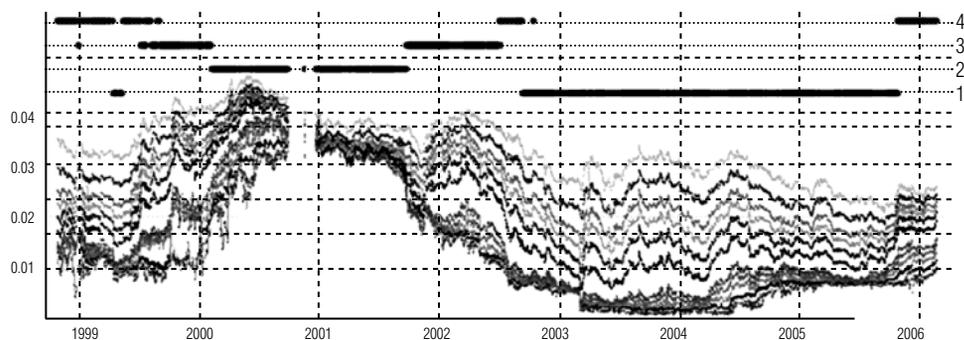


Рис. 5. Результаты классификации, соотношенные с исходными данными. Верхние жирные линии и точки соответствуют найденным 4 кластерам (правая ось)

мацию и предполагает вероятностное распределение по кластерам. Эта модель учитывает то, что каждый ген следует по одной из нескольких основных траекторий, форма и число которых зависит от экспериментальных условий. Был использован начальный фильтр на основе факторного анализа для снижения размерности данных. Результаты иллюстрируются с помощью двух экспериментов, осуществленных на дрожжах. Показатель эффективности данной модели – это апостериорная вероятность совместного поведения генов в процессе экспрессии, которая соответствует вероятности того, что два или более гена попадут в одну группу.

Для результатов первого эксперимента по спорообразованию дрожжей были использованы данные шести временных интервалов. Были оценены уровни совместной экспрессии генов в этом процессе и получены 12 «характеристических кривых» генной экспрессии по всему временному промежутку. На рис. 6 отмечены все кластеры, сплошная темная линия соответствует усредненным значениям, выделенным для описания каждого кластера.

Было обнаружено, что число кластеров, а иногда и вероятность совместной экспрессии объектов, могут быть весьма чувствительны к их априорному распределению.

Существует ряд совместных проектов, выполняемых российскими учеными совместно с европейскими коллегами, финансируемых Европейской комиссией, точнее ее подразделением CORDIS (Community Research and Development Information Service).

Например, проект INTAS 2004-77-7067 «Medical image mining: Theoretical foundation and technological aspects» направлен на развитие теоретических основ специализированного математического аппарата [37]. В нем также предложена реализация различных технологий и методологий для автоматического распознавания и анализа медицинских изображений.

Объективный анализ биомедицинских (цитологических и гистологических) изображений были предметом исследований в течение многих лет. Анализ медицинских изображений необходим для извлечения информации о состоянии больного и обеспечения правильного диагноза его заболевания. Одной из самых сложных задач в области анализа биомедицинских изображений является автоматизированное распознавание объектов из изображений и последующая классификация. Критическая проблема для автоматизации анализа биомедицинских изображений заключается в создании и исследовании алгебраических структур для представления изображения в качестве объектов анализа и распознавания, моделей преобразования изображения в качестве инструмента для эффективного синтеза и реализации основных процедур обработки и анализа изображений.

Практическая цель исследования – разработка:

- ◆ инструментария для решения основных задач автоматизированного анализа цветных биомедицинских изображений с использованием новых методов алгебры изображения, эффективных алгоритмов соответствия изображения, сегментации, нейронных сетей и автоматического преобразования растровых в векторные изображения;

- ◆ новых информационных технологий для морфологического анализа лимфоидных клеточных ядер у больных с опухолями кровеносной системы на основе комбинированного использования методов распознавания паттернов и методов анализа изображений.

Для оценки и проверки разработанных подходов и методов использовались медицинские базы данных изображений, содержащие более 12000 изображений. Кроме того, в целях разработки правил принятия решений для диагностики заболеваний включены морфометрические оценки биологиче-

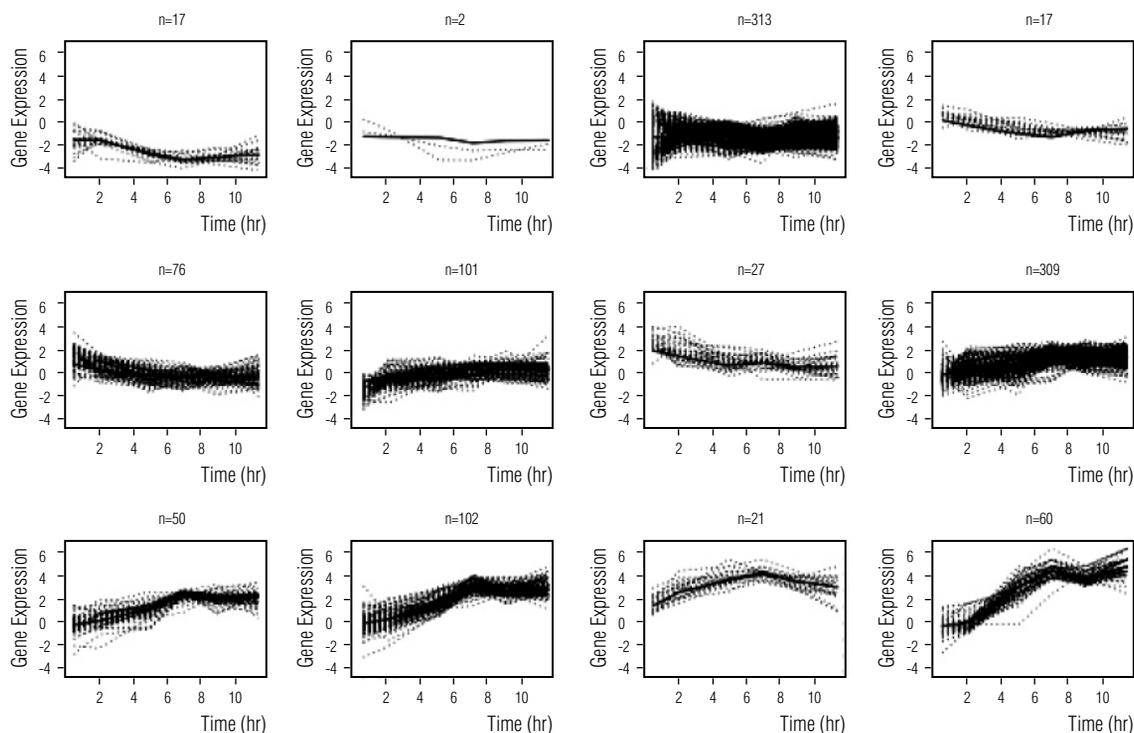


Рис. 6. Кластеры генной экспрессии

ских объектов в виде количественных параметров, характеризующих изменения клеточных ядер, волокон и тканей. В качестве иллюстрации эффективности полученных теоретических методов авторы проводили анализ изображений клеток крови и лимфатической ткани.

Еще один проект по смежной тематике – INTAS 2004-77-7347 «Principles of dissimilarity-based pattern recognition in signals, symbolic sequences and images (PRINCESS)» [38]. Научной целью проекта PRINCESS является разработка теоретической основы для алгоритмических технологий распознавания образов и паттернов данных. Особое внимание уделяется распознаванию объектов с сигналами, символьными последовательностями и изображениями. В нем применяется традиционный подход кластерного анализа, заключающийся в выделении такого рода информативных признаков в объектах, которые представляют каждый объект в векторном пространстве признаков, и использовании эвристических способов для оценки различия между сигналами, символьными последовательностями или изображениями, используя различные функции, которые обладают всеми свойствами метрики.

Задача проекта заключается в создании общей теоретической и методологической базы в следующих областях:

◆ нахождения эмпирических закономерностей и

паттернов данных в наборах сигналов, символьных последовательностей и изображений,

◆ структурное выделение паттернов в данных на основе различных метрик для описания сходства/различия объектов,

◆ распознавание образов с помощью классификаторов.

Практическая часть исследования состояла в применении рассмотренных методов к задаче распознавания человеческого лица на изображении, использования структурированной модели для 3D-реконструкции с изображения 2D и для задачи распознавания математических формул в печатных документах.

Два других проекта Европейской комиссии FP6-IST, а именно «Pattern analysis, statistical modelling and computational Learning» [39] и FP7-ICT «Pattern analysis, statistical modelling and computational Learning 2» [40], легли в основу создания полномасштабного общеевропейского распределенного института Паскаля (сайт института <http://www.pascal-network.org>), развивающего методы распознавания образов, статистического моделирования и машинного обучения.

Среди практических задач, которые рассматривают участники института Паскаля, можно выделить:

■ Компьютерное зрение

◆ Распознавание человека/объекта

◆ Слежение за человеком/объектом

■ Обработку и исследование текстов на естественных языках

- ◇ Машины поисковых запросов
- ◇ Переводы с разных языков
- ◇ Фильтрация, в том числе от спама
- ◇ Адаптивное управление веб-контентом

■ Машинное обучение

- ◇ Задачи классификации
- ◇ Задача кластерного анализа
- ◇ Задачи регрессионного анализа
- ◇ Задачи построения ранжирования

■ Анализ данных

- ◇ Проблема баз данных большой размерности
- ◇ Обнаружение закономерностей (Knowledge discovery)

■ Дополнительные задачи

- ◇ Распознавание речи
- ◇ Биоинформатика
- ◇ Портфельное инвестирование и риск-менеджмент
- ◇ Теория управления

Проект Паскаль активно развивается, в институте существуют программа поддержки и финансирования исследовательских программ «The Harvest Programme», проводятся соревнования (Challenges) по упомянутым темам (<http://pascallin.ecs.soton.ac.uk/Challenges/>), библиотека насчитывает свыше 6 тысяч публикаций и научных работ (<http://eprints.pascal-network.org/view/year/>).

5. Уточнение понятия паттерн и его использование для анализа динамики

Возможны различные уточнения понятия паттерн, связанные с вариативностью способов описания того, какие погрешности допускаются как в значениях признаков, так и количествах объектов, не покрываемых найденными паттернами. В частности, имеются три эквивалентных математических способа представления паттернов. Эти способы апеллируют к различным когнитивным подсистемам: один – к образной, другой – к логической, а третий – к геометрической.

Для примера рассмотрим одно из самых популярных множеств данных в машинном обучении и распознавании образов, так называемые Ирисы. Это множество задаётся таблицей 150×4 (см. <http://archive.ics.uci.edu/ml/datasets/Iris>), характеризующих 4 измерения 150 экземпляров цветка ирис, представленных в статье Р. Фишера по дискриминантному анализу [11], которая рассматривается как самая первая основополагающая статья по распознаванию образов.



Рис. 7. Чашелистик (*sepal*) и лепесток (*petal*) в цветке ириса

Цветы ириса в этих данных относятся к трем видам (таксонам): I *Iris setosa* (диплоид), II *Iris versicolor* (тетраплоид) и III *Iris virginica* (гексаплоид). Два признака таблицы относятся к измерениям длины и ширины чашелистика (w_1, w_2), а два – к таким же измерениям лепестков (w_3, w_4). Эти элементы на рис. 7 показаны стрелками.

Подсчитаем средние значения признаков на таксонах, нарисуем для каждого из них вертикальную ось, отложим средние на этих осях и соединим отрезками прямых средние, относящиеся к одному и тому же таксону (рис. 8). Такой способ представления многомерных объектов в литературе называют «параллельными координатами» (см., например, [10, 13]). Полученные ломаные в определённой степени характеризуют паттерны данных таксонов. Почему в определённой степени? Потому что здесь они проявляются на уровне средних, и не очень понятно, насколько эти паттерны характерны для всех объектов таксонов. Как видим, два паттерна более или менее похожи, тогда как паттерн, показанный сплошной линией (первый таксон), обнаруживает несколько иную структуру средних. Кроме того, совершенно очевидно, что по признаку w_2 паттерны практически не отличаются, так что этот признак следует исключить, чтобы паттерны действительно отличались друг от друга по всем признакам, участвующим в паттернах.

Рис. 9, выполненный с помощью операции *parallelcoords* системы МатЛаб, представляет паттерны трёх данных таксонов в различных системах признаков. Это, прежде всего, средние внутривидовые паттерны в исходных признаках (график слева), а также модификации этих паттернов, полученные после удаления «неразделяющего» признака w_2 , — ширина чашелистика (второй слева график). Рассматривая этот график, можно заметить, что средние одинаково упорядочены по всем трем признакам, причем по признакам w_3 и w_4 разница между видами 1 и 2

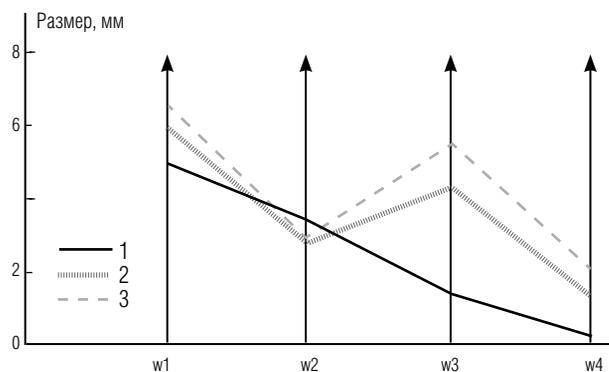


Рис. 8. Паттерны трёх таксонов по данным об ирисах на уровне средних

больше, чем между видами 3 и 2. Это означает, что межкластерные различия станут ещё больше, если сформировать новый признак путём умножения w_3 на w_4 . Действительно, паттерны в пространстве двух признаков, w_3 и w_4 , сильно расходятся (третий слева график). Это позволяет предположить, что в системе этих двух признаков таксоны ирисов будут иметь достаточно убедительно различающиеся паттерны. Рисунок справа подтверждает оправданность этого предположения. На нём представлены «ломаные» (в данном случае просто отрезки), соответствующие всем 150 объектам таблицы данных.

Как видим, представление в параллельных координатах даёт удобную форму для визуального анализа и преобразования данных с целью построения различающихся паттернов для заранее заданных таксонов.

Обратимся теперь к другому способу представления паттернов — через конъюнктивные предикаты. Рассмотрим для примера паттерны таксонов ириса на параллельных координатах самого правого графика на рис. 9. Приблизительно оценим, в ка-

ких интервалах, в основном, сосредоточены точки первого таксона, отмеченного сплошной линией, на обеих координатных осях. Скорее всего, это будут, например, интервалы: [4.3, 5.9] на оси w_1 и [0,1] на оси w_3*w_4 . Каждый из этих интервалов определяет «интервальный» предикат, значения которого — «истина» для точек интервала и «ложь» для точек вне его. Естественно обозначить первый предикат через $w_1[4.3, 5.9]$, а второй — как $w_3*w_4[0,1]$. Тогда предикат, соответствующий первому таксону будет не что иное, как конъюнкция этих двух, истинная на пересечении множеств истинности этих интервальных предикатов, $P_1 = w_1[4.3, 5.9] \& w_3*w_4[0,1]$. Это и есть паттерн первого таксона, представленный в виде предиката P_1 . Аналогично формируются конъюнкции интервальных предикатов, задающие паттерны для второго и третьего таксонов, $P_2 = w_1[5, 8] \& w_3*w_4[4,8]$ и $P_3 = w_1[6,8] \& w_3*w_4[7,16]$.

В принципе, можно рассматривать нечеткие предикаты, представляющие таксоны предикатами нечёткой логики, но здесь мы этого делать не будем. Заданные в форме предикатов паттерны образуют совокупность чётких классификаторов, которые для каждого конкретного объекта и заданного таксона определяют, принадлежит ли объект таксону или нет. При этом, как всегда, могут возникнуть ошибки в форме ложных плюсов и ложных минусов. Суммарная характеристика решений в данном примере представлена в табл. 1.

Как видим, только один предикат, а именно P_1 , в точности соответствует своему таксону. Два других предиката «покрывают» соответствующие таксоны на $41/50=82\%$ и $43/50=86\%$, причём выдают сбои трёх типов — ложные плюсы, ложные минусы и отказы от решения. На самом деле подобные уровни

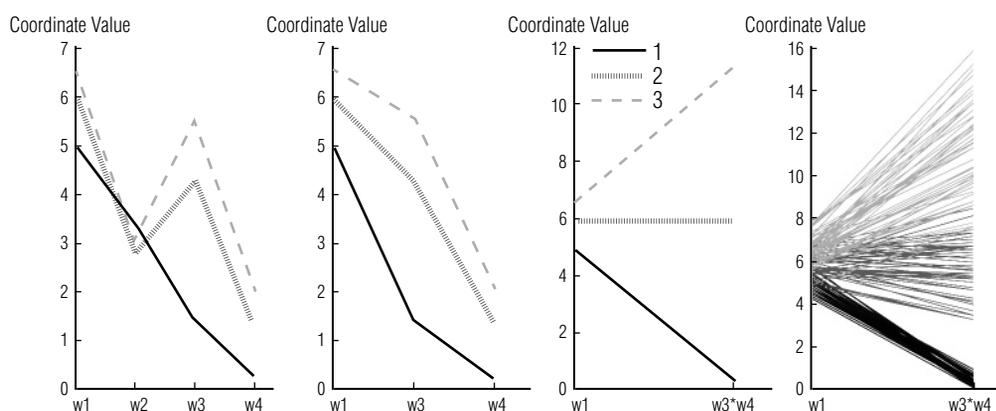


Рис. 9. Средние внутривидовые паттерны ирисов в разных системах признаков: (а) все четыре признака ирисов; (б) три признака; ширина чашелистика w_2 исключена; (в) два признака, w_1 и произведение измерений лепестка w_3 и w_4 , а также паттерны всех 150 объектов в разрезе этих двух признаков

Таблица 1.

Соответствие предикатов и таксонов ириса.
 «Остаток» представляет собой объекты,
 не покрытые ни одним из предикатов

Предикаты \ Таксоны	T1	T2	T3	Итого
P1	50	0	0	50
P2	0	41	3	44
P3	0	7	43	50
Остаток	0	2	4	6
Итого	50	50	50	

покрытия вполне удовлетворительны, если рассматривать паттерн как закономерность. Многие объекты удовлетворяют паттернам неточно, характеризуя как бы «пограничные» состояния. Кроме того, уровни покрытия и точности можно отчасти контролировать, изменяя границы интервалов, входящих в формируемый паттерн.

Следует указать, что конъюнкция интервальных предикатов взаимно-однозначно отображается в предикат, истинный только на симплексе прямого (декартова) произведения соответствующих интервалов. Это потенциально приводит к геометрическому формализму, выражающему паттерны многомерными симплексами (см. рис. 10). Геометрически паттерн на рис. 10(a) есть не что иное как декартово произведение соответствующих интервалов, а именно $[a1, a2] \times [b1, b2]$.

Проблемы формирования паттернов для заданных подмножеств в терминах семейств однородных ломаных на системах параллельных координат трактуются в книгах основоположника метода А. Инсельберга [13]. Метод формирования конъюнктивных предикатов описан в [19], но в этом методе предикаты формируются для каждого кластера независимо и, следовательно, могут использовать

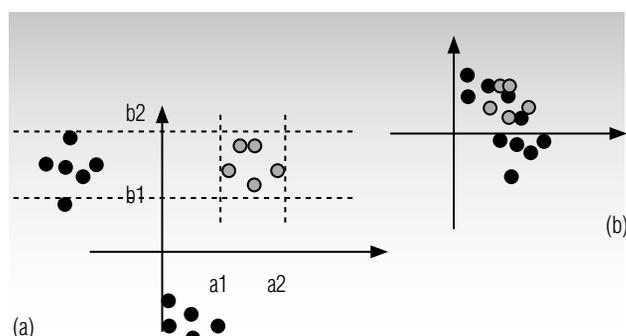


Рис. 10. Кластер, состоящий из светлых кружков, на графике (a) описывается паттерном, задаваемым предикатом $P = a1 < x < a2 \ \& \ b1 < y < b2$, без ошибок.

Напротив, аналогичный кластер на графике (b) зашумлён присутствием чёрных кружков и не может быть хорошо отделён от них конъюнктивным предикатом без использования преобразований координат

различающиеся системы признаков. При построении паттернов все множества должны рассматриваться в одной и той же системе признаков.

Ситуация усложняется тем, что на самом деле задача состоит вовсе не в том, чтобы описывать имеющиеся кластеры, а в том, чтобы определять кластеры, описанные паттернами. В этой более сложной задаче, к сожалению, пока похвастаться нечем. Формализм параллельных координат приводит к очень сложным проблемам подбора близких по форме ломаных как совокупностей отрезков прямых на плоскости, на которой изображены параллельные координаты [13], и для практического использования остаётся только разработка как можно более гибких систем, обеспечивающих возможность интерактивной работы исследователя с многообразными данными, визуализируемыми на экране компьютера. Подобных систем, в том числе бесплатных, в Интернете имеется довольно много; одна из последних разработок такого рода [22]. N.Mishra, D. Ron and R. Swaminathan утверждают, что их статья [21] посвящена проблеме автоматического формирования кластеров в виде конъюнктивных предикатов. Но это не так: название статьи не отражает существа дела – речь идет всего навсего о поиске бикластеров в бинарной матрице. Получается, что разработок по формированию кластеров в виде конъюнкций интервальных предикатов и, тем более, многомерных симплексов, т.е. декартовых произведений интервалов, вообще не делалось.

Это не оставляет нам иного выбора как использовать следующий двухэтапный метод для автоматизации формирования паттернов. На первом этапе происходит формирование кластеров с использованием обычного кластер-анализа, на втором – ищутся паттерны, достаточно полно представляющие полученные кластеры [1-3].

Когда данные характеризуют динамику функционирования социально-экономических объектов, к этому добавляется третий этап: оценка устойчивости их поведения с течением времени. В действительности смена паттерна во времени означает изменение стратегии деятельности субъекта или динамики развития объекта. В условиях относительного постоянства факторов внешней среды частота смены паттернов объектами отражает существенные внутренние параметры их структуры и системы управления. Таким образом, устойчивые по моделям поведения, т.е. не меняющие паттерна, объекты представляют особый интерес, поскольку могут представлять те объекты, которые относительно приспособились к среде и

определились со стратегией своего развития – в рамках присущего им паттерна. Напротив, социально-экономические объекты, часто меняющие паттерн, по-видимому, не могут найти свое место в относительно стабильных рыночных условиях и представляют специальный интерес с точки зрения изучения их поведения, поскольку могут выступать в качестве элементов повышенной волатильности в системе и обладать признаками банкротства организаций. Во второй части статьи мы попытаемся применить эту логику к анализу некоторых конкретных ситуаций.

6. Заключение

В работе предложено использовать термин «паттерн» для обозначения комбинаций значений значи-

мых признаков, характерных для определённых групп объектов, соответствующих отличающимся от других типам поведения. Проведен обзор литературы в каждом из трёх основных аспектов этого понятия. Среди них были выделены следующие: использование понятия «паттерн» в науке и технике, методы кластер-анализа, динамика многомерных объектов. Указаны математические формализмы, позволяющие адекватно представлять понятие паттерна в соответствии с каждой из трёх когнитивных подсистем: образной, логической и геометрической. Для данных о динамике функционирования социально-экономических объектов предложено использовать типологию функционирования объектов по частоте смены паттерна. Этот подход будет применён к анализу реальных данных во второй части статьи. ■

Литература

1. Aleskerov F., Ersel H., Yolalan R. Personnel allocation among bank branches using a two-stage multicriterial approach // *European Journal of Operational Research*. – 2003. – Vol. 148/1. – P. 116-125.
2. Aleskerov F., Alper C.E. A clustering approach to some monetary facts: a long-run analysis of cross-country data // *The Japanese Economic Review*. – 2000. – Vol. 51, No. 4. – P. 555-567.
3. Aleskerov F., Ersel H., Yolalan R. Clustering Turkish commercial banks according to structural similarities // *Yapi Kredi Discussion Paper Series*. – 1997. – No. 97-02. – Istanbul, Turkey.
4. Apresian Y.D. An algorithm for finding clusters by a distance matrix // *Computer. Translation and Applied Linguistics*. – 1966. – Vol. 9. – P. 72-79 (in Russian).
5. Bandyopadhyay S., Maulik U. An evolutionary technique based on K-means algorithm for optimal clustering in RN // *Information Sciences*. – 2002. – 146. – P. 221-237.
6. Bassanini A., Duval R. Employment patterns in OECD countries: Reassessing the role of policies and institutions // *OECD Economics Department Working Papers*, No. 486, 2006. – OECD Publishing. Источник в Интернет: <http://dx.doi.org/10.1787/846627332717>.
7. Biernacki C., Celeux G., Govaert G. Assessing a mixture model for clustering with the integrated completed likelihood // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 1990. – 22 (7). – P. 719-725.
8. Cheng Y., Church G.M. Biclustering of expression data // *Proceedings of 8th International Conference Intelligent Systems for Molecular Biology*. – 2000. – P. 93-103.
9. Ciaccio E.J., Dunn S.M., Akay M. Biosignal pattern recognition and interpretation systems. Part 4 of 4: Review of applications // *IEEE Engineering in Medicine and Biology Magazine*. – 1994. – Vol. 13, 2006, Issue 2. – P. 269-273.
10. Few S. Multivariate analysis using parallel coordinates, *Perceptual Edge*. Источник в Интернет: http://www.perceptualedge.com/articles/b-eye/parallel_coordinates.pdf.
11. Fisher R.A. The use of multiple measurements in taxonomic problems // *Annals of Eugenics*. – 1936. – 7. – P. 179-188.
12. Huang Z., Ng M.K., Rong H., Li Z. Automated variable weighting in k-means type clustering // *IEEE Transactions on Pattern Analysis and Machine Learning*. – 2005. – 27 (5). – P. 657-668.
13. Inselberg A. Parallel coordinates: Lecture notes, 2004. Источник в Интернет: <http://astrostatistics.psu.edu/su06/inselberg061006.pdf>.
14. Kanevski M., Timonin V., Pozdnoukhov A., Maignan M. Evolution of interest rate curve: Empirical analysis of patterns using nonlinear clustering tools // *ESTSP 2008 Proceedings*. Источник в Интернет: <http://www.mafy.lut.fi/timeseries/ESTSP/PDF/26.pdf>
15. Lance G.N., Williams W.T. A general theory of classificatory sorting strategies: 1. Hierarchical Systems // *Comp. Journal*. – 1967. – 9. – P. 373-380.
16. Laurillard D. *Teaching as a design science: Building pedagogical patterns for learning and technology*. – Routledge, 2012.
17. Lu Y., Lu S., Fotouhi F., Deng Y., Brown S. Incremental genetic algorithm and its application in gene expression data analysis // *BMC Bioinformatics*. – 2004. – 5. – P. 172.

18. von Luxburg U. A tutorial on spectral clustering // *Statistics and computing*. – 2007. – 17 (4). – P. 395-416.
19. Mirkin B. Clustering for data mining: A data recovery approach. – Chapman and Hall/CRC, Francis and Taylor, Boca Raton, FL., 2005.
20. Mirkin B. Mathematical classification and clustering. – Dordrecht-Boston-London: Kluwer Academic Publishers, 1996.
21. Mishra N., Ron D., Swaminathan R. A new conceptual clustering framework // *Machine Learning*. – 2004. – 56. – P. 115-151.
22. Steinberger M., Waldner M., Streit M., Lex A., Schmalstieg D. Context-preserving visual links // *IEEE Transactions on Visualization and Computer Graphics*. – December 2011. – Vol. 17, No. 12. – P. 2249-2258.
23. Wakefield J.C., Zhou C., Self S.G. Modelling gene expression data over time: Curve clustering with informative prior distributions. *Bayesian Statistics 7* / J.M. Bernardo, M.J. Bayarri, J.O. Berger (eds). – Oxford University Press, 2003.
24. Ward J.H., Jr. Hierarchical grouping to optimize an objective function // *Journal of American Statist. Assoc.* – 1963. – 58. – P. 236-244.
25. Бауман Е.В. Методы размытой классификации (вариационный подход) // *Автоматика и телемеханика*. – 1988. – №12. – С. 143-156.
26. Гамма Э., Хелм Р., Джонсон Р., Влссидес Дж. Приемы объектно-ориентированного программирования. Паттерны проектирования (*Design Patterns: Elements of reusable Object-Oriented Software*). – СПб: Питер, 2007.
27. Гапонова О.В. Электроэнцефалографические паттерны синдрома Веста // *Медицинский совет*. – 2008. – № 1-2.
28. Елкина В.Н., Загоруйко Н.Г. Об одном алфавите распознавания // *Вычислительные системы*. – 1966. – №12. – Новосибирск: Институт математики СО АН СССР.
29. Кузнецов Е.Н., Мучник И.Б. Монотонные системы для анализа организационных структур // *Методы анализа многомерной экономической информации*. – Новосибирск: Наука, Сиб. отд., 1981. – С. 71-83.
30. Лбов Г.С., Пестунова Т.М. Группировка объектов в пространстве разнотипных переменных // *Анализ нечисловой информации в социологических исследованиях*. – М.: Наука, 1985. – С. 141-149.
31. Малая медицинская энциклопедия. – М.: Медицинская энциклопедия. 1991-96 гг.
32. Миркин Б.Г. Группировки в социально-экономических исследованиях. – М.: Финансы и статистика, 1985.
33. Миркин Б.Г. Методы кластер-анализа для поддержки принятия решений: обзор. Препринт WP7/2011/03. – М.: Изд. дом НИУ ВШЭ, 2011.
34. Небожин А.И., Ситель А.Б. Паттерны боли при биомеханических нарушениях шейного отдела позвоночника // *Мануальная терапия*. – 2007. – № 1 (25). – С. 2-8.
35. Паттерны двигательных и чувствительных расстройств при патологии нервных структур в дистальных отделах верхней конечности // *Медицинский портал для врачей и студентов doctorspb.ru*. 2010. Источник в Интернет: http://doctorspb.ru/articles.php?article_id=1477
36. Поляков К. Прогноз – не роскошь, а инструмент управления // *Директор информационной службы*. – 2012. – № 2. – С. 40-44.
37. Проект INTAS 2004-77-7067 «Medical image mining: Theoretical foundation and technological aspects». 2005-2007. Источник в Интернет: http://cordis.europa.eu/search/index.cfm?fuseaction=proj.document&PJ_RCN=9945206
38. Проект INTAS 2004-77-7347 «Principles of dissimilarity-based pattern recognition in signals, symbolic sequences and images (PRINCESS)», 2008. Источник в Интернет: http://cordis.europa.eu/search/index.cfm?fuseaction=proj.document&PJ_RCN=9947067
39. Проект FP6-IST «Pattern analysis, statistical modelling and computational Learning». 2003-2008. Источник в Интернет: http://cordis.europa.eu/search/index.cfm?fuseaction=proj.document&PJ_RCN=6533866
40. Проект FP7-ICT «Pattern analysis, statistical modelling and computational Learning 2», 2008-2013. Источник в Интернет: http://cordis.europa.eu/search/index.cfm?fuseaction=proj.document&PJ_RCN=9905200
41. Сайт 19ой международной конференции по распознаванию образов <http://www.icpr2008.org/index.html>
42. Смит Н. Современные системы психологии. История, постулаты, практика (*Current systems in psychology. History, Theory, Research, and Applications*) / Пер. с англ., под общ. ред. А.А.Алексеева – М.: ОЛМА-ПРЕСС, 2003.

КРОССЛЕКСИКА: УНИВЕРСУМ СВЯЗЕЙ МЕЖДУ РУССКИМИ СЛОВАМИ

И.А. Большаков,

независимый исследователь, доктор технических наук, профессор,
почетный профессор Национального политехнического института Мексики

E-mail: iabolshakov@gmail.com

Адрес: г. Москва, ул. Лесная, 45

Создан сверхбольшой русский компьютерный словарь из 292 тысяч слов и выражений и 8 миллионов разнообразных связей между ними. Словарь предельно политематичен, рассчитан на любую аудиторию и запросы на русском и английском языках. Посредством лингвистической и энциклопедической информации он помогает в диалоге редактировать тексты и обучает языку. К словарю также могут обращаться внешние программы.

Ключевые слова: русский язык, компьютерный словарь, связи между словами, политематичность, диалоговый и программный доступ, запросы по-английски.

1. Введение

За последнюю четверть века русский язык существенно изменился.

- Пополнилась лексика. Накапливавшиеся в обществе разговорные слова и жаргонизмы выплеснулись на страницы изданий, в рекламу, на телеэкран, в Интернет. Появилось множество новых заимствований, многие слова приобрели новый смысл.

- Соответственно, изменился и пополнился состав словосочетаний, которыми, по формулировке И.Мельчука [1], только и говорит человек.

- Поляризовалась ситуация в части владения языком. На одном полюсе возросло число обозревателей, журналистов и ученых-гуманитариев, вир-

туозно владеющих языком и более не стесненных советскими речевыми штампами. На другом полюсе появилась масса «афтаров», которые демонстрируют в Интернете убогий язык, попирающий все нормы орфографии и приличия.

Речь, конечно, не идет о разрушении языка или «нервном срыве» [2], смежные поколения всегда договорятся друг с другом, но имеющиеся академические словари сильно устарели. Появившиеся крупные словари, например, [3, 4, 5, 6], как-то успевают истолковывать новации, но слабо отражают словосочетания, особенно – новые.

Создание словарей, отражающих допустимые и недопустимые связи между словами, стало крайне актуальным. Даже высоко грамотные люди могут не помнить или не знать, например, что *иск вчиняют*,

а католицизм практикуют. Людей с грамотностью пониже (а таких миллионы) нужно уберечь от «некультурных» выражений, как совсем примитивных (*более лучше, очень прекрасный, играть значение...*), так и более тонких (*поединок команд, более оптимальный, нелицеприятный разговор, возвести баню, раскаться о случившемся, потерпеть несправедливость, болтнуть глупость, принести урон...*). Таких «ляпов» множество и в текстах они нередки. Но есть еще больше языково-специфичных и притом вполне нормативных словосочетаний, которыми плохо владеют как многие с родным русским языком, так и особенно иностранцы: *решение можно принимать и выносить, тревогу испытывать и разделять, внимание обращать и уделять, сочувствие выразить и испытывать*. Чтобы понять русскую специфичность таких выражений, достаточно перевести их на другой язык пословно: правильный перевод обычно не получается. При этом не существует четкой границы между специфичными и неспецифичными (свободными) словосочетаниями. Поэтому желательно подобрать их как можно больше, допустимых и недопустимых. И пусть их число измеряется миллионами и постоянно растет, а цель исчерпывающего подбора всех словосочетаний недостижима.

Возможность отобразить миллионы межсловных связей в едином словаре появилась сравнительно недавно — благодаря быстрому совершенствованию компьютеров. Типовой объем их дискового накопителя увеличился за четверть века в тысячи раз. В памяти десктопа, ноутбука и смартфона умещаются тексты любого нужного объема. При выдаче словарей на экран уже не обязательно повторять их бумажный формат с его обычными сокращениями. Можно активнее использовать цвет, привлекать мультимедиа.

Для нас важно то, что теперь нет необходимости придерживаться привычного *линейного* построения словарей. Согласно этому принципу словарь — это последовательность статей, характеризующих смысл и грамматические категории своего заглавного слова (титула). Связи с другими словами при этом указываются лишь эпизодически. Отстают от этого принципа лишь в особых случаях, например, в тезаурусах и словарях синонимии или антонимии. Повышенный интерес к межсловным связям диктует *комбинаторный* принцип построения словаря, при котором титул, т.е. слово или устойчивое выражение, включается в него вместе со всеми обнаруженными на данный момент его связями с другими титулами.

Настоящая статья описывает новый компьютерный словарь русского языка *КроссЛексика* (CrossLexica), отвечающий возникшим потребностям и воплощающий следующие идеи:

♦ **Комбинаторный** принцип: титул включается только вместе со своими связями.

♦ **Декомпозиционный** принцип: титул — это однокорневое слово или словосочетание; в последнем случае полнозначные слова титула являются и отдельными титулами.

♦ **Политематичность**: словарь покрывает лексику большинства сфер использования языка, от высокой теории до низкого быта, содержит как лингвистические, так и энциклопедические знания.

♦ **Всеохватность** целевой *аудитории*: «от генерала до доярки».

♦ **Покрывание трех известных типов связей между титулами — синтаксических** (в словосочетаниях), **семантических** (= смысловых) и **паронимических** (= внешнего сходства).

♦ **Амбивалентность**: информация выдается словарем как пользователю в диалоге, так и внешней программе по ее запросу.

♦ **Встроенность в современный информационный мир**: словарь позволяет составить запрос и послать его в Интернет-поисковик.

♦ **Двуязычность**: есть *англо-русский* подсловарь, позволяющий по-английски запросить данные на русском языке. Можно также видеть английские переводы титулов.

♦ **Языковая многоуровневость**: кроме синтаксических и семантических связей, для каждого титула выдается *морфопарадигма*, т.е. список всех падежных форм для существительного и прилагательного и личных форм для глагола.

На июль 2013 г. КроссЛексика включает 292 тыс. титулов и более 8 млн. связей между ними, что по объему раз в 15 больше словаря Даля и уже совсем не подходит для печати. Словарь правится и пополняется постоянно, и очередная его версия всегда готова к работе с пользователем (*см. рекламу в данном номере*).

2. Источники и тематика контента

Источниками создания и пополнения КЛ явились:

♦ Академические словари русского языка, десятки словарей по экономике, бизнесу, электронике,

вычислительной технике, строительству и другим областям.

◆ Поток новостей, политические и научные статьи портала *газета.ру*.

◆ Десятки тысяч справок относительно слов и словосочетаний в Гугле и Яндексe.

◆ Рекламные буклеты, объявления по ремонту и строительству, издания о модах, туризме, персоналиях, автомобилях.

Все указанное накапливалось, структурировалось, размечалось и вводилось в базу словаря с 1990 г., в основном – вручную. Нас интересовала любая тематика, как то:

- Экономика, финансы и бизнес;
- Общественно-политическая сфера: политика, политология, социология, публицистика;
- Техника и технологии: радиоэлектроника, компьютеры, программирование, Интернет, автомобили, авиация, военная техника, бытовая техника, строительство;
- Точные и естественные науки: математика, физика, химия, биология, геология, география, экология;
- Гуманитарные и смежные сферы: психология, философия, история, лингвистика, лингвострановедение, искусство, религия;
- Медицина;
- Спорт;
- Кулинария;
- Бытовой язык, включая обценную лексику без слов трех жестко табуированных русских корней. (Но синонимы и эвфемизмы таких слов Госдума не запрещала!)

В итоге КЛ покрывает нужды самого широкого круга пользователей, включая ученых, преподавателей, инженеров, журналистов, бизнесменов, военных, студентов, школьников, пенсионеров, домохозяек.

3. Титулы, связи между ними, общая структура словаря

Титулы в КЛ относятся к четырем главным частям речи:

■ **Субстантивные:**

- Отдельное существительное: *абазур, битва, бифштекс, благо, блины...*
- Устойчивое именное словосочетание: *алко-*

гольные напитки, ближнее зарубежье, сельское хозяйство, точка зрения, уровень жизни...

■ **Глагольные в инфинитивной или личных формах:**

- Одиночный глагол: *говорить, идти, обсуждать, спать, ругать...*
- Глагольный оборот: *навести страх, оказывать внимание, испытать ужас...*

■ **Адъективные**, т.е. прилагательные или причастия:

- Отдельное прилагательное: *абстрактный, абхазский, авансовый, автономный, авантюрный, беж, воздушно-реактивный...*
- Отдельное причастие: *задвинутый, перевезенный, желающий...*
- Адъективный оборот: *бросающийся в глаза, хорошо одетый, бойцовской породы, большой дальности, бывший в употреблении, в елочку, как бархат...*

■ **Адвербиальные**, т.е. наречия или деепричастия:

- Отдельное наречие: *абсолютно, абстрактно, адски, аляповато, быстро...*
- Отдельное деепричастие: *базируясь, надев, торопясь, шепча...*
- Адвербиальный оборот: *аккуратным образом, более или менее, как банный лист, как выжатый лимон, в особой степени, куда попало, мелкой дрожью...*

Предлоги встроены в словосочетания и титулами обычно не являются, а многие прочие служебные слова условно считаются адвербиальными.

Упорядоченную по алфавиту совокупность титулов назовем **словником**.

Связи между титулами бывают трех базовых типов:

■ **Синтаксические** связи формируют словосочетания. Словосочетание у нас – это два полнозначных слова, синтаксически связанных и устойчиво совместимых по смыслу. Связь полнозначных слов может включать служебное слово (предлог или союз) по формуле *полнозначное слово1 → (служебное слово) → полнозначное слово2*, напр., *сотрудничество → ради → мира*, где стрелка означает синтаксическое управление.

■ **Семантические** связи соединяют слова со смысловым сходством. Это синонимы, антонимы и др., см. ниже.

■ **Паронимические** связи соединяют слова либо с буквенным сходством (например, отличающиеся лишь одной буквой слова *кадка* и *каска*), либо с морфемным сходством (например, однокоренные слова *кредитка* и *кредиторка*).

Глобальной структурой КЛ является матрица

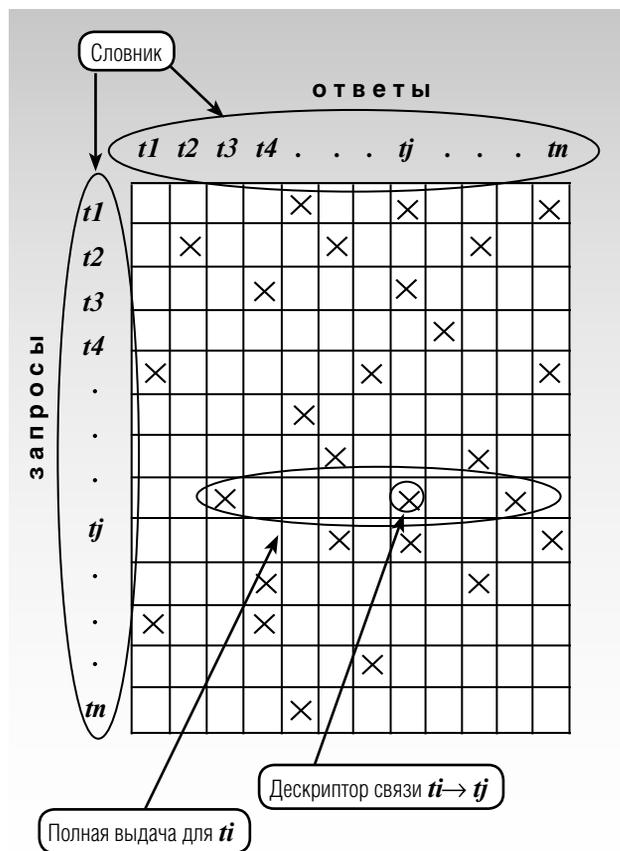


Рис. 1. КроссЛексика как матрица связей

{Словник × Словник} (рис. 1). Слева указаны титулы запросов, сверху – титулы ответов, крестиками отмечены дескрипторы связей между титулами. Связи бывают любого из указанных типов, а их дескрипторы указывают конкретный тип и подтип и включают, где следует, код соединяющего служебного слова. Запрос в виде титула t_i выхватывает из матрицы титулы ответов t_j , имеющих связи с t_i .

На допустимые связи между словами накладывает ограничения как сам язык, так и реалии внешнего мира. В итоге громадная матрица связей из $8.4 \cdot 10^{10}$ элементов оказывается крайне разреженной: непуст в среднем лишь каждый 10500-й ее элемент. Но это значит, что уже найдено 8 млн. непустых элементов.

4. Словосочетания

Словосочетания составляют самую важную и самую объемную часть КЛ. Их разнообразие сформировано различными подтипами синтаксических связей. Основная масса словосочетаний включает подтипы, зафиксированные в КЛ сотнями тысяч:

- ◆ Определительная пара «существительное – прилагательное»: *краснокочанная капуста, явный наглец, полная ясность, платье беж...*
- ◆ Определительная пара «глагол / прилагательное / наречие – наречие»: *полностью ясный, резко высказаться, ужасно страшно...*
- ◆ «Причастие/прилагательное – его дополнение–существительное»: *рассмотревший вопрос, ковырявший в носу, красный от гнева,...*
- ◆ «Глагол – его дополнение–существительное»: *ковырять в носу, рассмотреть вопрос, остаться из-за погоды, купить на рынке...*
- ◆ «Деепричастие/наречие – его дополнение–существительное»: *решив вопрос, ковыряя в носу, близко от города, купив на рынке...*
- ◆ «Существительное – подчиненное ему существительное»: *сердце матери, отличия в произношении, борьба против терроризма...*
- ◆ «Подлежащее–существительное – сказуемое в виде личной формы глагола или краткого адъектива»: *самолет вылетел, доклад краток, враг напал, глазки бегают...*

Прочие словосочетания представлены в КЛ десятками или единицами тысяч. Вот наиболее представительные из них:

- ◆ «Устойчивые сочиненные пары»: *наука и техника, власть и бизнес, ясный и четкий, взвесить и решить, в срок и в полном объеме, авиа- и железнодорожный транспорт...*
- ◆ «Глагол – его инфинитивное дополнение»: *собраться поехать, хотеть перекусить...*
- ◆ «Существительное – его инфинитивное дополнение»: *соблазн взять, желание уйти...*

На сегодня накоплено 2.26 млн. словосочетаний, и каждое из них доступно с двух сторон, что эквивалентно 4.52 млн. односторонних связей между словами.

5. Семантические связи

Семантических связей суммарно 2.77 млн. Наиболее широко представлены:

- **Синонимы:** 22.1 тыс. групп по 5.4 элементов, 1.291 млн. связей;
- **Смысловые производные:** 4.2 тыс. групп по 14.6 элементов, 1.096 млн. связей. Вот пример группы смысловых производных: {*извлечение, извлечения; извлекать, извлечь; извлеченный, извлекающий, из-*

влекший; *извлекая, по извлечении, путем извлечения*};

● **Когипонимы** («одноклассники»): 224 тыс. связей. Например, когипонимы для *мясо* – *вырезка, грудинка, гуляш, котлеты, паштет, фарш*;

● **Ассоциации**: 81 тыс. связей. Например, ассоциации для *аденоиды* – *аллергия, бассейн, гомеопатия, кашель, лазеротерапия, слух*.

Прочих семантических связей заметно меньше. Это

○ **Меронимы/холонимы** (= части/целые), напр., *террариум* – *зоопарк*;

○ **Гипонимы/гиперонимы** (= виды/роды), напр., *диплом* – *документ*;

○ **Антонимы**.

Все приведенные типы семантических связей хорошо известны в лингвистике, кроме ассоциаций, являющиеся нашей новацией. Они основаны на сочиненных парах, широко представленных в запросах к Интернет-поисковику или в их базах данных. Например, ассоциация *аденоиды* – *аллергия* включена, поскольку в поисковиках замечены пары *аденоиды и аллергия, аллергия и аденоиды*. Такие ассоциации полезны для составления запросов к Интернету, на их основе можно строить различные онтологии, а их статистика характеризует профиль массового пользователя Рунета.

Семантические связи в целом полезны следующим:

■ Они помогают понять смысл титула. Среди синонимов много толковательных, например, смысл титула *халляльный* хорошо поясняется его синонимом *отвечающий мусульманским нормам*. Гиперонимы (родовые понятия) давно служат в науке ядрами определений. Увидев, например, для *эндометриоз* гипероним *акушерско-гинекологическая болезнь*, пользователь едва ли захочет полного толкования.

■ Вместе с синтаксическими связями семантические позволяют прогенерировать словосочетания, в словаре отсутствующие. Так, исходя из содержащихся в базе словосочетания *букет цветов* и соотношения **Гипероним** (*каллы*) = *цветы*, для *каллы* порождается словосочетание *букет калл*, в базе отсутствующее. Такие «логические» выводы КЛ совершает миллионами, но из-за негарантированной правильности результаты выводов выдаются на экран низким контрастом.

■ Семантические связи позволяют выдавать разнообразные энциклопедически справки:

○ Названия континентов, океанов, крупных морей, горных цепей и других геообъектов;

○ Названия крупнейших городов мира в привязке к странам;

○ Сведения о 60 странах, по 20 крупнейшим из них – более подробные: столица, валюта, способ правления, титул главы государства, титульная нация, государственный язык, единица административного деления, преобладающая религия;

○ Названия и другие сведения о десятках городов и регионов России, включая наименования жителей: *Тула* – *туляки*, *Архангельск* – *архангелогородцы*, *Курск* – *куряне...*; для Москвы приведены названия всех ее районов;

○ Около 300 наиболее частых русских имен вместе с их диминутивами (*Сергей* – *Серезжа*);

○ Имена известнейших политических, деловых, научных и культурных деятелей мира.

○ Названия наиболее крупных организаций (корпораций) мира.

○ Названия наиболее известных художественных произведений (романов, фильмов, опер, мюзиклов и пр.)

Энциклопедической является также широко представленная терминология экономики, финансов, бизнеса, медицины, точных, естественных и гуманитарных наук.

6. Шаги навстречу рядовому пользователю

Для удобства рядового пользователя нужно было упростить интерфейс и сделать ненужным освоение сугубо лингвистических терминов. Для этого было предпринято следующее.

Исключены привычные пометы, используемые в печатных словарях. Оставлено два типа помет – степени разговорности (стиля) и степени фигуральности. Они служат пользователю рекомендациями и побудительными стимулами.

У разговорности пять градаций, размеченных цветом:

<нет пометы> Хорошо бы знать и уметь употребить это слово или выражение (*книга, налоги, конституция, роуминг...*).

<зеленый буллит> Специальное, книжное или забытое слово или выражение (*парадигма, афедрон, квадруольный...*); пользуясь им, опасайтесь непонятности.

<желтый буллит> Чисто разговорное слово или выражение (*мотать нервы, жевать сопли, мочить в сортире...*); не пользуйтесь им в официальных документах. Данная градация получилась слишком обширной, поскольку мы пока не выделяем из нее слова и выражения, имеющие сильную эмоциональную окраску.

<красный буллит> Обсценное слово или выражение (*говно, жопа, мудака, взять за яйца...*); не пользуйтесь им при дамах, детях и в официальной обстановке.

<серый буллит> Бытующее «некультурное» выражение (*оплатить за проезд, пошить пальто...*); правильнее передавать его смысл иначе. (Эквивалентный нормативный вариант, как правило, дается в виде синонима.)

Элементы любой расцвеченной градации можно отбросить в рамках данной сессии.

У фигуральности три градации:

<нет пометы> Понимается как есть (*идти в школу, вызвать слесаря*);

(*fig*) Понимается только фигурально (*сесть в галошу, висеть на волоске*);

(*mb-fig*) Понимается либо фигурально, либо в прямом смысле (*сесть в лужу, первая ракетка*).

Перед началом работы можно выбрать версию интерфейса – научную либо народную. В обеих версиях элементы меню, толкования омонимов и хелп-информация даются по-русски, но разделы выдачи именуются в научной версии строгими терминами (*Синонимы, Гипонимы, Когипонимы...*), а в народной – их «народными» синонимами (*Сходные по смыслу, Подклассы, Одноклассники...*). Реализованы и две аналогичные версии с английскими заголовками.

Малоподготовленному пользователю и иностранцу трудно выделить из множества выдаваемых словосочетания те, что подлежат первоочередному освоению. Но он может выбрать в рамках сессии частотный порядок выдачи определительных конструкций вместо алфавитного. Тогда наиболее многочисленные в КЛ определения предшествуют остальным. Можно вообще отсечь в данной сессии редкие словосочетания.

Запрос можно ввести по-разному: набором с клавиатуры; выбором строки в словнике; выбором строки выдачи на экране (это шаг навигации по словнику); шагом назад–вперед по списку *История*; вводом эквивалентного английского слова.

7. Различные приложения словаря

Предусмотрено несколько видов приложений КЛ:

♦ **Диалоговое** (интерактивное) приложение: пользователь вводит запрос в интерактивном режиме и использует выдачу, например, для углубленного изучения русского языка или параллельного редактирования текста за компьютером. Особо ценно для редактирования то, что словарь не только помогает устранять неуместные по стилю обороты, но и дает массу адекватных синонимических перифраз. Поскольку пассивное знание у многих носителей языка значительно превосходит объем активно используемых ими языковых средств, при показе того, как можно выразить ту же мысль иначе, обычно легко обнаруживается более подходящий вариант.

♦ **Интерфейсное** приложение: с помощью КЛ пользователь формирует запрос к Интернету, обращается к нему прямо из словаря и получает результаты поиска.

Оба этих вида приложений реализуются текущей версией КЛ.

В **недиалоговых** приложениях внешняя программа обращается к словарю через специальную утилиту КЛ и использует выдачу самостоятельно. Сами приложения не входят в КЛ и находятся на разных этапах отдельной разработки. В первую очередь эти приложения включают:

♦ **Обнаружение и исправление смысловых ошибок** типа *истерический центр*. В тексте ищутся синтаксически связанные пары слов, не представленные в КЛ и, скорее всего, неосмысленные. Просматриваются паронимы для обоих слов сомнительной пары, и ищутся имеющиеся в КЛ их сочетания. Найденные словосочетания предлагаются пользователю на выбор в качестве исправлений.

♦ **Разрешение неоднозначности омонимов**. Просматриваются словосочетания и семантические связи для омонимичных слов теста, и выбирается тот омоним, для которого в контексте найдено наибольшее число синтаксически и семантически сочетающихся соседей.

♦ **Стеганография и стеганализ**. Текстовые словосочетания и синонимы текстовых слов используются для регулируемой замены одних синонимов другими, дабы этими заменами закодировать стороннюю информацию, которая тем самым тайно передается несущим текстом без изменения смысла носителя и затем декодируется.

♦ **Фильтрация синтаксических разборов**. В ана-

лизируемом предложении отмечаются все словосочетания, наличествующие в КЛ, и чем больше обнаружено таких словосочетаний в данном варианте разбора предложения, тем вероятнее он считается.

Среди других возможных чисто автоматических приложений упомянем определение авторской принадлежности текстов и их общей тематики, рациональное деление текста на абзацы, идиоматичный перевод английских словосочетаний. На базе КЛ можно создавать тематически специализированные «мини-КроссЛексиксы», например, по экономике или вычислительной науке. Можно также строить онтологии, отражающих те многочисленные связи между понятиями современного мира, которые выходят за рамки синонимии, антонимии, гиперонимии / гипонимии и смысловой производности.

8. Утилиты создания рабочей версии и интерфейс словаря

При создании КЛ не использовались стандартные СУБД, и матрица связей была воплощена программой на C++. Сейчас КЛ работает под ОС Windows XP и Windows 7.

Отдельно программировались многочисленные утилиты морфоклассификации титулов и кодирования дескрипторов связи на основе исходной текстовой базы. Создана и специальная утилита сборки очередной рабочей версии.

Интерфейс КЛ кардинально отличается от привычных словарей. На рис. 2 представлена типовая выдача и размечены основные составляющие интерфейса: закладки сверху отображают словники, а закладки справа указывают начала различных

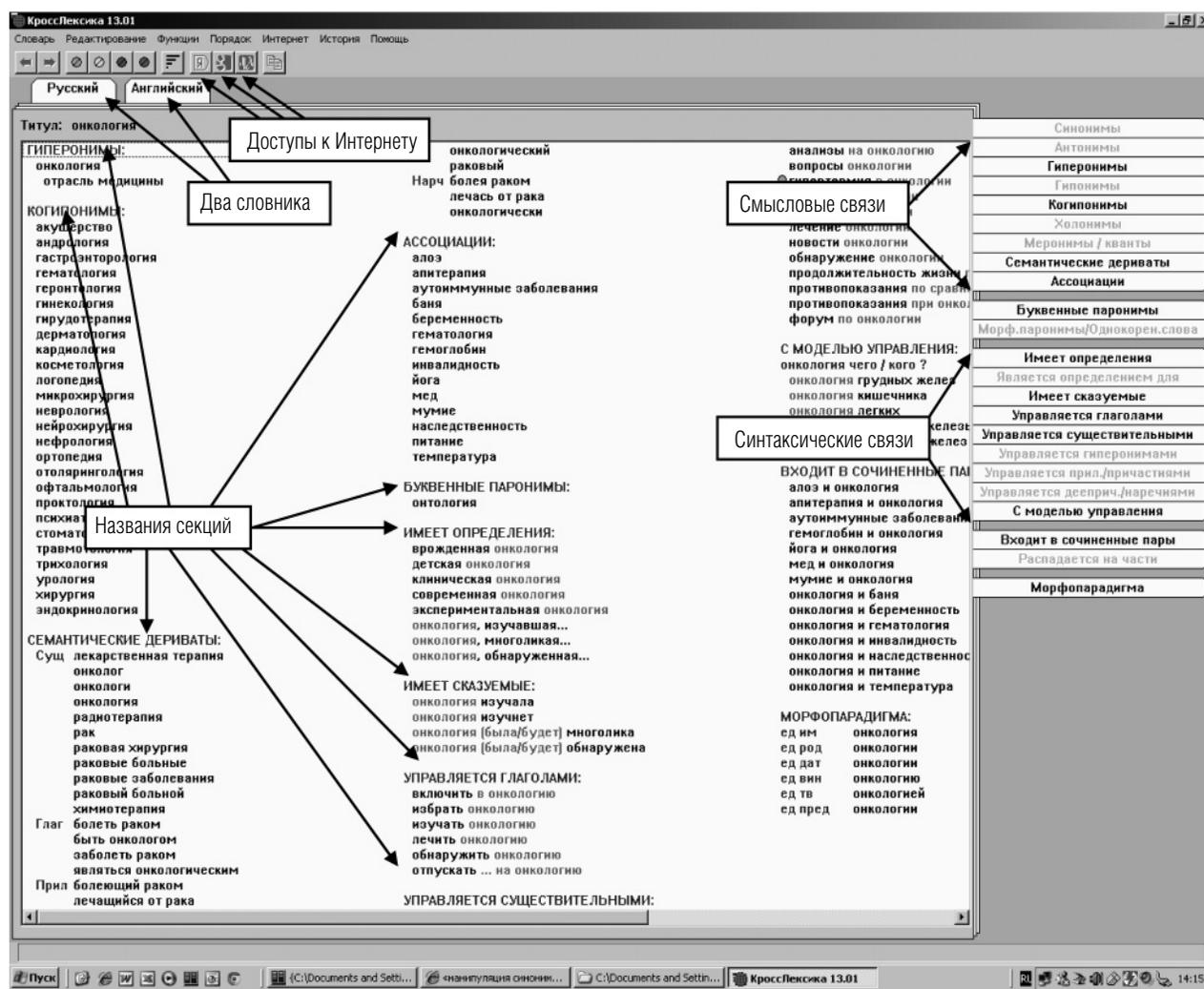


Рис. 2. Выдача для титула «онкология»

секций выдачи, в которых собраны титулы с одинаковым типом дескрипторов связи. В данном примере все секции уместились на одном экране. По верхнему краю экрана идут доступные пользователю функции управления и дублирующие их кнопки.

Если собрать у всех титулов секции синонимов, получим синонимический словарь, по объему превышающий все известные печатные словари. Тем же способом могут быть получены крупные словари антонимов и семантических дериватов, тезаурусы гипонимов / гиперонимов (= надклассов / подклассов) и холонимов / меронимов (= частей / целых), словари глагольно-именного управления, сочиненных пар и др. Тем самым, КЛ объединяет всех их под одной крышей.

КЛ не содержит толкований всех титулов подряд. Есть лишь краткие толкования омонимов, многочисленные семантические связи и английские переводы титулов. За толкованиями в процессе диалога можно обращаться через КЛ, например, в Викисловарь, но по объему обработанного словника он пока существенно уступает КЛ. В недиалоговых приложениях КЛ вполне самодостаточна, поскольку обычные словарные толкования, пусть даже оцифрованные, для современных автоматических средств обработки текстов бесполезны.

9. Заключение

Создан новый компьютерный словарь – ресурс, по объему и структуре не имеющий аналогов ни для одного языка мира. В то время как в [7] собрано 270 тыс. русских словосочетаний для 2.5 тыс. слов, а в [8] – 250 тыс. английских словосочетаний для 9 тыс. слов, словник КроссЛексики на порядок больше, словосочетаний здесь 2.26 млн., отражены также миллионы несинтаксических связей, и все это выдается в рамках единого компьютерного интерфейса.

КроссЛексика уже готова как вспомогательное средство редактирования сложных текстов любой тематики и обучения русскому языку в его деловой, литературной и разговорной ипостасях. Также КЛ позволяет поднять существующие типы систем автоматической обработки русских текстов на уровень, ныне недоступный.

Применительно к грядущему словарю [9], реализующему заметно более глубокий, подлинно академический анализ связей между русскими словами, КроссЛексику можно рассматривать как обширнейший массив сырого материала, который, правда, уже отвечает нуждам широкого круга пользователей-лингвистов.

Выражаю благодарность А.Ф.Гельбуху, кто спроектировал и запрограммировал интерфейс словаря и много лет бескорыстно помогал мне в многотрудном деле. ■

Литература

1. Mel'čuk I. Phrasemes in Language and Phraseology in Linguistics // M.Everaert et al. (Eds.) Idioms: Structural and Psychological Perspectives. – Lawrence Erlbaum Associates Publ.: NJ / Hove, 1995. – P. 169-252.
2. Кронгауз М.А. Русский язык на грани нервного срыва. – М.: Языки славянской культуры. Знак, 2008.
3. Толковый словарь русского языка начала XXI века. Под ред. Г. Н. Складчиковой. – М.: Эксмо, 2007.
4. Крысин Л.П. Толковый словарь иноязычных слов. – М.: Эксмо, 2008.
5. Большой толковый словарь русских существительных. Под ред. Л.Г.Бабенко. – М.: АСТ-Пресс Книга, 2009.
6. Толковые словари LINGVO, компания ABBY, 2012.
7. Словарь сочетаемости слов русского языка. Под ред. П.Н.Денисова и В.В.Морковкина. – М.: Русский язык, 1983.
8. Oxford Collocation Dictionary for students of English. – Oxford, NewYork: Oxford University Press, 2009.
9. Апресян В.Ю., Апресян Ю.Д., Бабаева Е.Э., Богуславская О.Ю., Галактионова И.В., Гловинская М.Я., Иомдин Б.Л., Крылова Т.В., Левонтина И.Б., Птенцова А.В., Санников А.В., Урысон Е.В. Проспект активного словаря русского языка. Под ред. акад. Ю.Д.Апресяна. – М.: Языки славянских культур, 2010.

ИСПОЛЬЗОВАНИЕ ПРИНЦИПОВ ЭРГОНОМИЧЕСКОЙ СЕМИОТИКИ ПРИ ПРОЕКТИРОВАНИИ ПОЛЬЗОВАТЕЛЬСКИХ ИНТЕРФЕЙСОВ В ПОЛИКУЛЬТУРНОМ КОНТЕКСТЕ

Ю.В. Таратухина,

кандидат филологических наук, доцент кафедры инноваций и бизнеса
в сфере информационных технологий Национального исследовательского
университета «Высшая школа экономики»

Д.А. Алдунин,

студент факультета бизнес-информатики Национального исследовательского
университета «Высшая школа экономики»

E-mail: jtaratuhina@hse.ru, daaldunin@edu.hse.ru

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

В настоящее время происходит активное внедрение информационных технологий в процесс обучения. Это могут быть как электронные учебные пособия, учебные среды, так и сайты, и сетевые сообщества. Очень часто ресурсы подобного рода подразумевают интернациональную пользовательскую аудиторию. Однако, в разных культурах существуют дифференцированные подходы к созданию эргономичного дизайна пользовательских интерфейсов. Данная работа посвящена анализу дифференциации подходов и выработке рекомендации по совершенствованию эргономичного дизайна электронных образовательных материалов, рассчитанных на поликультурную аудиторию.

Ключевые слова: эргономическая семиотика, эргономический дизайн, электронные образовательные материалы, поликультурная аудитория.

1. Введение

С началом использования информационных технологий в образовании появились новые средства и методы обучения, которые отчасти облегчают, а отчасти и увеличивают психологическую

напряженность образовательного процесса. По мнению Г.Я. Узилевского [6], эргономическая семиотика является научно-практической специализацией, исследующей проблемы общие для семиотики, лингвистики и эргономики и нацелена на разработку принципов проектирования и создания эффективных

языков человеко-компьютерного взаимодействия. В эпоху интернет-обучения, связанного с проектированием электронных обучающих сред и пособий, изучение кросс-культурных аспектов эргономической семиотики представляется наиболее важным.

Исходя из этого, к целям эргономической семиотики можно отнести определение природы пользовательского интерфейса и установление принципов создания эффективных и удобных знаковых средств для взаимодействия с разнообразными видами техники. В данной работе наибольший интерес будет представлять межкультурный контекст, который, в силу необходимости адаптировать учебные материалы для когнитивной специфики представителей разных культур, на сегодняшний день является самой проблемной областью. В вышеназванном контексте эргономическая семиотика предлагает следующие направления исследований:

- ◆ выявление и изучение подходов к созданию и использованию пользовательского интерфейса как интерактивной многоуровневой системы в разных культурах;

- ◆ цветовой код в разных культурах;
- ◆ иконический язык в разных культурах;
- ◆ музыкальный код в разных культурах;
- ◆ специфические параметры поведения пользователей на образовательных сайтах в разных культурах, обусловленные спецификой мышления и деятельности в разных культурных группах. Междисциплинарная область исследования обуславливает использование терминологического аппарата из разных предметных областей.

2. Пользовательский интерфейс как знаковая система

Пользовательский интерфейс в контексте эргономической семиотики может быть определен как интерактивная многоуровневая информационная система, состоящая из естественных языков, иконического языка, цветового и музыкального кодов. Термин «дружественный интерфейс» можно рассматривать применимо к прагматической составляющей с учетом удобного представления вербального и иконического сегмента. В различных культурах дифференциация вышеобозначенных параметров будет заметной. Если говорить об эргосемиотических требованиях и оценке пользовательского интерфейса, то представляется целесообразным их дифференцировать на:

- ◆ **Прагматические требования**, которые включают в себя легкость изучения, удобство восприятия и ис-

пользования информации, способствующие повышению эффективности деятельности пользователя и т.п. Именно в контексте этих требований мы в дальнейшем планируем рассматривать различные аспекты проектирования пользовательских интерфейсов с учетом культурных особенностей пользователей. С позиции прагматики должно быть соответствие когнитивным, психомоторным, эмоциональным, мотивационным характеристикам пользователей, а также соответствие их потребностям и задачам;

- ◆ **Семантические требования:** устойчивость к семантическим ошибкам, системообразующие связи, наличие обратной связи о результатах деятельности пользователя. С позиции семантики они должны отвечать нормам той или иной предметной области знания с максимальным учетом предметного языка пользователей;

- ◆ **Синтаксические требования:** гибкость, свобода выбора путей поиска информации и реализация данных связей в конкретной информационной системе;

- ◆ **Сигматические требования:** отношения между объектом, мыслительным отражением данного объекта и знаками, представляющими данное отражение; иначе говоря, логические связи и смыслы.

Таким образом, поликультурный контекст ставит перед нами следующие основные задачи, которые планируется решить в ходе дальнейшей работы:

1. Определение критериев пользовательской пригодности интерфейса для представителей разных культур;
2. Описание соответствия функциональных и графических характеристик интерфейса специфике когнитивной деятельности пользователей различных культур, во избежание возможных фрустраций, ошибок, снижения скорости выполнения задач;
3. Описание основных принципов национальных мультимедийных культур.

Если обратиться к практике, то несложно заметить, что дизайн пользовательских интерфейсов в разных странах демонстрирует культурные различия в части набора цветов, степени помощи в навигации, насыщенности информацией, степени ее сгруппированности и т.д.

Исследуя культурные различия, Р.Нисбетт [9] выделил несколько факторов, различающихся в зависимости от культурной принадлежности и оказывающих влияние на поведение индивидов: на Востоке это «внимание к полю», а на Западе — это

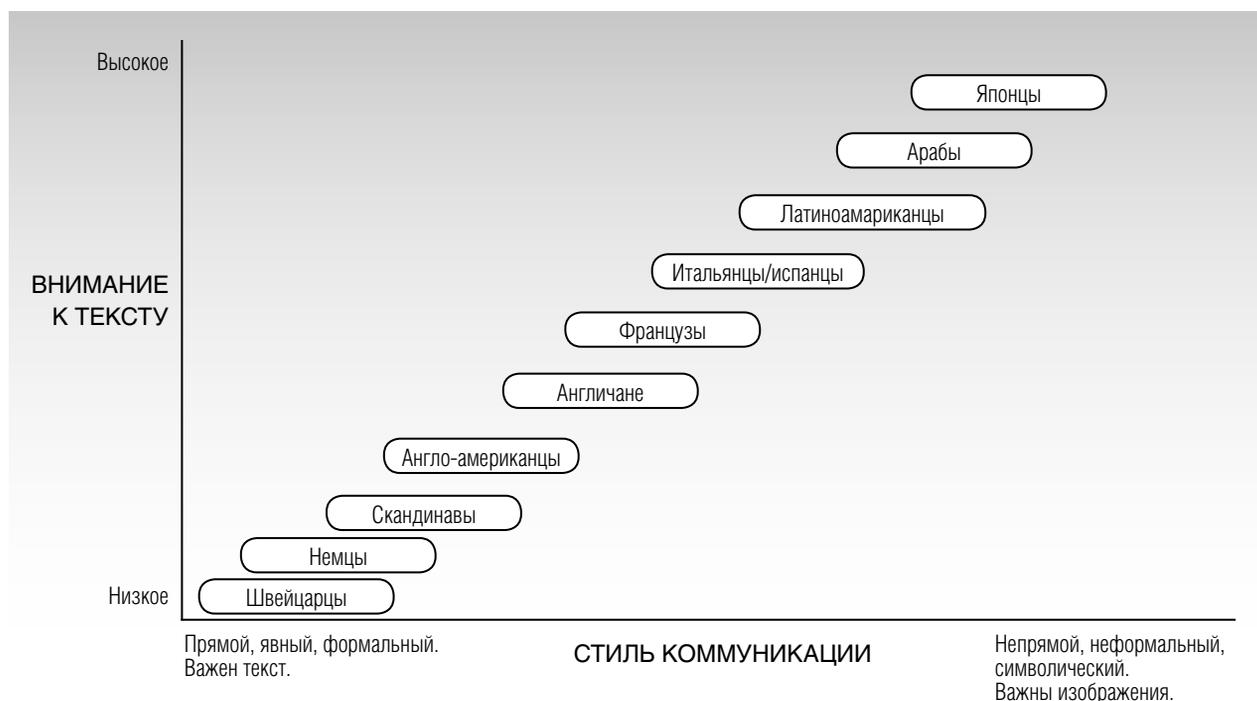


Рис. 1. Зависимость предпочитаемого стиля коммуникации в разных культурах от контекста

«внимание к основным объектам». Автор проводил ряд экспериментов с американскими и японскими пользователями, в результате которых выяснилось, что японцы обращают внимание на свойства окружающего фона примерно на 70% чаще американцев, хотя и те, и другие были одинаково склонны упоминать детали. Кроме того, японцы почти в два раза чаще отмечали взаимосвязи и отношения, включающие неодушевленные аспекты окружающей среды. Нисбетт увидел причины данных различий в когнитивно-семиотических механизмах, унаследованных либо от холистической (древнекитайской), либо от аналитической (древнегреческой) систем познания мира. На *рис. 1* мы можем наблюдать зависимость доминирующего стиля коммуникации в разных культурах от контекста.

Так же, опираясь на исследования Г.Хофстеде [10], можно отметить, что азиаты больше склонны к восприятию целостной картины, практически не используя разделение на категории и формальную логику. Они больше опираются на «диалектически» и эмпирические аспекты. В то же время, европейцы предпочитают строгую каталогизацию и сохранение формальной структуры. Согласно исследованиям М.А. Холодной [7], когнитивный стиль отражает способ восприятия, анализа, структурирования и категоризации мира, стиль учения: в большинстве западных культур преобладает «вы-

сокая когнитивная сложность» — многомерная модель реальности во множестве взаимосвязей. Для восточных культур характерна «низкая когнитивная сложность» — однозначная, упрощенная интерпретация реальности. Вся эта специфика так или иначе должна быть отражена на специфике взаимодействия пользователя с сетевыми образовательными ресурсами, и ее понимание поможет адаптировать электронные образовательные ресурсы под индивидуальные и социокультурные предпочтения пользователей. Вышеперечисленные особенности, безусловно, должны быть учтены разработчиками, имеющими дело с пользователями различных культурных групп. В первую очередь это должно учитываться при структурировании и каталогизации информации: представителям западных культур чаще требуется детальная информация по конкретному аспекту, тогда как представители восточных культур скорее захотят изучить вопрос в целом.

Изначально было предположено, что фактором, объединяющим человека и интерфейс, являлась информация. Сейчас же считается, что это деятельность [5]. Мы считаем, что оба фактора одинаково важны: работу с информацией, формирование представлений о предмете следует рассматривать как активность стратегическую, а взаимодействие с интерфейсом, направленное на получение этой информации, — как активность тактического ха-

рактера.

Деятельность представляет собой совокупность действий, с помощью которых достигается цель. Можно предположить, что в разных культурах структура деятельности будет неоднозначна по пространственно-временным показателям. Если следовать классификации Р.Льюиса [8], то специфике деятельности в разных культурах можно классифицировать на моноактивный, полиактивный и реактивный стили.

Таким образом, нам представляется важным выделить следующие аспекты:

- ◆ Учет взаимосвязи специфики мышления и деятельности и видов интерфейсных элементов в культурно-специфичном ключе;
- ◆ Тип и содержание информации (культурно-специфичные параметры);
- ◆ Структура и последовательность расположения элементов на экране, количество и детальность элементов в поле восприятия – прагматическая специфика;
- ◆ Семантический анализ высказываний и профессиональной терминологии – инструкции, подсказки, названия элементов системы и их прагматическая адекватность;
- ◆ Приемы невербального кодирования в виде иконок, сигналов, цветовых изображений;
- ◆ Наличие адекватной обратной связи.

3. Специфика проектирования культурноспецифичного интерфейса образовательного ресурса

В прагматическом ключе создание и проектирование интерфейса можно отразить метафорой «модель мира». Следовательно, основной задачей моделирования эргодизайна интерфейса и текста общеобразовательного сетевого ресурса и будет являться анализ данных критериев для построения своеобразной модели – общих дизайнерско-оформительских аспектов текста, наиболее адекватных для восприятия текстовой информации пользователем.

Все требования, представленные для эргономичной организации интерфейса сетевого ресурса, можно разбить на четыре основные категории:

- ◆ Навигация;
- ◆ Архитектоника и структура страницы;
- ◆ Возможность обратной связи, доступа, попол-

нения контента; доминирующий стиль репрезентации информации;

- ◆ Цветовое и шрифтовое решение;

Нами был определен ряд наиболее существенных параметров для дальнейшего анализа образовательных вебсайтов.

В первую очередь важна **иконическая символика**. В разных когнитивных контекстах будет возникать сложность в установлении однозначного соответствия между конкретными понятиями и иконическими знаками. Соответственно для создания дружественного интерфейса, рассчитанного на поликультурную аудиторию необходим профессионализм в проектировании иконических знаков, так как необходимо проанализировать предметную область и культурные особенности целевой аудитории и определить состав и отличительные особенности иконических знаков.

Второй класс проблем в иконике будет связан с репрезентацией абстрактных понятий в виде объектов, выраженных иконическими знаками или представленными в виде визуальных метафор. Необходимо учитывать репрезентацию абстрактных понятий в разных культурах, дифференцирование внутри иконической символики, образных объектов, пиктограмм, указывающих на характер выполненных действий, пиктограмм, используемых для функционального аналога и обозначающих результат выполненных действий.

Как известно, основные функции иконических знаков следующие: свойства подобия тому или иному объекту; замена или репрезентация объекта; иллюстративная и коммуникативная функции.

Совершенно очевидно, что изображения (значки, «иконки»), могут совершенно по-разному восприниматься представителями разных культур. Например, арабские ресурсы используют очень многонациональной и религиозной символики, независимо от темы, которой посвящен материал, очень активно национальная символика используется на японских ресурсах. Не все иконические знаки, понятные представителям западных культур могут быть адекватно поняты в культурах Востока. И, соответственно, наоборот. Во избежание прагматических несоответствий при работе с поликультурной аудиторией следует использовать как можно более нейтральную символику. Например, знаки «V» или «OK», изображенные пальцами имеют негативную, а нередко, и антисоциальную коннотацию в ряде культур (страны Латинской Америки). В США очень часто

пользуются красным флажком как обозначением нового *e-mail*'а, этот символ абсолютно не понятен на Востоке. Сюда же можно отнести и изображения животных (в Индии корова – священное животное, в мусульманских культурах свинья – «грязное животное»). Поэтому, создавая интернациональный ресурс, лучше, по возможности, свести к минимуму количество символов и иконок. Другим решением проблемы может являться создание адаптированного интерфейса с предварительным выбором культурной принадлежности и последующим предоставлением информации в адаптированном к конкретной культуре интерфейсе. В данном случае стоит добавить в интерфейс специфические знаки, что повысит уровень комфорта пользователя.

Что касается цветовой гаммы, то, в первую очередь, необходимо учитывать принцип функционального, физиологического и эмоционального соответствия, а лишь затем фактор культурной специфики. Если говорить о цветовых решениях, то можно отметить, что в большинстве европейских электронных ресурсов используются серые, коричневые оттенки, в то время как на азиатских ресурсах преобладают оттенки красного. Однако символика цвета может быть интерпретирована так же неоднозначно в зависимости от культурной принадлежности, соответственно, чтобы избежать нежелательной дополнительной смысловой нагрузки данный фактор должен быть проанализирован и учтен в соответствии с культурным контекстом и анализом целевой аудитории. Простой пример: белый цвет обозначает траур в Японии, желтый – траур в Латиноамериканских культурах и т.д.

С точки зрения **навигационных параметров**: в выбранном нами контексте имеет значение расположение меню, каким образом размещается текст в пространстве в зависимости от культурной принадлежности. Например, арабский и израильский текст читается справа налево; до сих пор можно встретить вертикальные надписи или целые статьи на японском и китайском. Имеют значение также параметры доступа на отдельные страницы. В некоторых культурах необходимо иметь право доступа для просмотра определенной информации, в то время как в других культурах информация может быть доступна всем. Существуют различия и во взаимодействии с пользователем: ошибки и указания по навигации могут быть указаны в жесткой форме, а могут сопровождаться вежливыми комментариями, из-за чего что-то могло пойти не так, и указаниями, как это исправить.

Специфика наполнения сайтов контентом тоже может быть дифференцированной в зависимости от культурных факторов. В культурах с высоким индексом индивидуализма, например, большинство образовательных ресурсов используют метод открытого контента, что не характерно для коллективистских культур.

Если брать во внимание такой критерий Хофстеде, как дистанцированность от власти, то на образовательных ресурсах культур с низкой дистанцией власти меню часто организовано достаточно просто. Высокодистантные культуры же, зачастую, будут иметь прямо противоположные характеристики. В них, как правило, интерфейс спроектирован таким образом, что доступ к информации часто бывает очень сложен, более того, часто можно заметить большую иерархичность в организации информации и специальные социальные роли для контролирования доступа к ней.

Соответственно, если принимать во внимание электронную форму подачи информационной составляющей сетевого ресурса, становятся актуальными вопросы, связанные с графическим оформлением как интерфейса так и контента веб-ресурса.

В данном контексте необходим учет структурной специфики организации информации на сетевых ресурсах образовательного характера в разных культурах.

Для образовательных ресурсов индивидуалистских культур, как правило, характерна простота изложения, ясность метафор, навигационное меню организовано так, чтобы не дать пользователю заблудиться и т.д. Как уже говорилось выше, на ресурсах данного формата присутствуют возможности пополнения контента самими пользователями.

Что касается меню, то у российских и азиатских образовательных ресурсов оно, как правило, вертикальное, у западных – горизонтальное. Здесь подразумевается логика построения сайта: например, большинство американских ресурсов занимают всю ширину страницы, в Арабских странах наблюдается структура «сверху-вниз».

Сравнивая дизайны сайтов разных культур также можно заметить, что, например, азиаты очень любят «всплывающие окна», использование которых крайне редко замечено на ресурсах северной Европы. Так, китайские сайты часто содержат по несколько синхронных анимаций, наложений, слайдеров. Это можно объяснить во многом тем, что всплывающие окна появляются не сразу, а возни-

кают через некоторое время, что нехарактерно для низкоконтекстных западных культур, поэтому, зачастую, раздражает и отвлекает внимание их представителей, и, напротив, положительно воспринимается азиатскими пользователями. По нашим наблюдениям, представители азиатских культур не любят печатать текст, они предпочитают «кликать» на ссылки, поэтому их сайты зачастую перегружены ссылками, картинками и другими элементами интерфейса, позволяющими избежать набора текста. Такое разнообразие на экране нередко приводит в шок представителей низкоконтекстных культур.

Согласно нашим наблюдениям, можно сделать вывод о том, что вебсайты культур высокого контекста содержат больше картинок и меньше текста, чем сайты культур низкого контекста.

Что касается специфики структуры страницы и шрифтовых решений, то небезынтересно будет отметить следующую разницу: в англоязычных текстах применяется так называемый флаговый набор (текст выравнивается по левому краю, а правый остается «рваным»). Абзацы текста отделяются друг от друга вертикальными отступами. В России традиционным является набор с выравниванием по ширине колонки набора и с вертикальными отступами между абзацами, более известными как «красная строка». При разработке поликультурного интерфейса необходимо учесть последствия смены стандартного отображения дат, времени, валюты и другой служебной информации на локальное. К примеру, в США даты традиционно отображаются в формате YYYY/MM/DD или MM/DD/YYYY, в то время как большинстве европейских стран принято представление DD/MM/YYYY. Таким образом, отображение даты не в том формате может вызвать недопонимание, особенно когда числа месяца меньше 12-ти (12.04.2012 может означать как 12 апреля, так и 4 декабря). Необходимо использовать кодировки и шрифты, позволяющие использование обозначений местных валют (£, \$, ¥, € и более специфичные, такие как ₪) и так далее. Использование специфических знаков для разных стран может быть значительно облегчено использованием CSS3: если раньше приходилось использовать только самые распространенные шрифты, а надписи с редкими знаками приходилось заменять на изображения, то теперь появилась возможность подгружать для веб-документа собственные шрифты. И уж совсем в прошлом остались те времена, когда для корректного отображения веб-ресурса пользователю приходилось самостоятельно устанавливать необходимые шрифты.

4. Учет культурно-прагматических ожиданий пользователей при проектировании интерфейсов

В дизайне веб-ресурсов, безусловно, воплощается национальная культура его создателей. По наблюдению авторов, для дизайна европейских сайтов характерно удобство навигации, логика и предсказуемость, дозированность информации, отсутствие скрытого контента. Для максимальной дидактической эффективности веб-интерфейс должен соответствовать культурно-прагматическим ожиданиям обучающихся (навигация сайта, графика и контент). Согласно наблюдениям Р. Зальцман [2], в современном кросс-культурном информационном пространстве существует тенденция перенесения западной веб-культуры в веб-пространство Востока, и восточная веб-культура во многом начала адаптироваться к западной, уйдя от использования сверх-контента (специфичной цветовой гаммы и иллюстраций).

В процессе создания и мониторинга образовательных ресурсов, нацеленных на поликультурную аудиторию в обязательном порядке должен быть учтен прагматический фактор. Существуют когнитивные модели восприятия и переработки информации, характерные для восточных (склонных к диалектике) культур и для западных (склонных к формальной логике) культур. На Западе, чаще всего, восприятие информации происходит непосредственного через призму индивидуального восприятия личности и накладывается на имеющийся информационный бекграунд. На Востоке часто при потреблении информационного контента пользователи просто собирают информацию без индивидуального критического осмысления.

Кросс-культурная проблема на настоящий момент является одной из основных, которые необходимо учитывать при создании открытых образовательных ресурсов, ориентированных на представителей разных культур.

Соответственно, если принимать во внимание электронную форму подачи информационной составляющей образовательного ресурса, становятся актуальными вопросы, связанные с графическим оформлением как интерфейса, так и контента веб-ресурса.

Существуют различия между представителями разных культур – они обращают внимание на разные объекты на образовательном вебсайте, считая важными разные виды информации. Более того,

по мнению Д. Мацумото [4], представители разных культурных групп могут применять различные стратегии работы с информацией. Например, в культурах с высокой степенью избегания неопределенности чаще проявляется тенденция принимать решения, основанные на репрезентативности. Это тенденции касаются и взаимодействия пользователя с интерфейсом. То есть, в дополнение к теоретической информации нужны графические приложения в форме презентаций, макетов и т.д. Представители высококонтекстных культур чаще всего предпочитают описательное фактическому, в то время как жители стран с высоким индексом избегания неопределенности стремятся получить исчерпывающую информацию. То есть должны быть предложены различные источники информации: описательный текст, статистика, картинки, видеоролики (также в различном стиле) и т.п.

5. Заключение

Таким образом, можно утверждать, что практические принципы эргономичного дизайна электронного образовательного ресурса должны поддер-

живаться с помощью походов, разрабатываемых в области семиотики, культурологии, психологии, педагогики и дидактики выявляющими законы сочетания содержания текста, изображения, семантики шрифтов и способов выделения компонентов текста и его восприятия обучающимся. Более того, одним из важных компонентов, которые необходимо учитывать, будет специфика цветовосприятия в разных культурах и социокультурный подход к созданию интерфейсов образовательных сетевых ресурсов, а также специфика деятельности и работы с информацией.

Использование современных технологий, таких как CSS3 и HTML5, открывает широкие возможности по практической реализации обозначенных рекомендаций по проектированию адекватного эргономического дизайна для поликультурной аудитории. По результатам дальнейших исследований предполагается сформировать сводную таблицу соотношений принадлежности к определенной культурной группе и соответствующих способов подачи контента, а также необходимости использования конкретных элементов пользовательского интерфейса. ■

Литература

1. Агеев В. Семиотика. — М.: Весь мир, 2002.
2. Зальцман Р. Транснациональное дистанционное образование: кооперация и (или) конкуренция? Международная конференция «Информационно-телекоммуникационные технологии в образовании — 2003, Москва. — № 6. — С. 149-153.
3. Лурия А.Р. Речь и интеллект в развитии ребенка. — М., 1927.
4. Мацумото Д. Психология и культура. Современные исследования. — СПб.: Прайм-Еврознак, 2002.
5. Сугак Е. Эргономические аспекты проектирования пользовательского интерфейса // Автореферат диссертации на соискание ученой степени кандидата психологических наук. — М., МГУ, 2011.
6. Узилевский Г.Я. Начала эргономической семиотики. — Орел: ОРАГС, 2000. .
7. Холодная М.А. Когнитивные стили. О природе индивидуального ума. 2-е изд. — СПб.: Питер, 2004.
8. Lewis R. When Cultures Collide. Managing successfully across cultures. — London. 1996.
9. Nisbett R.E. The geography of thought. — Free Press, 2003.
10. Hofstede G. Culture's consequences, international differences in work related values. — Sage Publications, 1980.

РАСПРЕДЕЛЕННАЯ ЭВОЛЮЦИОННАЯ СЕТЬ ДЛЯ РЕШЕНИЯ МНОГОКРИТЕРИАЛЬНЫХ ОПТИМИЗАЦИОННЫХ ЗАДАЧ В СИСТЕМАХ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ

М.А. Хивинцев,
аспирант кафедры бизнес-аналитики Национального
исследовательского университета «Высшая школа экономики»

А.С. Акопов,
доктор технических наук, профессор кафедры бизнес-аналитики
Национального исследовательского университета «Высшая школа экономики»

E-mail: mkhivintsev@hse.ru, aakopov@hse.ru
Адрес: г. Москва, ул. Кирпичная, д. 33/5

В статье представлен новый подход к решению многокритериальных оптимизационных задач большой размерности, реализуемых, в частности, в системах имитационного моделирования класса AnyLogic с помощью распределенных вычислений. Предложена новая концепция построения распределенной эволюционной сети, основанная на разбиении пространства искомым переменных на кластеры и присвоения каждому вычислительному элементу сети своего кластера, по которому осуществляется поиск промежуточных результатов с помощью взаимодействующих генетических алгоритмов.

Ключевые слова: имитационное моделирование, генетические алгоритмы, распределенные вычисления, многокритериальная оптимизация.

1. Введение

Существует известная проблема поиска субоптимальных решений в имитационных моделях «больших систем», таких как, система управления технологическими режимами сложного производства, система оптимального управления финансовыми и материальными потоками крупной корпорации, система управления

инвестиционным портфелем и др. [1]. Поиск оптимальных решений в подобных системах должен осуществляться по множеству, как правило, конкурентных и разномасштабных критериев со сложными ограничениями.

Проблема заключается в том, что оптимизация сразу по множеству критериев требует значительных временных и вычислительных ресурсов. При

этом, так как целевые функции многих переменных (более 10) порождают многомерное пространство поиска решений, вычислительная сложность оптимизационной задачи, как правило, становится преградой к ее решению за допустимое время. Аналитическое описание и строгое решение такой задачи на практике, как правило, не осуществимо.

При решении такого класса задач хорошо зарекомендовали себя алгоритмы метаэвристического класса, позволяющие, не исследуя полностью область допустимых решений, с определенной вероятностью находить глобальный экстремум в однокритериальных задачах и фронт Парето в многокритериальных.

Одним из распространенных классов метаэвристических алгоритмов являются генетические алгоритмы (ГА), теория которых была разработана J. H. Holland в 1975 г. [2]. Доказано, что ГА обладают значительно меньшей вычислительной емкостью, чем простой метод перебора, и рядом плюсов по сравнению с другими способами решения оптимизационных задач, однако всегда имеют риск схождения к локальному оптимуму вместо глобального [3 – 5].

Следует отметить, что имеется ряд работ в области проектирования параллельных генетических алгоритмов [6 – 11]. Первой работой в этой области можно считать работу Grosso P.V. [6] выполненной еще в 1985 г.

Одним из основных преимуществ предлагаемой распределенной эволюционной сети является возможность разбиения процесса решения одной задачи на несколько параллельных вычислительных подпроцессов, объединенных в распределенную сеть.

В качестве средства имитационного моделирования (ИМ) в данной работе выбран программный продукт AnyLogic, так как это один из немногих инструментов, который поддерживает все основные методы ИМ: процессно-ориентированный, системно-динамический и агентный, а также любую их комбинацию. Отметим, что в программном продукте AnyLogic реализован однокритериальный оптимизационный алгоритм OptQuest (<http://www.opttek.com/OptQuest>), основанный на метаэвристиках рассеянного поиска (scatter search) и поиска «табу» (tabu search). Однако хорошо известно, что заложенные в основе OptQuest алгоритмы неэффективны при размерности оптимизационной задачи более 10 переменных. Кроме того, в AnyLogic отсутствуют решения для многокритериальной оптимизации, нахождения и визуализации фронта

Парето, а также распараллеливания вычислений.

Итак, практически все современные системы имитационного моделирования (в том числе, AnyLogic, Powersim и др.) не имеют в своем арсенале достаточно эффективных средств решения многокритериальных оптимизационных задач большой размерности. Поэтому построение РЭС является весьма актуальной задачей.

Целью данной статьи является разработка распределенной эволюционной сети, предназначенной для решения многокритериальных оптимизационных задач большой размерности. Представлен новый подход к построению комплексного программного решения по интеграции распределенной эволюционной сети с системой имитационного моделирования AnyLogic.

2. Концепция распределенной эволюционной сети

Итак, для ускорения поиска решений в ГА при распараллеливании вычислительных процессов необходимо добиться снижения суммарного числа расчетов функций приспособленности. Для этого предлагается наиболее действенный способ — сокращение размера популяции, с которой будет работать ГА на каждом из процессов. В этом случае на всех шагах эволюции в результате операции скрещивания будет создаваться меньшее число потомков и, следовательно, для меньшего их количества будут рассчитываться функции приспособленности.

Важной задачей при сокращении размера популяции является обеспечение высокой вероятности схождения алгоритма к глобальному экстремуму. Это выполнимо только при сохранении оптимального соотношения между длиной хромосомы и размером популяции. Поэтому сокращение размера популяции предлагается достичь за счет уменьшения длины хромосомы, разбив пространство переменных на кластеры и присвоив каждому вычислительному элементу сети свой кластер, по которому он будет искать промежуточные результаты. Этот принцип заложен в новую предлагаемую методику эффективного решения многокритериальных оптимизационных задач с функциями многих переменных, используя *распределенную эволюционную сеть (РЭС, или distributed evolutionary network - DEN)*.

РЭС включает в себя $N+1$ вычислительных элементов (процессоров). На одном из них выполняется управляющий процесс, на N других - вычислительные. В основу механизма управления РЭС лег

алгоритм управления распределенной эволюционной сетью (АУРЭС, или *distributed evolutionary network algorithm - DENA*), а за базу для каждого из N вычислительных процессов - распространенный метод решения многокритериальных задач путем построения фронта Парето - SPEA2 (*Strength Pareto Evolutionary Algorithm*) [12; 13]. Каждый из вычислительных процессов будет выполнять оптимизацию по своему кластеру переменных, что позволит выполнять ГА с меньшим размером популяции, и, как следствие, снизить время на поиск решения для каждого из процессов. В этом заключается основное отличие от классической островной модели в теории ГА [11]. Все острова работают с общим набором переменных и хромосомами равной длины.

Другим отличием РЭС от островной модели является наличие центрального элемента управления, который управляет ходом эволюции всей системы и использует архив решений, полученных на прошлых стадиях эволюции. Согласно АУРЭС с определенной периодичностью управляющий процесс аккумулирует промежуточные результаты каждого из вычислительных процессов, сопоставляет их с архивными решениями, обрабатывает и перераспределяет генетический материал между вычислительными процессами, попутно давая им команду к запуску следующей итерации эволюционного процесса. После этого вычислительные процессы начинают асинхронно следовать ГА поиска Парето-оптимальных решений согласно методу SPEA2 (*Strength Pareto Evolutionary Algorithm*) [12; 13]. Когда они заканчивают очередную итерацию, то передают результаты управляющему процессу и ожидают от него нового генетического материала, возникшего в результате обработки результатов работы ГА всех параллельных процессов – через управляющий процесс происходит обмен генетическими материалами между вычислительными процессами. С каждым из описанных эволюционных циклов система движется к фронту Парето наивысшего ранга. Схема функционирования РЭС представлена на рис. 1.

Основной функцией РЭС является решение многокритериальных оптимизационных задач в ИМ. Для ее выполнения РЭС имеет установленную связь с ИМ.

На вход имитационной модели от РЭС поступают новые значения для переменных, для которых по установленным модельным связям должны быть рассчитаны значения целевых функций, которые затем передаются из имитационной модели в РЭС. Таким образом, оцениваются значения функций приспособленности ГА, выполняемых на каждом из вычислительных процессов РЭС. Далее распределенные ГА формируют новые (лучшие) значения для переменных имитационной модели, и осуществляется их очередное присваивание искомым переменным с последующим прогоном модели для оценки фитнес-функций.

3. Механизм управления распределенной эволюционной сетью

К сожалению, классическая островная модель ГА [11] не может существенно снизить потребность в большом размере популяции, так как длина хромосомы остается неизменной, а в границах каждого острова для оптимального хода эволюционного процесса пропорция между длиной хромосомы и числом членов популяции должна также сохраняться, в противном случае на локальном уровне каждого из островов будет значительно меньше генетического материала, что повысит вероятность схождения к локальным, а не глобальным, оптимумам. Если на каждом из островов останется столько же членов популяции, сколько и при одном вычислительном процессе, то время поиска решений в обоих случаях будет практически одинаковым.

Чтобы добиться действенного снижения эффективного размера популяции, в рамках предлагаемой методики решения рассматриваемого класса задач предлагается новый алгоритм управления РЭС (АУРЭС). Разберем его по стадиям.

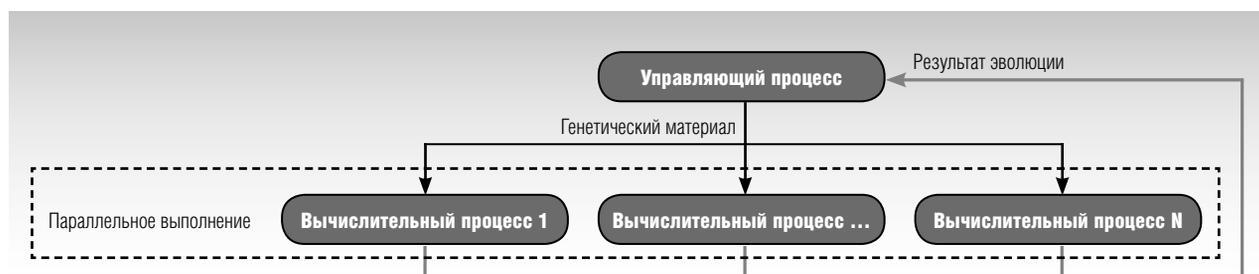


Рис. 1. Схема функционирования распределенной эволюционной сети

Стадия подготовительная.

Прежде чем приступать к решению задачи в РЭС необходимо провести статистический анализ переменных, построить их корреляционную матрицу и выявить, как непрямые связи между ними влияют на конечный результат. В конечном счете, нужно объединить переменные в кластеры. В каждом кластере будут находиться те переменные, изменение которых вызывает наиболее явную потребность в изменении других для достижения наилучшего результата.

Каждому из вычислительных процессов нужно определить свой кластер переменных, поэтому число кластеров должно совпадать с количеством вычислительных процессов.

При разбиении переменных на кластеры нужно стремиться к тому, чтобы размеры кластеров были схожими, так как в этом случае для каждого из вычислительных процессов можно ожидать примерно равную длину хромосом, а, значит и вычислительная мощность распределится равномерно по процессам.

Стадия исходная.

На исходной стадии управляющий процесс осуществляет формирование популяции случайным образом. Размер общей популяции равен сумме эффективных размеров популяций для каждого из вычислительных процессов, зависящий от длины хромосомы и настроек ГА.

Далее управляющий процесс делит популяцию на N подпопуляций согласно алгоритму, поддерживающему разнообразие генетического материала в подпопуляциях на одинаковых аллелях, и передает особей из одной подпопуляции на один вычислительный процесс.

Эффективный размер подпопуляции для каждого из вычислительных процессов необходимо выбрать исходя из числа переменных его кластера, определяющего длину хромосомы. После распределения подпопуляций управляющий процесс передает вычислительным процессам персональные настройки и команду к запуску ГА. Далее следует переход к первой стадии.

Стадия 1.

С первой стадии стартует первая итерация поиска решений. Вычислительные процессы работают параллельно и асинхронно. Поэтому итерации, равно как и стадии 1-4, выполняются независимо для каждого из вычислительных процессов.

На первой стадии происходит первичное вычисление функций приспособленности для каждой

из особой подпопуляции на всех вычислительных процессах.

Стадия 2.

Вторая стадия является наиболее ресурсозатратной. В ходе нее каждый из вычислительных процессов реализует построение фронта Парето высшего ранга только по тому набору переменных, которые включены в его кластер. Остальные переменные принимают фиксированные значения и остаются неизменными на протяжении всей второй стадии. В результате сокращается фактическая длина хромосомы, так как в ней появляется «неактивная» часть, не участвующая в эволюции и соответствующих вычислительных процессах.

Нахождение фронта Парето осуществляется при помощи одного из распространенных эффективных методов на основе ГА – SPEA2. На рис. 2 показаны исходные подпопуляции и конечный результат построения фронта Парето, варьируя только выделенные переменные у каждого из процессов.

В момент достижения критерия остановки любым из вычислительных процессов, только этот процесс переходит к Стадии 3. Таким образом, каждый из процессов независимо от других переходит к следующей стадии.

Стадия 3.

Вычислительный процесс передает результат второй стадии управляющему процессу. Далее в рамках управляющего процесса обеспечивается сравнение полученных на второй стадии результатов с имеющимися архивными результатами. При этом управляющий процесс опрашивает остальные вычислительные процессы для обновления результа-

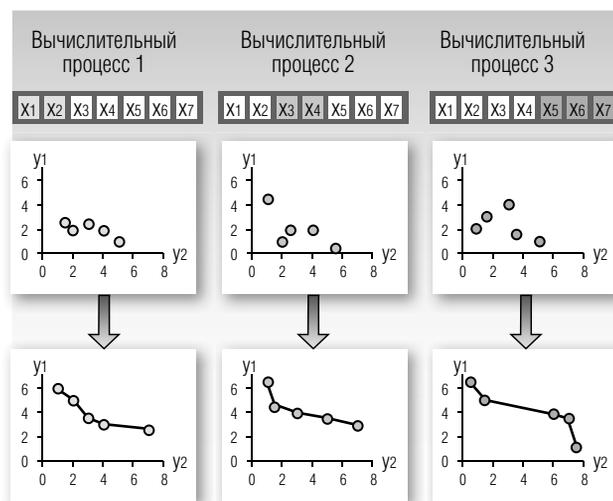


Рис. 2. Процесс построения фронта Парето на каждом из вычислительных процессов

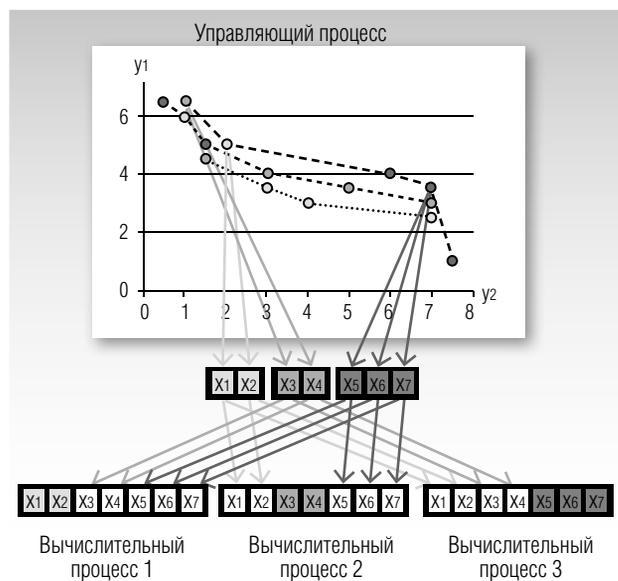


Рис. 3. Перестроение фронтов Парето нескольких рангов управляющим процессом на основе имеющихся у него данных

тов в архиве. В результате управляющий процесс из всех известных ему решений строит фронт Парето нескольких рангов. Может оказаться, что фронт высшего ранга будет состоять из решений нескольких процессов, как это показано на рис. 3.

Стадия 4:

Для каждого из кластеров переменных случайным образом выбирается по одному набору решений (по одной точке на границе Парето), но вероятность выбора того решения выше, который находится на фронте Парето более высокого ранга.

У процесса, инициировавшего обмен генетическим материалом с управляющим процессом, происходит изменение неактивной части хромосомы – тех переменных, которые были фиксированы в ходе второй стадии. Эти переменные принимают значения, соответствующие значениям этих переменных из случайно выбранного набора решений (рис. 4).

На этом завершается очередная итерация работы АУРЭС и, если не достигнут критерий остановки вычислительной процедуры, то происходит переход к следующей итерации поиска решений на глобальном уровне и возврат инициировавшего обмен вычислительного процесса к стадии 1, но с новыми значениями переменных. Стадии 1-4 повторяются до тех пор, пока управляющий процесс в течение определенного числа итераций не обнаружит улучшение хотя бы одной из точек фронта Парето наивысшего ранга.

4. Пример использования распределенной эволюционной сети, интегрированной с моделью AnyLogic

Спроектированная РЭС позволяет решать многокритериальные задачи в имитационных моделях высокого уровня сложности. Данный класс задач характеризуется несколькими свойствами:

1. имитационная модель имеет сложные внутренние нелинейные зависимости между переменными;
2. требуется обеспечить решение многокритериальной оптимизационной задачи в системе AnyLogic;
3. оптимизация производится в пространстве большой размерности (10 и более искомым переменных, несколько целевых функций);
4. вычисление Парето-оптимальных решений должно происходить за приемлемое время.

Для демонстрационного примера решения задачи подобного класса далее рассматривается имитационная модель управления производственной компанией, занимающейся производством кабельно-проводниковой продукции и обладающей типичными для своей деятельности бизнес-процессами.

Основные параметры компании, которые были заложены в имитационную модель:

- ◆ Номенклатура включает в себя 24 вида изделий;
- ◆ 2 производственных цеха;

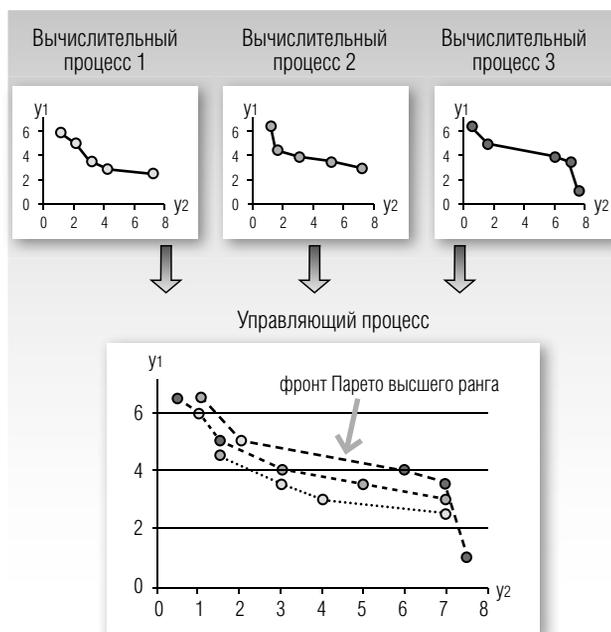


Рис. 4. Присваивание неактивным на каждом из вычислительных процессов переменным новых значений

- ◆ 103 вида оборудования;
- ◆ работа строится в 2 смены;
- ◆ 38 единиц сырья;
- ◆ 380 производственных сотрудников;
- ◆ 15 должностных (функциональных) ролей для сотрудников.

Важно отметить, что взаимосвязь между основными объектами модели является «многие-ко-многим» — другими словами, для производства одного вида продукции может использоваться несколько видов сырья и один вид сырья может быть использован для нескольких видов продукции. В результате переменные рассматриваемой модели являются многомерными (имеют от 1 до 3 размерностей), что порождает трехмерное пространство поиска решений и повышает вычислительную сложность оптимизационной задачи.

Имитационная модель предприятия была реализована в системе AnyLogic.

Оптимизационная задача формулируется следующим образом: необходимо определить цены, обеспечивающие наилучшую динамику продаж для максимизации двух целевых критериев на следующий квартал: объема продаж (в натуральном выражении) и чистой прибыли. Данный набор целей является распространенным в бизнесе и отражает интересы акционеров — иметь не только прибыльную компанию, но и быть лидером по доле рынка. Оба эти критерия оказывают весомое влияние на рыночную стоимость компании.

В результате будет построена граница Парето, которую образуют недоминируемые Парето-решения. Другими словами, остальные решения не позволят добиться одновременно более высоких значений чистой прибыли и объема продаж.

Итак, в рассматриваемой бикритериальной оптимизационной задаче имеется две целевые функции: объем продаж в натуральном выражении и чистая прибыль. Конечной целью максимизации объема продаж является увеличение доли рынка. При этом данная цель является конкурентной по отношению к чистой прибыли.

Целевая функция максимизации объема продаж в натуральном выражении имеет следующий вид:

$$\sum_{t=t_0}^{t_0+T} \sum_{i=1}^I Q_i(p_i(t)) \rightarrow \max_{p_i(t)}, \quad (1)$$

$$Q_i(t) = \frac{Q_i(t-1)}{(p_i(t) / p_i(t-1))^{\xi_i}}. \quad (2)$$

Целевая функция чистой прибыли:

$$\sum_{t=t_0}^{t_0+T} \sum_{i=1}^I p_i(t) Q_i(p_i(t)) - Costs(Q_i(t)) - Expenses(t) \rightarrow \max_{p_i(t)}, \quad (3)$$

где: $t = t_0, t_0+1, \dots, t_0+T$ — время в модели (по месяцам), T — год.

$i = 1, 2, \dots, I$ — индекс товара (по умолчанию $I=24$).

$p_i(t)$ — цена i -ого товара,

$Q_i(t)$ — объем продаж в натуральном выражении,

$Costs(t)$ — себестоимость производства (переменные затраты),

$Expenses(t)$ — суммарные постоянные затраты: коммерческие, управленческие, прочие расходы и налоги.

ξ_i — эластичность объема продаж i -ого товара по соответствующим ценам ($\xi_i \geq 0$).

Следует отметить, что все слагаемые в формуле чистой прибыли имеют связь между собой, установленную через отношения внутри модели. Так, например, заложенные функции спроса, издержек и производства позволяют найти оптимальный баланс между планом продаж и планом производства с учетом производственных ограничений.

В общей сложности модель состоит из 125 переменных размерностью от 1 до 2(цеха)×103(оборудования)×24(продукта) и 159 связями между ними. Однократный просчет модели занимает 0.2 сек на процессоре Intel core i7 с тактовой частотой 2000 МГц.

Встроенного функционала в AnyLogic недостаточно для решения поставленной оптимизационной задачи, поэтому рассматриваемая имитационная модель была интегрирована с разработанной РЭС с использованием технологии IntelliJ IDEA (на языке программирования Java).

Согласно алгоритму управления РЭС на первом этапе 24 искомые переменные были разбиты на 3 кластера в зависимости от используемого оборудования для их производства.

РЭС была составлена из 4 вычислительных элементов: управляющий процессор и 3 вычислительных, обеспечивающих поиск решений по своему кластеру переменных.

В результате, формирование фронта Парето заняло всего 3 минуты. При этом каждый из вычислительных процессоров произвел 720 итераций вычисления целевых функций рассматриваемой имитационной модели. Таким образом, реализованная по предложенной концепции РЭС система обеспечила решение поставленной многокритериальной оптимизационной задачи за допустимое время.

5. Заключение

В данной статье был предложен и реализован подход к решению многокритериальных оптимизационных задач большой размерности в системах имитационного моделирования класса AnyLogic. Такой подход основан на использовании распределенной эволюционной сети (РЭС), в которой эффект от распараллеливания вычислительных процессов выше, чем при классической островной модели. Это достигается благодаря способности предложенного алгоритма к разбиению всего пространства решений на кластеры, каждый из которых передается на свой вычислительный процесс, исполняя генетический алгоритм (ГА) на локальном уровне. В этом случае для ГА, работающего с целевыми функциями из меньшего числа переменных, для качественных решений требуется меньший размер популяции и, как следствие, со-

кращению числа вызовов функций приспособленности, что ведет к значительному снижению времени схождения алгоритма к конечному решению.

При координации управляющего процесса с определенной периодичностью происходит обмен генетическим материалом между процессами. Это обеспечивает возможность формирования фронта Парето наивысшего ранга за допустимое время, что свидетельствует о высокой эффективности данной методики решения.

Разработан платформенно-независимый программный комплекс на основе предложенной концепции РЭС на языке программирования Java, позволивший интегрировать РЭС и имитационные модели, разработанные в средстве имитационного моделирования AnyLogic.

Функциональность РЭС была апробирована на демонстрационном примере. ■

Литература

1. Akopov A.S. Designing of integrated system-dynamics models for an oil company // International Journal of Computer Applications in Technology. – 2012. – Vol. 45, No. 4. – P. 220-230.
2. Holland J.H. Adaptation in natural and artificial systems. – Ann Arbor: University of Michigan Press, 1975.
3. Hajena P., Lin C.-Y. Genetic search strategies in multicriterion optimization design // Structural Optimization. – June 1992. – Vol. 4. – New York: Springer, P. 99-107.
4. Емельянов В.В., Курейчик В.В., Курейчик В.М. Теория и практика эволюционного моделирования. – М.: Физматлит, 2003.
5. Курейчик В.М., Лебедев Б.К., Лебедев О.К. Поисковая адаптация: теория и практика. – М.: Физматлит, 2006.
6. Grosso P.B. Computer simulations of genetic adaptation: Parallel subcomponent interaction in a multilocus model // Unpublished doctoral dissertation. – The University of Michigan (University Microfilms №8520908), 1985.
7. Alba E., Troya J.M. A survey of parallel distributed genetic algorithms // Complexity. – 1999. – Vol.4, № 4. – P. 31-52.
8. Alba E., Troya J.M. Influence of the migration policy in parallel distributed GAs with structured and panmictic populations // Application Intelligence. – 2000. – Vol.12, №3. – P. 163-181.
9. Alba E., Troya J.M. Analyzing synchronous and asynchronous parallel distributed genetic algorithms // Future Generation Computer Systems. – 2001. – Vol.17, №4. – P. 451-465.
10. Golub M., Jakobovic D. A new model of global parallel genetic algorithm // Proceedings of 22nd International Conference on Information Technology Interfaces IVI. – 2000. – P. 363-368.
11. Silva J.D., Simoni, P.O. The island model parallel GA and uncertainty reasoning in the correspondence problem // Proceedings of International Joint Conference on Neural Networks (IJCNN '01). – 2001. – Vol. 3. – P. 2247-2252.
12. Zitzler E., Thiele L. Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach // IEEE Transactions on Evolutionary Computation. – 1999. – Vol. 3, No. 4. – P. 257-271.
13. Bleuer S., Brack M., Thiele L., Zitzler E. Multiobjective genetic programming: Reducing bloat by using SPEA 2 // Proceedings of the 2001 Congress on Evolutionary Computation (CES-2001). – 2001. – P. 536-543.

МЕТОД ВЗВЕШЕННОЙ СУММЫ КРИТЕРИЕВ В АНАЛИЗЕ МНОГОКРИТЕРИАЛЬНЫХ РЕШЕНИЙ: PRO ET CONTRA

В.В. Подиновский,

доктор технических наук, профессор кафедры кафедры высшей математики
на факультете экономики Национального исследовательского
университета «Высшая школа экономики»

М.А. Потанов,

кандидат физико-математических наук, ведущий научный сотрудник
Института автоматизации проектирования РАН

E-mail: podinovski@mail.ru, potanov@icad.org.ru

Адрес: г. Москва, ул. Мясницкая, 20

Дан обзор приведенных в научной литературе результатов анализа метода взвешенной суммы критериев (МВСК), широко применяемого для решения многокритериальных задач.

Ключевые слова: многокритериальные задачи принятия решений, взвешенная сумма критериев, шкалы критериев, веса критериев, важность критериев, нормализация критериев.

1. Введение

В подавляющем большинстве случаев задачи принятия решений оказываются *многокритериальными*: варианты решения оцениваются при помощи (частных) критериев f_1, \dots, f_m , $m \geq 2$. Поскольку, как правило, каждый из критериев выделяет «свой» наилучший вариант, т.е. не бывает варианта, который одновременно является лучшим по каждому из критериев, то многокритериальные задачи принципиально сложнее однокритериальных (когда $m = 1$) и требуют для своего решения специальных методов и подходов.

Самым распространенным, давно известным и до сих пор чаще других используемым является ме-

тод, основанный на свертывании всех критериев в один-единственный обобщенный (глобальный, интегральный, агрегированный, составной, комплексный, синтетический, компромиссный, ...) критерий F , представляющий собой сумму критериев, взвешенных коэффициентами их относительной важности, или весами. Этот метод называется методом взвешенной суммы критериев (далее – МВСК), на английском языке – weighted sum method (WSM).

МВСК – давно известный и активно применяемый метод. Широкая его распространенность вызвана целым рядом причин. Одни из них обусловлены *привлекательными достоинствами* МВСК:

♦ метод представляется *простым и понятным*,

♦ он удобен для расчетов,

♦ применим для решения задач принятия решений в разных постановках: выбрать один наилучший или несколько лучших вариантов, упорядочить (ранжировать) все варианты по предпочтительности, и т.п.

Другие причины иного рода и состоят в следующем.

♦ Пользователи метода обычно не имеют представления о современной теории принятия решений при многих критериях и, в частности, о результатах научных исследований МВСК.

♦ Корректные (теоретически обоснованные) методы решения многокритериальных задач требуют достаточно серьезной работы по получению и обработке информации о предпочтениях лица, принимающего решение (ЛПР), значительных затрат сил и времени (по сравнению с МВСК).

♦ Анализ многокритериальных задач проводится на основе информации о предпочтениях ЛПР, и поэтому у таких задач нет объективных решений, с которыми, как с эталонами, можно было бы сравнить результаты, полученные методом взвешенной суммы.

♦ Многокритериальные задачи весьма разнообразны, и не для всех видов таких задач разработаны корректные и эффективные методы их решения.

Цель статьи — представить всесторонний обзор результатов критического анализа МВСК, которые «рассыпаны» по целому ряду научных публикаций (см., например, [1 – 14]).

2. Очевидные недостатки МВСК

Ряд недостатков МВСК известен всем, кто его вдумчиво анализировал или пользовался им:

• При использовании взвешенной суммы критериев приходится оперировать с ее значениями, которые обычно лишены содержательного («физического») смысла. Это затрудняет пояснения и обоснования рекомендаций, полученных в результате решения многокритериальных задач. Для иллюстрации [7] предположим, что Вы — лицо, принимающее решение (ЛПР). После анализа сложной и ответственной задачи со многими критериями консультант принес Вам результаты, согласно которым из трех возможных вариантов решения он в качестве оптимального предлагает выбрать первый, так как значение обобщенного показателя для него равно, скажем, 0,76; для двух остальных вариантов его значения меньше — соответственно 0,69 и 0,58.

На просьбу пояснить смысл этих чисел консультант ответил, что они получены в результате расчетов и являются значениями взвешенной (при помощи назначенных Вами коэффициентов важности) суммы нормированных значений критериев для этих вариантов. Насколько убедителен будет для Вас представленный довод в пользу первого варианта? И как Вы будете защищать указанный выбор, если, например, это попросит сделать комиссия или потребует общественность?

• Использование постоянных весов или коэффициентов важности означает, что соотношение критериев по важности одно и то же при любых значениях критериев. Однако нередко случается так, что даже упорядоченность рассматриваемых двух критериев по важности зависит от того, какие значения приняли другие критерии [7]. Так, выбирая загородный дом, покупатель или арендатор может полагать, что при наличии на участке бассейна близость леса важнее близости «воды» для купания (пруда, озера или речки), а при отсутствии бассейна, наоборот, что близость «воды» важнее близости леса. Что уж говорить о постоянстве величин весов критериев. Можно было бы, конечно, ввести в рассмотрение и переменные веса. Но это был бы уже совсем другой метод.

3. Взвешенная сумма критериев как математическая модель предпочтений

Перейдем к научному анализу МВСК. Прежде всего, введем исходные определения и формулы. Пусть Z_i — множество значений, или градаций критерия f_i (его «шкала»). Далее для определенности и без ограничения общности будем считать, что с ростом значения каждого критерия предпочтения возрастают (как иногда говорят, критерии желательнее максимизировать, или, что критерии ориентированы положительно). Каждый вариант (альтернатива, план, стратегия, ...) x характеризуется значениями критериев $y_1 = f_1(x), \dots, y_m = f_m(x)$, составляющими векторную, или критериальную оценку этого варианта $y = (y_1, \dots, y_m)$. Обозначим через X множество всех вариантов.

Сравнение вариантов по предпочтительности осуществляется на основе сопоставления их векторных оценок. Понятно, что если значения одних критериев больше для первого из сравниваемых вариантов, а значения других критериев — для второго, то возникают очевидные трудности (вспомним шуточную дилемму: что лучше — быть бедным

и здоровым или богатым, но больным?). Для их преодоления согласно МВСК векторный критерий «свертывается» – вводится в рассмотрение взвешенная сумма критериев (которые при необходимости вначале нормализуются – см. раздел 5):

$$F(f|w) = w_1 f_1 + \dots + w_m f_m, \quad (1)$$

где положительные числа w_i (обычно в сумме равные 1) – это коэффициенты важности, или веса, предназначенные для учета, вообще говоря, разной относительной важности, весомости, значимости критериев. После назначения величин этих коэффициентов каждый вариант x с использованием формулы (1) характеризуется одним числом – значением взвешенной суммы критериев

$$F(f(x)|w) = w_1 f_1(x) + \dots + w_m f_m(x). \quad (2)$$

Вариант считается тем более предпочтительным, чем больше соответствующее ему значение взвешенной суммы (2). Поэтому, например, оптимальным считается тот вариант, для которого эта сумма принимает наибольшее значение.

Подчеркнем, что функция (1) является аддитивной функцией ценности специального вида (см. раздел 5). Поэтому для ее применения должны выполняться условия, гарантирующие ее существование [15]. В частности, должно выполняться необходимое условие взаимонезависимости критериев по предпочтению. Однако *в практике использования МВСК проверка таких условий обычно не только не делается, но о них даже и не вспоминают.*

4. Взвешенная сумма однородных критериев

Вначале будем рассматривать наиболее простой случай, когда все критерии изначально имеют общую шкалу: $Z_1 = \dots = Z_m = Z_0$. Предположим, что эта шкала является балльной, причем «цена балла» для всех критериев одна и та же (самый известный пример – школьные оценки успеваемости учеников).

Для простоты сначала будем полагать, что все критерии одинаково важны, так что $w_1 = \dots = w_m = 1/m$. При этом взвешенная сумма (2) превращается во всем известный «средний балл»:

$$F(f(x)|w) = \frac{1}{m}(f_1(x) + \dots + f_m(x)). \quad (3)$$

❖ *Необоснованное рассмотрение общей шкалы критериев как количественной.* Формула (3), как и все последующие, предполагает выполнение арифме-

тических операций со значениями критериев. Согласно математической теории измерений [15], это допустимо только в том случае, когда критерии имеют количественную шкалу (для (3) шкала критериев должна быть не менее совершенной, чем шкала интервалов). Только при выполнении указанного условия могут быть адекватными (осмысленными) утверждения о соотношении величин функции (3), например, что если $F(f(x')|w) > F(f(x'')|w)$, то вариант x' предпочтительнее варианта x'' . Подчеркнем, что шкала должна быть количественной именно для характеристики предпочтений.

Однако балльная шкала не является количественной. Еще хуже обстоит дело, когда критерии являются лингвистическими, т.е. градации их общей шкалы описываются словесно и нумеруются в порядке возрастания их предпочтительности, а затем эти номера используются в качестве значений критериев в формуле (3). В этом случае шкала критериев является всего лишь порядковой.

❖ *Необоснованное принятие допущения о равномерности общей шкалы критериев.* Проанализируем формулу (3) как модель предпочтений. Прежде всего, заметим следующую ее особенность: если значение одного из критериев, пусть f_i , уменьшить на некоторую величину δ , а значение другого критерия, пусть f_j , увеличить на ту же величину δ , то величина суммы (3) не изменится. Поэтому уменьшение (ухудшение) значений одних критериев компенсируется таким же в сумме увеличением (улучшением) значений других критериев. И так для всего диапазона шкалы критериев Z_0 . Это означает, помимо всего прочего, что шкала критериев считается равномерной: предпочтения возрастают вдоль нее «с постоянной скоростью». Насколько реально это утверждение о равномерности шкалы? Для ответа на этот вопрос достаточно задуматься над таким примером: при оценке успеваемости учеников по двум предметам применение (3) означает, что одинаковыми по успеваемости будут считаться ученики с такими парами оценок: (3, 3), (2, 4), (4, 2), (1, 5), (5, 1)! Даже если исключить неудовлетворительные оценки, то, имея в виду (3), придется признать, что одинаковыми по успеваемости будут ученики с такими парами оценок: (4, 4), (3, 5), (5, 3). Иначе говоря, согласно (3), улучшение успеваемости (знаний, умений и навыков) при переходе от оценки 3 к оценке 4 такое же, как при переходе от 4 к 5. Поскольку средний балл используется для массовых задач оценивания, например, успеваемости, то с этим недостатком приходится смириться: пока не

придуман другой простой, понятный и эффективный метод для решения таких задач. Но можно ли игнорировать указанный недостаток применительно к сложным и ответственным задачам принятия решений и «экономить при выстреле за счет прицеливания»?

♦ *Необоснованное удаление некоторых вариантов из рассмотрения.* Использование формулы (3) в силу указанной выше возможности компенсации уменьшения худших значений одних критериев увеличением лучших значений других критериев может исключить из числа выбираемых те варианты, которые при внимательном рассмотрении могут быть признаны оптимальными. Вот такой пример [4]. Абитуриент осуществляет выбор одного из трех университетов, одинаковых по качеству обучения, руководствуясь двумя критериями – условиями проживания в общежитии и условиями для занятий спортом. Шкала критериев десятибалльная. Пусть для первого университета значение первого критерия равно 10, а второго – 1; для второго университета значения критериев – это соответственно 1 и 10, а для третьего университета значения обоих критериев равны 5. Таким образом, первый университет имеет отличное общежитие, но очень плохие условия для занятий спортом. У второго университета, наоборот, очень плохое общежитие, но отличные условия для занятий спортом. Общежитие и условия для занятий спортом у третьего университета – удовлетворительные. Согласно формуле (3) средние оценки университетов таковы: 5,5; 5,5; 5. Следовательно, третий университет со сбалансированными условиями проживания и занятий спортом останется без рассмотрения, хотя именно он мог бы оказаться наиболее предпочтительным для абитуриента.

Теоретически установлены условия исключения возможности отбрасывания вариантов, которые могут претендовать на роль лучших (при надлежащих значениях весов w_i). Например, достаточно, чтобы множество вариантов X было выпукло, а все критерии $f_i(x)$ были вогнутыми (в частности, линейными) функциями. Стоит, однако, иметь в виду [12], что методы линейного программирования выделяют лишь вершины многогранного множества ограничений, так что при решении линейных многокритериальных задач при помощи МВСК с использованием компьютерных программ линейного программирования отличные от вершин точки будут потеряны.

Отметим еще следующий пример [2], в котором

рассматриваемый недостаток проявляется особенно выпукло: ухудшение качества изображения телевизора не может быть компенсировано улучшением качества его звука.

♦ *Необоснованное рассмотрение коэффициентов важности как количественных оценок важности критериев.* Перейдем к рассмотрению задач, в которых критерии имеют общую шкалу, но разную важность. Согласно МВСК, это должно быть учтено при помощи коэффициентов важности – положительных чисел, в сумме равных единице. Поскольку формула (2) предусматривает, что с ними производятся арифметические операции, то эти коэффициенты, согласно теории измерений, следует рассматривать как результаты количественного измерения важности (причем в шкале отношений), т.е. должны иметь смысл утверждения типа: «Если $w_1 = 0,4$ и $w_2 = 0,2$, то первый критерий важнее второго в 2 раза». Предложено множество методов оценивания важности критериев, т.е. назначения величин w_i . Но все они страдают общим тяжелым пороком – *точного (формального) определения понятию важности критериев не дается*, и полагается, что человек должен исходить из своего понимания, что такое важность. (Кстати, это замечание касается и предыдущего случая, когда предполагалась равная важность критериев.) Поэтому для оценивания коэффициентов важности человеку предлагается отвечать на вопросы, сводящиеся в итоге к таким: «Во сколько раз критерий f_i важнее критерия f_j ?» или «Какая доля общей (единичной) важности приходится на данный критерий?». Проблема здесь состоит в том, что невозможно установить (а потому и обоснованно формализовать) точный смысл, вкладываемый конкретным человеком в ответы на указанные вопросы [13]. Для пояснения заметим еще, что коэффициенты важности используются в целом ряде методов решения многокритериальных задач, которые либо опираются на обобщенные показатели вида, отличного от (1) (пример – мультипликативный обобщенный критерий

$$F_{\Pi}(f|w) = f_1^{w_1} \cdot \dots \cdot f_m^{w_m},$$

либо используются совершенно иначе (например, в схемах «голосования» в методах ELECTRE [16]). Понятно, что разные способы использования одних и тех же величин коэффициентов важности могут привести к совершенно разным результатам. Однако методы оценивания коэффициентов важности не учитывают этого. Более того, часто ЛПП

(или экспертам) даже не сообщается, какие манипуляции будут осуществляться с этими коэффициентами.

Положение кардинально изменилось с появлением теории важности критериев (ТВК), которая была создана в России и продолжает активно развиваться [7, 17]. Эта теория опирается на точные определения понятия превосходства в важности одного критерия над другим (качественная важности) и понятия превосходства в важности одного критерия над другим во столько-то раз (количественная важности), а для оценки коэффициентов важности предлагает человеку сравнивать по предпочтительности векторные оценки специального вида. Рассмотрение основных положений этой теории выходит за рамки данной статьи.

Интересные результаты анализа МВСК получены в [18]. В частности, вычислительные эксперименты показали, что в задачах с числом критериев от 5 до 9 и числом вариантов от 5 до 9 при неравномерной пятибалльной шкале критериев МВСК может приводить к ошибочным результатам не менее чем в 35 – 40% случаев. Мы, в свою очередь, отметим, что указанные нижние оценки занижены, т.е. число ошибок может быть и большим, так как эти оценки основаны, по сути, на неявном допущении, что величины коэффициентов важности критериев получены методами ТВК. Кроме того, количество ошибок будет возрастать с увеличением числа градаций балльной шкалы и увеличением числа критериев.

5. Взвешенная сумма нормированных неоднородных критериев

Перейдем к рассмотрению общего случая, когда критерии имеют разные шкалы в силу своей «природы» (например, один критерий может характеризовать денежные затраты, другой – надежность системы, третий – ее массу, четвертый – воздействие на окружающую среду, и т.д.), а также иметь разную важность. В этом случае непосредственное использование формулы (1) невозможно. Поэтому вначале исходные критерии приводят к сопоставимому виду, или, как говорят, нормализуют. Нормализованные критерии \hat{f}_i безразмерны и их значения лежат в одинаковых пределах, обычно от 0 до 1. В итоге вместо формулы (1) возникает ее обобщение – формула, имеющая следующий вид:

$$F(\hat{f}|w) = w_1 \hat{f}_1 + \dots + w_m \hat{f}_m. \quad (4)$$

В рассматриваемом общем случае методу взвешенной суммы критериев помимо отмеченных выше присущ ряд других недостатков.

♦ «Интеллектуальная ошибка», вызванная независимостью процедур нормализации критериев и назначения их весов. Вначале зададимся резонным вопросом: «Насколько правомерен подход, основанный на формуле (4)?». Для получения ответа на него обратимся к теории аддитивных функций ценности, имеющих вид:

$$v(f) = v_1(f_1) + \dots + v_m(f_m). \quad (5)$$

Поскольку все критерии желательно максимизировать, то каждая частная функция ценности v_i является возрастающей. При помощи функции ценности (5) каждый вариант оценивается ее значением – числом $v(f(x))$: чем это число больше, тем вариант предпочтительнее. Установлены (при некоторых технических предположениях) условия существования аддитивной функции и доказано, что если $v' = v'_1 + \dots + v'_m$ – другая такая функция, то существуют числа $k > 0$ и l_i такие, что $v'_i = kv_i + l_i$, $i = 1, \dots, m$ (см. [15]). Подчеркнем, что слагаемые l_i для разных критериев могут быть разными, но положительный множитель k – общий для всех критериев! Разработаны научно обоснованные методы построения аддитивных функций ценности [19]. Но они достаточно сложны и поэтому не получили широкого распространения.

Пусть f_i^* и f_i^* – наибольшее и наименьшее значения критерия $f_i(x)$ на множестве вариантов X . При необходимости добавив к каждой функции v_i из (5) число $-v_i(f_i^*)$, можно полагать, что $v_i(f_i^*) = 0$, $i = 1, \dots, m$. Функцию ценности (5) можно представить в таком виде:

$$v(f) = w_1 \hat{v}_1(f_1) + \dots + w_m \hat{v}_m(f_m), \quad (6)$$

$$\text{где } \hat{v}_i(f_i) = v_i(f_i) / v_i(f_i^*),$$

$$w_i = v_i(f_i^*) / \sigma,$$

$$\sigma = v_1(f_1^*) + \dots + v_m(f_m^*).$$

Заметим, что все $\hat{v}_i(f_i^*) = 1$. Теперь понятно, что «по своему происхождению» числа w_i являются шкалирующими множителями.

Сравнивая (4) и (6), видим, что $\hat{v}_i(f_i)$ можно рассматривать как нормализованные критерии \hat{f}_i , а множители w_i – как веса критериев, которые учитывают не только относительную важность критериев f_i , но и соизмеряются размахи значений функций v_i (т.е. играют более сложную роль, чем коэффициенты важности критериев, изначально

имеющих общую шкалу). Из изложенного становится ясным, что нельзя назначать величины весов w_i , не учитывая значений f_i^* и $f_i^!$. Назначение величин w_i с нарушением этого положения названо в [11] *интеллектуальной ошибкой*. Но именно так и назначаются веса при практическом использовании обобщенных показателей (4)!

Эту ошибку нельзя исключить за счет проведения анализа чувствительности (т.е. выяснения, в каких пределах можно изменять веса критериев, сохраняя при этом неизменным полученное решение задачи, например, выделенный наилучший вариант).

Заметим, что в [11] предложен метод SMARTS, предусматривающий корректное назначение весов путем сравнения по предпочтительности векторных оценок специального вида. Но этот метод рассчитан только на тот частный случай, когда предпочтения вдоль шкал критериев возрастают равномерно, или линейно (пояснение понятия равномерности роста предпочтений см. ниже).

♦ *Необоснованность выбора функций для нормализации критериев.* Еще одной серьезной проблемой для МВСК в случае критериев с разными шкалами является подбор вида функции нормализации критериев φ , при помощи которой рассчитываются значения нормализованных критериев: $\hat{f}_i = \varphi_i(f_i)$.

Обычно предлагается использовать линейные функции нормализации:

$$\varphi(f_i) = \frac{f_i}{f_i^*}, \quad \varphi(f_i) = \frac{f_i - f_i^*}{f_i^! - f_i^*}. \quad (7)$$

Использование таких функций означает, что принимается допущение о равномерном росте предпочтений вдоль шкал критериев. Это допущение весьма сильное и часто не выполняется на практике.

Для иллюстрации рассмотрим такую задачу [7]. Предположим, что Вы хотите купить автомашину, и Вам, выслушав все Ваши пожелания, продавец предложил три машины *A*, *B* и *B*. Цена и все интересующие Вас основные их характеристики примерно одинаковы, кроме максимальной скорости f_1 (км/час) и экономичности f_2 (км/л). Значения этих двух характеристик Вы записали в *табл. 1*. Затем рассчитали значения нормализованных критериев по левой из формул (7). Вы решили, что обе характеристики для Вас примерно одинаково значимы, и поэтому приняли значения весов $w_1 = w_2 = 1/2$. Наконец, рассчитали значения обобщенного показателя по формуле (4) для всех марок машин и результаты записали в ту же *табл. 1*.

Таблица 1.

Оценки автомобилей по двум критериям

Марка автомобиля	f_1	f_2	\hat{f}_1	\hat{f}_2	$F(\hat{f} w)$
A	240	10	1	0,667	0,83
Б	140	14	0,583	0,933	0,76
B	120	15	0,5	1	0,75

Поскольку самое большое значение F оказалось для машины *A*, то стало ясно, что именно ее «рекомендует» приобрести использованный МВСК. Однако эта рекомендация Вам почему-то не понравилась. Поразмыслив, Вы заметили, что иметь возможность развить скорость более 200 км/час, конечно, заманчиво. Однако реально со скоростью более 140 км/час Вам поехать если и захочется, то лишь в каких-то исключительных случаях. Но вот за бензин приходится платить постоянно, да и цена на него растет и растет. А максимальная скорость 120 км/час все-таки маловата. Приглядевшись, Вы заметили, что значение обобщенного показателя для машины *A* оказалось больше, чем для машины *B*, именно за счет солидной прибавки в максимальной скорости, которая «перевесила» существенное ухудшение экономичности. А эта прибавка в скорости хотя и велика, но для Вас привлекательности прибавляет немного. Напротив, прибавка в скорости всего на 20 км/час при отсчете от 120 км/час для Вас весьма ощутима и делает машину *B* (при небольшой разнице в экономичности) предпочтительнее машины *B*. Вам стало понятно, что результаты расчетов Вас не удовлетворили потому, что на самом деле Ваши предпочтения вдоль шкалы скорости растут явно неравномерно, и это не было учтено в МВСК. Поэтому, взвесив все «за» и «против», Вы решили выбрать машину *B*.

Для учета неравномерности роста предпочтений нужно использовать нелинейные функции нормализации. Но их построение – новая сложная проблема: ведь они должны количественно описывать изменение предпочтений вдоль шкал критериев. А обычно имеются лишь качественные соображения по поводу подбора таких функций.

♦ *Нарушение аксиомы независимости от посторонних альтернатив.* Отметим еще одну неприятность, связанную с использованием МВСК: если к множеству вариантов добавить еще один заведомо не лучший вариант или, наоборот, удалить из него заведомо не лучший вариант, то рекомендация по выбору лучшего варианта может измениться! В тео-

Таблица 2.

Оценки вариантов по критериям

Варианты	f_1	f_2	\hat{f}_1	\hat{f}_2	$F(\hat{f} w)$	\hat{f}_1	\hat{f}_2	$F(\hat{f} w)$
x^1	5	3	1	0	0,6	1	0	0,6
x^2	3	5	0,5	1	0,7	0	1	0,4
x^3	1	3	0	0	0			

рии принятия решений это называется нарушением аксиомы независимости от посторонних альтернатив [20]. Причина этого явления состоит в том, что удаление или добавление не лучшего варианта может привести к изменению величин f_i^* для некоторых критериев и, как следствие, к изменению значений нормированных критериев для вариантов из оставшегося или же пополненного множества вариантов.

Для иллюстрации рассмотрим двухкритериальную задачу с тремя вариантами, исходные данные для которой, величины нормализованных с использованием правой из формул (7) критериев и значения обобщенного показателя при $w_1 = 0,6$ и $w_2 = 0,4$ приведены в левой части табл. 2. Получилось, что вариант x^2 предпочтительнее варианта x^1 . Поскольку вариант x^3 хуже варианта x^1 (значения второго критерия для них одинаковы, но значение первого критерия больше для x^1), то x^3 не может претендовать на роль наилучшего. Исключим его из рассмотрения. Результаты расчетов для полученной модифицированной задачи приведены в правой части табл. 2. Теперь лучшим оказался вариант x^1 .

6. Заключение

МВСК является привлекательным эвристическим методом, которому, однако, присущ целый ряд неустраняемых недостатков принципиального характера. И это надо иметь в виду всем, кто его применяет или планирует использовать для принятия решений.

И, тем не менее, взвешенная сумма критериев очень часто используется при анализе прикладных многокритериальных задач. Почему? Известный ученый и популяризатор науки профессор Вентцель Е.С. так ответила на этот вопрос. «Здесь мы встречаемся с очень типичным для подобных ситуаций приемом – «переносом произвола из одной инстанции в другую». Простой выбор компромиссного решения на основе мысленного сопоставления всех «за» и «против» каждого решения кажется слишком произвольным, недостаточно «научным». А вот маневрирование с формулой, включающей (пусть столь же произвольно назначенные) коэффициенты – совсем другое дело. Это уже «наука»! По существу же никакой науки тут нет, и нечего обманывать самих себя» [21]. ■

Литература

1. Гафт М.Г. Принятие решений при многих критериях. – М.: Знание, 1979.
2. Горский П. Введение в дисциплину «Поддержка принятия решений». <http://www.pavel.gorskiy.ru/Articles/Dmss/d0.html> (дата обращения: 21.06.2013).
3. Ларичев О.И. Теория и методы принятия решений: Учебник. Изд. третье, перераб. и доп. – М.: Логос, 2006.
4. Лотов А.В., Поспелова И.И. Многокритериальные задачи принятия решений: Учебное пособие. – М.: Издательский отдел факультета ВМиК МГУ им. М.В. Ломоносова; МАКС Пресс, 2008.
5. Орлов А.И. Принятие решений. Теория и методы разработки управленческих решений: Учебное пособие. – М.: MapT, 2005.
6. Подиновский В.В. Аксиоматическое решение проблемы оценки важности критериев в многокритериальных задачах принятия решений // Современное состояние теории исследования операций / Под ред. Н.Н. Моисеева. – М.: Наука, 1979. – С. 117-145.
7. Подиновский В.В. Введение в теорию важности критериев в многокритериальных задачах принятия решений: Учебное пособие. – М.: Физматлит, 2007.
8. Теория выбора и принятия решений: Учебное пособие / Макаров И.М., Виноградская Т.М., Рубчинский А.А., Соколов В.Б. – М.: Физматлит, 1982.

9. Belton V., Gear T. On the meaning of relative importance // Journal of multi-criteria decision analysis. – 1997. – Vol. 6. – P. 335-338.
10. Belton V., Stewart T.J. Multiple criteria decision analysis: an integrated approach. – Boston: Cluwer, 2003.
11. Edwards W., Barron F.H. SMARTS and SMARTER: improved simple methods for multiattribute utility measurement // Organization behavior and human processes. – 1994. – Vol. 60. – P. 306-325.
12. Multiobjective optimization: interactive and evolutionary approaches / J. Branke, K. Deb, K. Miettinen, R. Slowinski (Eds.). – NY: Springer, 2008.
13. Roy B., Mousseau V. A theoretical framework for analyzing the notion of relative importance of criteria // Journal of multi-criteria decision analysis. – 1996. – Vol. 5. – P. 145-159.
14. Миркин Б.Г. Проблема группового выбора. – М.: Физматлит, 1974.
15. Foundation of measurement. Vol. 1 / D.H. Krantz, R.D. Luce, P. Suppes, A. Tverski. – NY: Academic Press, 1971.
16. Roy B. The outranking approach and the foundation of ELECTRE methods // Theory and decision. – 1991. – Vol. 31. – P. 49-73.
17. Подиновский В.В., Потапов М.А. Важность критериев в многокритериальных задачах принятия решений: теория, методы, софт и приложения // Открытое образование. – 2012. – № 2. – С. 55-61.
18. Салтыков С.А. Экспериментальное сопоставление методов взвешенной суммы, теории полезности и теории важности критериев для решения многокритериальных задач с балльными критериями // Управление большими системами. – 2010. – Вып. 29. – С. 16-41.
19. Кини Р., Райфа Х. Принятие решений при многих критериях: предпочтения и замещения / Пер. с англ. – М.: Радио и связь, 1981.
20. Льюс Р., Райфа Х. Игры и решения. Введение и критический обзор / Пер. с англ. – М.: ИЛ, 1961.
21. Вентцель Е.С. Исследование операций: задачи, принципы, методология: Учебное пособие. 2-е изд., стер. – М.: Высшая школа, 1988.

ЖДЕТ СВОИХ ПОЛЬЗОВАТЕЛЕЙ НОВЫЙ УНИКАЛЬНЫЙ В МИРОВОЙ ПРАКТИКЕ

Универсум связей между русскими словами CrossLexica

- ◆ К сверхбольшому компьютерному словарю, по объему уже не воспроизводимому в печати, обращается либо человек – интерактивно, либо внешний софт – программно.
- ◆ Словарь предельно политематичен: от математики и философии до политики, спорта, строительства, кулинарии, автомобилей, моды и бранной лексики, допущенной Госдумой.
- ◆ Включает информацию как языкового, так и энциклопедического характера.
- ◆ Беспрецедентен по объему: 292 тыс. элементов словника и 8 млн. связей между ними: 4.5 млн. связей в словосочетаниях, 2.8 млн. смысловых связей и 0.7 млн. связей внешнего сходства слов. 47% элементов словника сами состоят из нескольких слов.
- ◆ Объединяет в себе словари словосочетаний, синонимов, антонимов, паронимов, словоизменения, тезаурусных иерархий, глагольно-именного управления и др.
- ◆ Рассчитан на любой класс пользователей – на ученого, преподавателя, инженера, журналиста, бизнесмена, военного, студента, школьника, пенсионера, домохозяйку.
- ◆ Обучает грамотному русскому языку как русскоговорящего, так и иностранца.
- ◆ Обеспечивает доступ иностранцу еще и через английские переводы словника.
- ◆ Быстро выбирает WEB-запрос из наличных миллионов, посылая его в Google / Яндекс.
- ◆ Отвечает на запросы с клавиатуры и те, что возникают при навигации по словнику.
- ◆ Размещается в ОП любого компьютера, запускается под Windows XP и Windows 7.
- ◆ Не повторяя существующие толковые словари, способен поднять как интерактивную, так и автоматическую обработку русских текстов на уровень, ныне не достижимый.

На эксклюзивную продажу выставляется пакет:

текущая версия, исходная текстовая база, грамматические таблицы, утилиты морфоклассификации слов и сборки рабочей версии, коды интерфейса с человеком и компьютером, технология правки и пополнения базы.

Автор и владелец: **доктор наук профессор ИГОРЬ АЛЕКСЕЕВИЧ БОЛЬШАКОВ.**

Контакт: **ibolshakov@gmail.com**

Подробнее: см. статью в данном номере.

МЕТОД ЭМПИРИЧЕСКИХ ВЕРОЯТНОСТЕЙ: АВТОМАТИЧЕСКАЯ СИСТЕМА ДЛЯ РЕКОМЕНДАЦИИ СЛЕДУЮЩИХ ДЕСЯТИ ЛЕКЦИЙ КУРСА ПОСЛЕ ПРОСМОТРА ТРЕХ ДАННЫХ ЛЕКЦИЙ

В.Н. Никулин, кандидат физико-математических наук, доцент кафедры математических методов в экономике Вятского государственного университета

С.А. Палешева, студентка кафедры математических методов в экономике Вятского государственного университета

Д.С. Зубарева, студентка кафедры математических методов в экономике Вятского государственного университета

*E-mail: vnikulin.uq@gmail.com, s.palesheva@gmail.com, zubarevadasha@yandex.ru
Адрес: г. Киров, ул. Московская, д. 36*

В статье представлен алгоритм, который был награжден призом за третий лучший результат, продемонстрированный в ходе международного соревнования по анализу данных VideoLectures.Net ECML/PKDD 2011 (Track 2). Мы предлагаем использовать две лекции (взятые из тройки данных лекций), для того чтобы определить направление прогноза. Соответствие всего предсказанного набора вычисляется согласно оставшейся третьей лекции.

Ключевые слова: рекомендательные системы, прогнозирование, оценки популярности, ансамбли и элементарные классификаторы, бэггинг, анализ данных.

1. Введение

VideoLectures.Net¹ – это открытый и общедоступный мультимедийный ресурс видео-лекций, в основном исследовательского и образовательного характера. Лекции предлагаются выдающимися учеными и исследователями в рамках самых значимых и известных событий, таких как: конференции, лет-

ние школы, семинары, а также научно-популярные мероприятия в различных отраслях науки. Целями интернет-ресурса являются продвижение научных идей, стимулирование обмена знаниями, которые достигаются посредством предоставления высококачественных учебных материалов не только научной общественности, но и более широкой аудитории. Все лекции, включая документы, информацию и ссылки, систематизированы и сгруппированы редакторами с учетом комментариев пользователей.

¹ <http://videlectures.net/>

Задача международного соревнования VideoLectures.Net в рамках Центральной европейской конференции по анализу данных ECML/PKDD 2011 состояла в подготовке списка рекомендованных лекций ресурса VideoLectures.Net на основе исторических данных с этого сайта. Описание методов, которые использовал победитель соревнования VideoLectures.Net даны в [1]. Наш метод [2] был награжден третьим призом (Track 2).

Согласно [3], открытые социально-образовательные системы предоставляют новые возможности для миллионов заинтересованных студентов, для того чтобы последние могли пользоваться высококачественными дидактическими материалами в режиме реального времени. В соответствии с известными оценками, более 100 миллионов студентов по всему миру имеют вполне достаточный уровень образования для поступления в университет в течение следующих десяти лет. Университеты откликаются на сложившиеся потребности посредством создания открытых образовательных ресурсов: тысячи общедоступных высококачественных online-курсов, подготовленных сотнями преподавателей университетов, используются миллионами людей по всему миру. К сожалению, учебные материалы, соответствующие курсу в режиме реального времени, не дают достаточный опыт для эффективного изучения, что является необходимым условием для поддержания заинтересованности студентов.

Однако в настоящее время студенты обеспокоены качеством своего образования. С целью стимулирования и облегчения процесса обмена опытом для этих студентов необходимо решить два важных вопроса: 1) создание и накопление библиотеки учебных материалов online; 2) стимулирование обмена опытом в режиме реального времени (общение).

Центральная проблема заложена как раз во втором вопросе: каким образом задать подходящее направление для студентов-слушателей, которые «живут» в Интернете, принимая во внимание существование многообразия доступных исследовательских / образовательных ресурсов.

Функциями рекомендательных систем является профилирование пользователей по определенным критериям предпочтения и моделирование соотношений между пользователями и предметами потребления. Задача подобной системы заключается в формировании рекомендаций для того, чтобы максимально удовлетворить вкусы пользователей и облегчить последним выбор из огромного разнообразия предлагаемых услуг и товаров [4]. Рекомендательные системы имеют огромное значение в таких

сферах деятельности, как: электронная торговля, подписка на базовые службы, отбор информации и др. Рекомендательные системы, формирующие персонализированные предложения, значительно повышают вероятность осуществления покупки клиентом. Индивидуальные рекомендации особенно важны на рынке, где выбор достаточно велик, вкусы потребителей играют значительно большее значение в сравнении с ценами, которые ограничены. Типичными сферами применения подобных систем являются: искусство (книги, фильмы, музыка), мода и одежда, еда и рестораны, игры и юмор.

Большинство методов, представленных в [4], были мотивированы известным соревнованием по анализу данных NetflixCup. Отметим, что методы, основанные на разложении матриц, не могут быть применены в нашем случае напрямую, поскольку данные имеют иную структуру. В нашем случае мы имеем дело не с конкретными, а с абстрактными пользователями, которые ранее ознакомились с содержанием трех лекций из предложенного набора лекций, причем их точная последовательность неизвестна.

Использование традиционных методов анализа данных (ассоциативные правила) позволило получить хорошие результаты на ранних стадиях разработки рекомендательных систем [5]. Наиболее часто используемые наборы лекций (или иных товаров/предметов потребления), определенные методом ассоциативных правил, представляют собой тип направляющих или ориентирующих образцов, поскольку они сконцентрированы на факте наличия лекций, нежели на их порядке, в котором происходит процесс рассмотрения или обучения [6]. Частотные методы (или методы, основанные на эмпирических вероятностях) являются основным инструментом в следующих 3.1 – 3.6 разделах. Заметим также, что модель Марковских цепей для принятия решений позволяет улучшить качество принятия решений для рекомендательных систем в случае, если последовательность состояний известна. Согласно теории Марковских цепей, мы имеем дело с пространством, в котором число состояний ограничено, и, используя оценку максимального правдоподобия (опирающуюся на исторические данные), мы формулируем прогноз или предсказание.

Метод бэггинг используется для вычисления множества элементарных предсказаний с целью формирования суммарного (совокупного) предиктора. Этот предиктор представляет собой усреднение относительно элементарных предикторов и дает прогноз в соответствии с большинством голосов. В разделе 3.6 мы рассмотрим метод случайных повторных выбо-

рок: предполагается, что, используя сотни предикторов (элементарных классификаторов), опирающихся на подмножества всей тренировочной выборки, мы сможем уменьшить эффект случайных факторов. Согласно принципам, на которых базируются однородные ансамбли, финальный предиктор представляет собой среднее элементарных предикторов. Отметим, что параллельно вычислению однородного ансамбля мы можем вычислить CV-паспорт [7] (cross-validation passport) для оценки качества решения. Заметим, что популярная модель случайных деревьев [8] является хорошо известным примером удачного однородного ансамбля. Однако структура случайных деревьев основана на другом методе, который опирается прежде всего на признаки, но не на подмножества.

2. Данные и некоторые определения

Тренировочная база данных состоит из двух подмножеств: 1) пары лекций P ; 2) тройки лекций T , каждая из которых включает две части (левую и правую), где левая часть содержит входные тройки лекций и соответствующие количества их просмотров, правая часть содержит выходные лекции и соответствующие количества их просмотров.

2.1. Пары лекций

Обозначим через I_p множество индексов, соответствующих парам данных. Любой элемент из I_p представляет неупорядоченный набор из двух индексов $I = \{a, b\}$, где $I \in I_p$, а индексы a и b однозначно определяют соответствующие лекции. Под выражением $P_I = P_{ab}$ мы будем понимать число тех случаев, когда лекции с индексами a и b были просмотрены вместе.

2.2. Тройки лекций

Обозначим через I_T множество индексов, соответствующих тройкам данных: $\tau_i = \{a, b, c\}$ – это тройка или набор из трех лекций a, b и c . Элемент с индексом $I \in T$ имеет два значения (одно для левой и одно для правой частей). Под выражением T_I мы будем понимать число случаев, когда соответствующие три лекции были просмотрены вместе; L_I – набор отдельных лекций, просмотренных после τ_i . А также обозначим через $T_I(\ell)$, $\ell \in L_I$, количество случаев когда лекция была просмотрена после тройки τ_i .

2.3. Графические иллюстрации

Рис. 1 иллюстрирует гистограмму эмпирических вероятностей или частот:

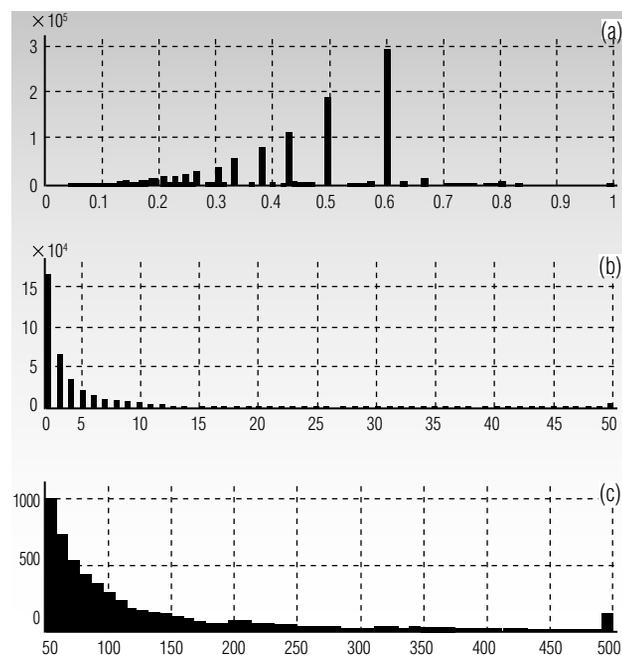


Рис. 1. Гистограммы частот, соответствующих (а) тройкам; (б) парам до 50; (с) парам от 50 до 500, см. раздел 2.3

$$r_\tau(\ell) = \frac{T_I(\ell)}{T_I}, \ell \in L_I, I \in I_T, \quad (1)$$

где мы заменили I на τ в левой части (1), поскольку имеет место взаимно однозначное соответствие между параметрами I и τ .

Рис. 1(б-с) иллюстрирует гистограмму частот: $P_I, I \in I_p$, где все значения на рис. 1(б) сокращены до уровня 50 (если значения превышают 50). Дополнительно рис. 1(с) демонстрирует гистограмму значений P_I от 50, где значения, превышающие 500, сокращены до 500.

3. Методы

3.1. Прогнозы при использовании двоек лекций

Задача соревнования заключалась в построении прогноза (предсказания) согласно тестовой базе данных V , которая имеет такую же структуру, как и T , (левая часть). Точнее говоря, было необходимо предложить рекомендацию десяти наиболее подходящих лекций после просмотра данных трех.

Замечание 1. В качестве особенного и наиболее важного фактора данного соревнования отметим отсутствие одинаковых троек в тренировочном T (левая часть) и тестовом V множествах. В то же время мы обнаружили значительное количество одинаковых пар в обоих множествах: тренировочном и тестовом.

В целом мы нашли $n_c = 34756$ двоек (которые были найдены в левой части T) с количеством повторов, соответствующих отдельной двойке, от 1 до $m_c = 4020$.

Отметим, что каждая тройка может быть рассмотрена как набор из трех двоек. Общее количество троек в тестовой выборке V составило $N_V = 60274$, и 1) мы не обнаружили соответствующих им двоек в тренировочной выборке только в 116 случаях, 2) мы нашли только одну двойку в 829 случаях, 3) мы нашли две двойки в 4705 случаях; 4) все три двойки были обнаружены в абсолютном большинстве (54624 случаев).

Замечание 2. В оригинальной базе данных каждая лекция была идентифицирована при помощи индекса, где наибольший индекс $n_L = 13251$. Однако не все n_L лекции были использованы. Мы предположили, что предсказанные лекции следует искать в правой части T , где определены только $n_s = 5209$ различных лекций.

3.2. Некоторые предварительные определения и обозначения

Мы объясним, как работает система в условиях n_s вторичных индексов, так как преобразование к первоначальным n_L индексам является тривиальной задачей. Наша база данных была организована следующим образом: квадратная матрица A размерности $n_s \times n_s$ содержит n_c различных адресов матрицы B размерности $n_c \times m_c$.

Сначала мы найдем три пары $\alpha_{ij}, j = 1, 2, 3$, для всех троек $\tau_i, i = 1, \dots, N_V$, содержащихся в тестовой базе данных V . Затем для каждой пары α_{ij} мы найдем (в соответствии с матрицей A) соответствующий адрес $\beta(\alpha_{ij})$ (номер строки в матрице B) и количество имеющихся записей $n(\alpha_{ij})$, где $1 \leq \beta \leq n_c, 1 \leq n \leq m_c$.

Под элементом матрицы B мы будем понимать предсказанную/рекомендованную лекцию ℓ и соответствующую частоту:

$$\{\ell, r_\tau(\ell)\}, \quad (2)$$

где отношение r_τ определено в (1).

Процесс обновления

Здесь мы опишем наиболее важный шаг вычислительного процесса. Каждая конкретная тройка τ тренировочной базы данных рассматривается идентичным образом, поэтому мы опускаем индекс i с целью упрощения обозначений.

Предположим, что первоначально все рейтинги равны нулю $s(\ell) = 0, \ell = 1, \dots, n_s$, где $s(\ell)$ – это рейтинг, соответствующие лекции ℓ , которые будут использованы для конечного ранжирования лекций в качестве результата модели. Далее представлена наиболее важная формула для обновления:

$$s(B_{\beta k}(1)) = s(B_{\beta k}(1)) + B_{\beta k}(2), k = 1, \dots, n(\alpha_j), j = 1, 2, 3, \quad (3)$$

где $B_{\beta k}(1)$ – индекс лекции, $B_{\beta k}(2)$ – соответствующая частота, которая была определена в (2).

После вычисления вектора s в соответствии с (3) мы отсортируем соответствующие элементы в порядке убывания, и индексы лекций, соответствующие десяти наибольшим значениям s , могут быть представлены в качестве решения.

Замечание 3. В случае если число положительных значений вектора s меньше, чем 10, мы генерировали оставшиеся индексы случайным образом, предполагая, что они отличны от 1) индексов входящей тройки τ_i , а также отличны от 2) тех индексов, что уже выбраны.

При помощи метода, описанного в этом разделе, был получен результат 0,49568 в терминах критерия качества, который использовался организаторами соревнования² для сравнения различных решений:

$$MAP_p = \frac{1}{|V|} \sum_{i \in V} AvgRp(\ell), \text{ где}$$

$$AvgRp(\ell) = \frac{1}{|Z|} \sum_{z \in Z} Rp@z(\ell),$$

$$Rp@z(\ell) = \frac{|relevant \cap retrieved|_z}{|relevant|_z},$$

где $Z = \{5, 10\}$ – первые 5 или все 10 лекций (имеется в виду, что рекомендованные 10 лекций отсортированы в порядке убывания в соответствии с рейтингом).

Замечание 4. Главное преимущество описанного выше метода, который основан на информационной базе данных, включающей матрицы A и B , состоит в его скорости. Согласно проведенным экспериментам, изложенный в этом разделе алгоритм прошёл через всю тренировочную выборку V и вычислил требуемое решение в течение пятидесяти одной секунды. Использовались 1) многопроцессорная рабочая станция с операционной системой Linux 3.2 GHz 16 GB RAM и 2) специально разработанная программа написанная на языке программирования C (все временные замеры осуществлялись в автоматическом режиме).

3.3. Прогнозы при использовании отдельных лекций

Отметим, что прогнозы при использовании отдельно взятых лекций работают схожим образом, что и прогнозы с парными предсказаниями. Однако имеются некоторые различия, которые могут быть рассмотрены как упрощения. Мы выяснили, что максимальное число записей, соответствующих отдельной лекции, – $m_s = 77798$. Соответственно

² <http://tunedit.org/challenge/VLNetChallenge>

матрица \hat{B} (замена матрицы B в предыдущем разделе 3.1) имеет размерность $n_s \times m_s$.

Модель работает следующим образом: по определению, любая тройка представляет собой множество (набор из трех лекций), состоящее из трех лекций ℓ_1, ℓ_2, ℓ_3 . Мы найдем количество записей для каждой отдельной лекции $1 \leq n(\ell) \leq m_s$, где $1 \leq \ell \leq n_s$.

Процесс обновления

Как и ранее, первоначально все рейтинги равны нулю: $s(\ell) = 0, \ell = 1, \dots, n_s$. Ниже представлена основная формула для обновления:

$$s(\hat{B}_{\ell_j k}(1)) = s(\hat{B}_{\ell_j k}(1)) + \hat{B}_{\ell_j k}(2), k = 1, \dots, n(\ell_j), j = 1, 2, 3. \quad (4)$$

После вычисления вектора s в соответствии с (4) сортируем значения в порядке убывания, и индексы лекций, соответствующие десяти наибольшим значениям s , представимы в качестве решения.

Метод, который описан в этом разделе, позволил нам получить результат 0,33278.

3.4. Прогнозирование при использовании пар лекций

Определение 1. Будем называть лекции a и b P -связанными, если $P_{ab} \geq 1$. В соответствии с симметричной матрицей P определим множество лекций $H(a)$, которые P -связаны с данной лекцией a .

Процесс обновления

Как и ранее, первоначально все рейтинги равны нулю: $s(\ell) = 0, \ell = 1, \dots, n_s$.

Далее мы представим формулу для обновления:

$$s(d) = s(d) + P(\ell_j, d), d \in H(\ell_j), j = 1, 2, 3, \quad (5)$$

где определение лекций ℓ_j является таким же, что и в (4).

После вычисления вектора s в соответствии с (5) мы сортируем значения в порядке убывания, и индексы лекций, соответствующие десяти наибольшим значениям s , представимы в качестве решения.

При использовании метода, представленного в этом разделе, был получен результат 0,12677.

Замечание 5. Решение, описанное в данном разделе, было рекомендовано организаторами форума как «simple pairs solution».

Заметим, что в ходе наших числовых экспериментов мы сделали довольно интересное наблюдение.

Замечание 6. Оценки, определенные в (3-5), представляют суммы частот. Очень интересно отметить, что результаты будут значительно слабее, если мы используем среднее в качестве альтернативы сумме.

3.5. Прогнозы при использовании двоек с весовыми коэффициентами

В соответствии с предыдущими тремя разделами, прогнозы при использовании двоек позволили нам получить лучшие результаты. Мы решили продолжить работу в этом направлении и принять во внимание оставшиеся третьи лекции φ и ψ в обеих тренировочной и тестовой выборках.

Основная идея метода: в случае если оставшиеся лекции φ и ψ подобны (имеют большое количество общих просмотров в соответствии с парными данными), направление прогноза, соответствующее двойке, приобретает больший вес.

Как было отмечено в замечании 1, лекции φ и ψ различны по определению. Иными словами, схожие тройки из тренировочной и тестовой баз данных могут быть представлены в следующем виде: $\alpha_j \cup \varphi_j, \alpha_j \cup \psi_j$, где $\varphi_j \neq \psi_j, j = 1, 2, 3$.

Процесс обновления

Как и ранее, первоначально все рейтинги равны нулю: $s(\ell) = 0, \ell = 1, \dots, n_s$. Затем мы можем переписать (3) следующим образом:

$$s(B_{\beta k}(1)) = s(B_{\beta k}(1)) + w(P(\varphi_j, \psi_j)) B_{\beta k}(2), \quad (6)$$

$$k = 1, \dots, n(a_j), j = 1, 2, 3,$$

где w – возрастающая функция весовых коэффициентов. При вычислении нашего финального решения мы использовали простейшую линейную функцию: $w(x) = 0.01 \cdot x + 0.05$.

После вычисления вектора s в соответствии с (6) мы сортируем полученные значения в порядке убывания, и индексы лекций, соответствующие десяти наибольшим значениям s , представимы в качестве решения.

При использовании метода с взвешенными двойками, который был описан нами в этом разделе, было достигнуто значительное улучшение: 0,58145.

3.6. Прогнозы при использовании отдельных лекций с весовыми коэффициентами

Данный раздел может быть рассмотрен как дополнение к разделу 3.2. В некотором смысле прогнозы с взвешенными отдельно взятыми лекциями схожи с прогнозами с взвешенными двойками (раздел 3.4). Однако есть некоторые отличия. В случае с отдельными лекциями мы определяем направление прогноза согласно одиночным лекциям. Соответственно мы имеем две другие (дополнительные) лекции, которые необходимо сравнить с двумя лек-

циями из соответствующей тройки лекций из тренировочного множества.

Процесс обновления осуществляется следующим образом:

$$s(B_{\ell,k}(1)) = s(B_{\ell,k}(1)) + w(\varphi_{1j}, \varphi_{2j}; \psi_{1j}, \psi_{2j}) \cdot B_{\ell,k}(2),$$

$$k = 1, \dots, n(\ell_j), j = 1, 2, 3, \text{ где} \quad (7)$$

$$w(\varphi_{1j}, \varphi_{2j}; \psi_{1j}, \psi_{2j}) =$$

$$= 0,0005 \cdot \left(P(\varphi_{1j}; \psi_{1j})P(\varphi_{2j}; \psi_{2j}) + P(\varphi_{1j}; \psi_{2j})P(\varphi_{2j}; \psi_{1j}) \right) + 0,01.$$

Идея данной формулы проста: мы должны быть уверены, что каждая дополнительная лекция из тестовой тройки лекций близка по крайней мере к одной дополнительной лекции из тренировочной тройки.

После вычисления вектора s согласно (7) мы сортируем полученные результаты в порядке убывания, и индексы лекций, соответствующие десяти наибольшим значениям s , представимы в качестве решения.

При использовании данного метода был получен результат 0,4529.

3.7. Метод случайных повторных выборок (финальная рекомендательная система)

Вычисление отдельного упорядоченного вектора s в данном разделе основано на 75% случайно выбранных выборках. В абсолютном большинстве всех 60274 тестовых образцов число положительных компонентов вектора s , определенных в (6), больше 100. Поэтому мы будем рассматривать только этот случай.

Таблица 1.

Различия между пятью решениями (в терминах дистанции (8)), представленными в разделах 3.1 - 3.7

N	Метод/Раздел	Результат	1	2	3	4	5
1	3.1	0,49568	0	0,2605	0,2137	0,6394	0,6517
2	3.3	0,33278	0,2605	0	0,5832	0,4269	0,4327
3	3.4	0,12677	0,2137	0,5832	0	0,1565	0,1664
4	3.5	0,58145	0,6394	0,4269	0,1565	0	0,91
5	3.7	0,58727	0,6517	0,4327	0,1664	0,91	0

Обозначим вектор вторичных рейтингов как z , представляющий собой множество нулей в начале всего процесса повторных выборок. Мы исследовали 200 случайных выборок (глобальные итерации). После каждой глобальной итерации только 100 лекций (компоненты вектора z) получили приращение в диапазоне от 1 до 100 голосов (чем больше, тем лучше).

Метод, который мы использовали внутри каждой глобальной итерации (элементарный классификатор), описан в разделе 3.4.

После завершения всех 200 глобальных итераций мы отсортировали вектор z в порядке убывания, и индексы лекций, соответствующие десяти наибольшим значениям z , представим в качестве решения.

Таким образом, при использовании метода повторных выборок, представленного в этом разделе, был получен результат **0,58727**. Этот результат был использован в качестве итогового результата.

3.8. Статистическое сравнение различных решений

Отметим, что любое решение представляет собой целочисленную матрицу $N_v \times 10$, $T = 10N_v$ целочисленных индексов в целом. Путем сравнения двух матриц мы найдем число общих индексов (пересечение) в каждом ряду. Общее число пересечений даст нам значение числителя R , и требуемое расстояние будет представлено в виде отношения:

$$D = \frac{R}{T} \quad (8)$$

3.9. Время вычисления

Для вычислений была использована многопроцессорная рабочая станция с операционной системой Linux 3.2 GHz 16 GB RAM. Вычисления были произведены на основе специально разработанного кода (алгоритма) в C. Для получения финального решения, описанного в разделе 3.6, потребовалось около 12 часов.

4. Заключительные замечания

Мы согласны с утверждением [9], что превосходство новых алгоритмов следует демонстрировать на независимых данных. В этом смысле важность соревнований по анализу данных является неоспоримой. Стремительно растущая популярность подобных соревнований свидетельствует о том, что они являются одним из самых эффективных способов оценки различных моделей и систем.

В целом, мы удовлетворены нашими результатами, продемонстрированными в ходе соревнования PKDD2011. В качестве одного из направлений для дальнейшего развития было бы интересно найти эффективный способ построения неоднородных ансамблей. Например, при использовании методов отдельных лекций со взвешенными коэффициентами (см. раздел 3.6) и двоек лекций (см. раздел 3.1) результаты имеют значительные расхождения. Тем

не менее, оба результата являются достаточно хорошими и заслуживают внимание с тем, чтобы найти, каким образом интерпретировать и использовать различия между исходными/базовыми решениями для построения решения более высокого уровня.

Кроме того, мы полагаем, что метод градиентной факторизации [10-11] вполне применим в этой задаче и может привести к принципиально новому высококачественному решению.

Отметим, что предложенный метод эмпирических вероятностей был мотивирован структурой данных Международного соревнования PKDD 2011 и имеет тесную связь с популярным методом ассоциативных правил, который нашел широкое применение в ряде областей, см. [12 - 14]. В качестве альтернативного примера и иллюстрации применения метода эмпи-

рических вероятностей мы можем рассмотреть задачу определения влияния принятия лекарств на последующие состояния пациентов [15]. Путем сравнения реальных событий (интенсивность которых измеряется путем метода эмпирических вероятностей) и аналитически ожидаемых событий мы можем выявить интересные закономерности и явления. Этой задаче было посвящено Международное соревнование по анализу данных, организованное американской компанией ОМОП (Observational Medical Outcomes Partnership). Кубок ОМОП 2010³ включал две секции, где описание метода победителя в первой секции опубликовано в статье [16]. Наш метод был официально признан лучшим в секции №2.

Авторы благодарны анонимному рецензенту за ряд полезных замечаний. ■

Литература

1. Дьяконов А.Г. Алгоритмы для рекомендательной системы: технология LENKOR // Бизнес-Информатика. – 2012. – Т. 1, № 19. – С. 32-39.
2. Nikulin V. OpenStudy: recommendations of the following ten lectures after viewing a set of three given lectures // ECML/PKDD workshop and conference proceedings, discovery challenge. Editors: Tomislav Smuc, Nino Antulov-Fantulin, Mikołaj Morzy. – Athens, Greece, 2011. – С. 59-70.
3. Ram A., Ai H., Ram P., Sahay S. Open social learning communities // In International conference on web intelligence, mining and semantics. – Sogndal, Norway, 2011.
4. Takacs G., Pillaszy I., Nemeth B., Tikk D. Scalable collaborative filtering approaches for large recommender systems // Journal of machine learning research. – 2009. - No. 10. – P. 623-656.
5. Agrawal R., Srikant R. Fast algorithms for mining association rules // Proceedings of 20th Int. conf. very large data bases. 1994. – 32 p.
6. Mobasher B., Dai H., Luo T., Nakagawa M. Using sequential and non-sequential patterns in predictive web usage mining tasks // ICDM. – 2002.
7. Ефимов Д.А., Никулин В.Н. Предсказание биологического состояния молекул исходя из их химических свойств // Advanced Science, Вятский Государственный Университет. – 2013. – Т. 2, № 2. – С. 107-123.
8. Breiman L. Random Forests // Machine Learning. – 2001. – Vol. 45, No.1. – P. 5-32.
9. Jelizarow M., Guillemot V., Tenenhaus A., Strimmer K., Boulesteix A.-L. Over-optimism in bioinformatics: an illustration // Bioinformatics. – 2010. – Т. 26, № 16. С. 1990-1998.
10. Nikulin V., Huang T.-H., Ng S.-K., Rathnayake S., McLachlan G. A very fast algorithm for matrix factorization // Statistics and probability letters. – 2011. – Vol. 81. – P. 773-782.
11. Rendle S. Factorization machines with libFM // ACM Transactions on Intelligent Systems and Technology (TIST). – 2012. – Vol. 3, No. 3. – P. 22.
12. Papender K., Deepak D., Nidhi P. Diagnosis of tuberculosis using association rule method // Journal of Information and Operations Management. – 2012. – Vol. 3, No. 1. – P. 133-135.
13. Martinez-Ballesteros M., Troncoso A., Martinez Alvarez F., Riquelme J. Mining quantitative association rules based on evolutionary computation and its application to atmospheric pollution // Integrated Computer-Aided Engineering. – 2010. – Vol. 17. – P. 227-242.
14. Gautam P., Pardasani K. Efficient method for multiple-level association rules in large databases // Journal of Emerging Trends in Computing and Information Sciences. – 2011. – Vol. 2, No. 12. – P. 722-732.
15. Norén G., Hopstadius J., Bate A., Star K., Edwards I. Temporal pattern discovery in longitudinal electronic patient records // Data Mining and Knowledge Discovery. – 2009.
16. Schuemie M. Methods for drug safety signal detection in longitudinal observational databases: LGPS and LEOP-ARD // Pharmacoepidmiology and Drug Safety. – 2010.

³ <http://omop.fnih.org/omopcup>

АВТОМАТИЗАЦИЯ ПРОЦЕССОВ РАЗРАБОТКИ ТЕХНОЛОГИЧЕСКИХ ДОРОЖНЫХ КАРТ. РАСЧЕТ ИНТЕГРАЛЬНЫХ ПОКАЗАТЕЛЕЙ ПРИМЕНИМОСТИ¹

О.В. Ена,

*директор центра информационно-аналитических систем
Института статистических исследований и экономики знаний
Национального исследовательского университета
«Высшая школа экономики»*

К.В. Нагаев,

*старший научный сотрудник центра информационно-аналитических
систем Института статистических исследований и экономики знаний
Национального исследовательского университета
«Высшая школа экономики»*

*E-mail: ovena@hse.ru, knagaev@hse.ru
Адрес: г. Москва, ул. Мясницкая, д. 18*

Рассмотрены способы расчета частных и интегральных показателей применимости элементов технологической дорожной карты и их атрибутов в целях применения автоматизированных средств для задач, связанных с форсайтными исследованиями, долгосрочным научно-техническим прогнозированием, определением инновационных траекторий развития предметных областей.

Ключевые слова: дорожная карта, форсайт, показатель применимости, технологическая траектория, формализация знаний, экспертный опрос, МАУТ.

¹ Работа проводилась при финансовой поддержке Министерства образования и науки Российской Федерации по государственному контракту № 07.524.12.4018 от 16 мая 2012 г. в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2013 годы»

1. Введение

В мировой практике дорожная карта является одним из наиболее эффективных инструментов долгосрочного стратегического планирования и может использоваться на всех уровнях принятия решений: в рамках формирования и реализации стратегий корпоративного уровня, конкурентных и функциональных стратегий [1].

Наиболее сложным видом дорожных карт является технологическая дорожная карта (далее, ТДК). ТДК позволяет оценить влияние на развитие предметной области большого количества разнообразных факторов: глобальные вызовы, перспективы коммерциализации ключевых инновационных технологий; желаемую конфигурацию продуктов и рынков и др. [2].

Дорожная карта формируется в виде набора слоев (групп однотипных элементов). Информационное наполнение элементов ТДК и взаимосвязи между ними определяются в соответствии с базовыми алгоритмами расчета элементов ТДК, а также консолидации экспертных сведений [3].

В качестве слоев дорожной карты в общем случае выступают: 1) научно-технологическое развитие; 2) технологии; 3) продукты; 4) рынки.

Автоматизация процессов создания и применения ТДК связана с решением трех классов научно-технологических задач:

1) формализацией предметной области дорожного картирования и автоматизированным сбором релевантных сведений из разнородных источников информации;

2) поиском экспертов в предметной области дорожного картирования и многоэтапной консолидацией экспертных мнений;

3) унификацией интегральных показателей, характеризующих элементы ТДК и их атрибуты, и расчетом основной технологической траектории ТДК.

В настоящей статье рассматриваются подходы к решению третьей научно-технологической задачи. Научно-технологические решения, представленные в настоящей статье, реализованы в составе программного комплекса «Интерактивная дорожная карта с обратной связью».

2. Постановка задачи определения частных и интегральных показателей применимости

Главным аналитическим результатом ТДК является основная технологическая траектория – набор взаимосвязанных элементов различных слоев ТДК, обеспечивающих оптимальный план долгосрочного развития предметной области дорожного картирования.

Классическим представлением основной технологической траектории является базовая схема, предложенная Европейской ассоциацией управления промышленными исследованиями (EIRMA)[4] (на рис. 1 элементы и связи основной технологической траектории выделены полужирным начертанием).

Как следует из рис. 1, элементы ТДК и связи между ними образуют связный направленный граф, следовательно, для построения основной техноло-

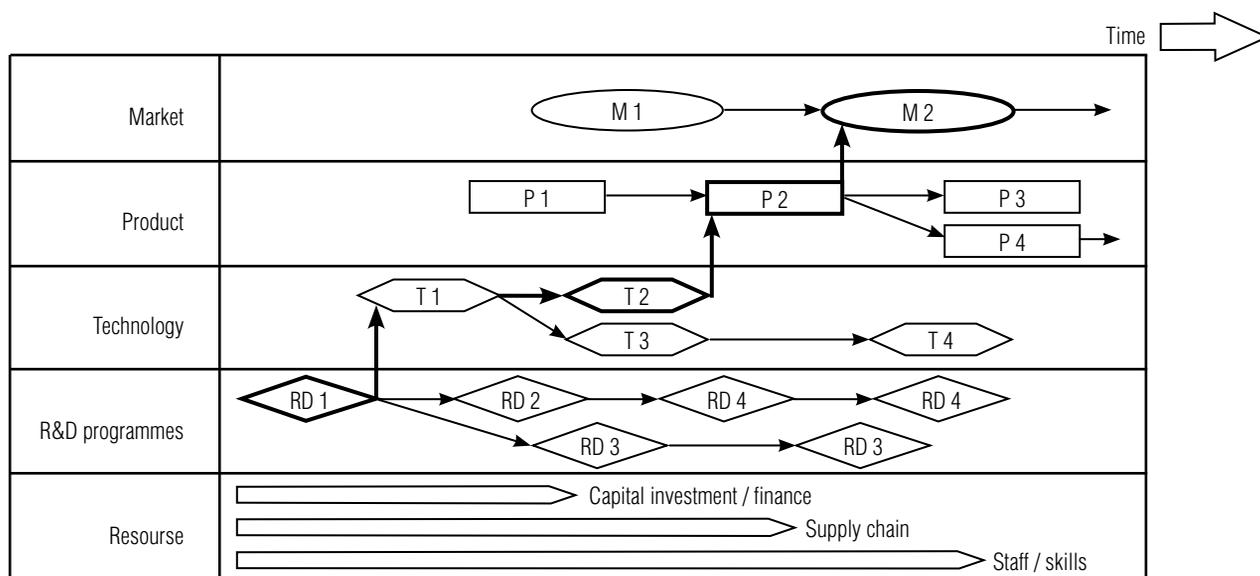


Рис. 1. Основная технологическая траектория на базовой схеме EIRMA

гической траектории могут быть применены алгоритмы расчета максимального потока от источника к стоку (с добавлением синтетических вершин источника и стока с неограниченными пропускными способностями), например, модифицированные алгоритмы Форда-Фалкерсона [5, 6].

Основные сложности при применении данных групп методов заключаются в необходимости определения весов ребер (дуг) графа – взаимно однозначных показателей, характеризующих отношения между элементами ТДК (от научных направлений к технологиям, от технологий к продуктам и от продуктов к рынкам).

Постановка задачи выглядит следующим образом.

Пусть заданы множества элементов ТДК (1), атрибутов элементов ТДК (2):

$$X = \{x_i\}, i = \overline{1, n} . \quad (1)$$

$$A = \{a_{ij}\}, i = \overline{1, n}, j = \overline{1, m} , \quad (2)$$

где n – количество элементов ТДК;

m – максимальное количество атрибутов элементов ТДК.

Значения атрибутов элементов ТДК формируются в результате сбора и консолидации экспертных оценок (3).

$$R = \{r_{ijk}\}, i = \overline{1, n}, j = \overline{1, m}, k = \overline{1, l} \quad (3)$$

где r_{ijk} – экспертная оценка значения j -го атрибута i -го элемента ТДК, представленная k -ым экспертом.

Необходимо для каждого элемента ТДК определить интегральный показатель применимости (полезности) Ψ_i , $i = \overline{1, n}$ с учетом отношений между элементами.

3. Использование многомерных функций применимости для оценки объектов комплексной природы

Элементы ТДК могут быть представлены как объекты, характеризующиеся набором атрибутов различной природы: технологической, экономической, политической и социальной. Постановка задачи по поиску многомерной функции оценки элемента ТДК определяется как поиск отображения вектора значений атрибутов элемента из множества допустимых альтернатив на n -мерное пространство последствий [7].

Таким образом, задача построения интегрального показателя применимости для элемента ТДК

может быть сведена к свертке вектора атрибутов в скалярное значение показателя применимости (4)

$$f(x_1, x_2, \dots, x_n) \geq f(x'_1, x'_2, \dots, x'_n) \leftrightarrow (x_1, x_2, \dots, x_n) \gg (x'_1, x'_2, \dots, x'_n), \quad (4)$$

где \gg – отношение предпочтительности, означающее, что набор значений (x_1, x_2, \dots, x_n) не менее предпочтителен, чем $(x'_1, x'_2, \dots, x'_n)$.

Данная функция может быть названа функцией предпочтения, оценки или применимости [8, 9].

В целях сокращения размерности функций применимости при расчете интегральных показателей применимости целесообразно реализовать независимость вариантов по предпочтению [10] и обеспечить сепарабельность при разбиении функции нескольких переменных на функции, определенные на непересекающихся подмножествах первоначальных переменных [11]. Достаточно простым способом построения функции применимости на многих критериях является использование аддитивной формы ее представления. Впервые это было показано Жераром Дебре в работе [12], где была предложена система аксиом, обеспечивающая существование аддитивной функции для трех и более критериев. Алгебраическое доказательство возможности построения аддитивных функций было дано в работе Льюса и Тьюки [13], рассмотревших случаи «совместных измерений» для случая двух критериев.

В случае полного декартова произведения $X_1 * X_2 * \dots * X_N$ Фишберн показал [14], что функция применимости может быть представлена в виде аддитивной функции только тогда, когда желательность любой лотереи, обозначаемой (x_1, x_2, \dots, x_n) , зависит только от маргинальных вероятностных распределений компонент и не зависит от их совместного распределения. В этом случае построение n -мерной функции применимости сводится к легко осуществляемому известными методами [15] построению n одномерных функций применимости.

Для формирования интегрального показателя применимости элемента ТДК в настоящей работе выбран подход многокритериальной теории полезности (MAUT – Multi-Attribute Utility Theory). В соответствии с этой теорией интегральная оценка объекта может быть получена как средневзвешенное оценок его атрибутов с учетом присущих им шкал [16].

В этих условиях атрибуты рассматриваются как измерения факторного пространства, а веса в агре-

гирующей функции задаются, исходя из степени значимости соответствующего атрибута для оценивающего (потребителя) [7, 17].

В теории MAUT для доказательства существования функции применимости в определенном виде применяется следующий набор аксиом, общих с классической теорией полезности:

а) аксиома полноты, в соответствии с которой принимается, что между применимостями любых элементов может быть установлено отношение превосходства или безразличия;

б) аксиома транзитивности: из превосходства применимости элемента X_i над применимостью элемента X_j и превосходства применимости элемента X_j над применимостью элемента X_k следует превосходство применимости элемента X_i над применимостью элемента X_k ;

в) Для соотношений между применимостями альтернатив X_i , X_j и X_k , имеющих отношения упорядоченности

$$f(X_i) > f(X_k) > f(X_j), \quad (5)$$

можно найти такие числа $0 \leq \alpha \leq 1$ и $0 \leq \beta \leq 1$, что:

$$\begin{aligned} \alpha \cdot f(X_i) + (1 - \alpha) \cdot f(X_k) &= f(X_j) \\ (1 - \beta) \cdot f(X_i) + \beta \cdot f(X_k) &> f(X_j) \end{aligned} \quad (6)$$

г) Независимость по предпочтению. Два атрибута X_i и X_j независимы по предпочтению от остальных атрибутов, если предпочтения между альтернативами, различающимися оценками по этим атрибутам, не зависят от константных значений по другим атрибутам. Это условие наиболее часто используется в задачах MAUT [14, 18].

Для использования в расчетах функции применимости элемента дорожной карты применяется основная теорема многокритериальной функции применимости в следующей формулировке: если условия независимости по применимости и независимости по предпочтению выполнены, то функция применимости является аддитивной или мультипликативной [7].

Современные исследования показали [19, 20, 21, 22, 23], что для практического применения целесообразно принять гипотезу о независимости атрибутов и аддитивной (средневзвешенного) формы представления агрегирующей функции в целях упрощения моделирования и повышения управляемости модели при оперировании экспертами значениями с некоторым уровнем неопределенности.

На текущий момент использование аддитивной функции применимости в задачах принятия решений в условиях неопределенности является наиболее распространенным подходом для оценки многомерных альтернатив [24, 25, 26].

4. Использование интегральных показателей применимости для оценки элементов ТДК

Канонической формой представления структуры ТДК является многослойная [4], в которой элементы группируются по слоям в соответствии с их принадлежностью к научно-технологической или бизнес-ориентированной стадии разработки и внедрения продукта. Такими слоями являются научные направления, технологии, продукты и рынки. Следовательно, потребителем при оценке элемента какого-либо слоя может выступать элемент следующего вышележащего слоя (рынок для продукта, продукт для технологии, и технология для научного направления).

Для получения интегральных показателей элементов ТДК парам классов «источник-потребитель» определяются функции применимости, значения которых используются в расчете интегральной оценки по MAUT, расширенной для оценки влияния аспектов атрибутов, а также специальных объектов ТДК слоя глобального контекста развития.

Для каждого значимого с точки зрения разработчика ТДК атрибута определенного класса элемента необходимо для получения значения частного показателя применимости определить функцию применимости с указанием класса потребителя

$$U = f(S, D, A), \quad (7)$$

где S – класс-источник, D – класс-потребитель, A – значение атрибута, а также поставить в соответствие вес ω для агрегации частных показателей в интегральный показатель применимости элемента ТДК.

Оценка такого сложного объекта, как элемент ТДК, сильно затруднена нелинейностью предпочтений, зависящих в первую очередь от точки зрения потребителя на качество интересующих его свойств элемента.

В целях имитации нелинейности предпочтений предусматривается разделение атрибутов по аспектному признаку: технологические (например, «Глубина крекинга»), экономические (например, «Текущая себестоимость»), политические (например, «Важность для технологического лидерства России»),

социальные (например, «Доля вновь созданных высокотехнологичных рабочих мест») и экологические (например, «Снижение доли серы в топливе»).

Помимо определения функций применимости с учетом класса-потребителя для учета взаимного влияния элементов ТДК применяется перебалансировка полученных значений показателей применимости с переносом применимости по установленным связям между элементами.

Перенос показателей применимости может происходить в соответствии с двумя типами эффектов, обусловленных природой элементов ТДК. В случае присутствия эффекта первого типа, называемого выталкивающим эффектом элемента, показатель применимости будет переноситься от нижележащих элементов, способствующих появлению перспективных продуктов. Степень переноса в этом случае будет определяться влиянием элемента-источника на вероятность создания продукта и улучшением его продуктовых свойств (изменению значений атрибутов в сторону увеличения потребительских свойств).

Второй тип переноса будет возникать в случае взаимодействия элементов, относящихся к рыночным предпочтениям и определяющим объемы спроса. В этом случае степень увеличения рейтинга применимости элемента с атрибутами потребительских свойств будет зависеть от степени их соответствия вышележащим элементам, определяющим ожидания потребителей и объемы рынка, потенциально осваиваемые этим элементом.

Коэффициенты переноса, определяющие масштаб эффекта, определяются параметрами связей между элементами ТДК. Объект типа связь имеет два обязательных атрибута: силу влияния и вероятность появления. Исходя из консолидированных значений этих атрибутов, заданных группой экспертов в предметной области, рейтинг элемента ТДК по применимости изменяется в ту или иную сторону на величину, зависящую также и от величины рейтинга связанного элемента.

При перебалансировке учитывается влияние глобального контекста дорожной карты на процессы создания применимости. Контекст может вносить корректировки не только на уровне значений атрибутов, но и в порядок расчета частных показателей применимости и весовых коэффициентов функций свертки, а также порядок переноса рейтинга применимости.

Основным источником поступления данных для формирования частных показателей применимо-

сти являются сведения, собираемые в ходе процедуры опроса экспертов предметной области по возможным значениям каждого атрибута каждого элемента ТДК. Учитывая массовость (несколько сотен экспертов), а также в целях облегчения для эксперта процедуры задания значений атрибутов, при решении задачи принимается допущение о взаимной независимости атрибутов одного элемента.

Инженер-разработчик дорожной карты, исходя из содержательного смысла атрибутов элемента ТДК, принимает решение о выборе аддитивной или мультипликативной (в исключительных случаях) формы функции применимости.

На первом этапе эксперт оценивает диапазон разброса значений атрибута. После оценки диапазона определяются значения функции применимости для всех точек изменения значений.

Максимальное значение функции для нормирования принимается равным единице, минимальное — нулю. Эксперт определяет зависимость направления роста функции применимости от направления шкалы атрибута и в соответствии с этим назначает максимальное и минимальное значения функции применимости на левую и правую границы диапазона. После этого для нахождения промежуточных точек итеративно проводятся типовые лотереи для нахождения детерминированного эквивалента. Для упрощения принятия решения эксперту может предъявляться ряд значений атрибута с вопросом о положении детерминированного эквивалента относительно этих значений.

После того, как инженер-разработчик дорожной карты определил функцию применимости, которая описывает упорядочение по предпочтению всех многомерных альтернатив, последующий анализ проводится с рассмотрением множества практически достижимых оценок по используемым критериям и выбором альтернатив, обладающих лучшими оценками.

Построение функции применимости для отдельного атрибута выполняется следующим образом. При задании метаданных атрибута должны быть указаны следующие параметры:

- ◆ вид шкалы (номинальная, ранговая, метрическая);
- ◆ реперные точки.

Реперные точки определяют интервалы, в которых ценность элемента по данному атрибуту имеет характерное поведение и изменяется с постоянной скоростью. В случае использования метрической шкалы необходимо задать значения показателя

применимости не менее чем для двух крайних точек в области определения атрибута (минимум и максимум). По двум крайним точкам задается линейное отображение значений атрибута на область допустимых значений показателя применимости.

Если для метрической шкалы задаются дополнительные реперные точки, отображение значений атрибута на интервал показателя применимости принимает кусочно-линейный вид.

При сборе фактических экспертных сведений по атрибуту с метрической шкалой значения атрибута, представленные экспертом, подставляются в вышеуказанные отображения, соответствующие интервалу области определения, и интерполируются в пределах указанного интервала.

При сборе сведений по данному атрибуту от всех экспертов, принимающих участие в опросе, и расчете значений функции применимости атрибута для каждого экспертного мнения выполняется обобщение экспертных сведений и вычисляется консолидированное значение данного атрибута. В силу требования монотонности функции применимости значение показателя применимости корректно рассчитывается как средневзвешенное с учетом компетенций экспертов. В случае нарушения монотонности функции инженер-разработчик дорожной карты должен рассмотреть соответствующий атрибут для выявления в нем смешанного влияния различных характеристик элемента ТДК.

Полученное значение частного показателя применимости изменяется с учетом корректировочных коэффициентов, соответствующих настройкам значимости аспектов атрибутов, а также коэффициентов влияния объектов глобального контекста развития. По окончании корректировки частных показателей применимости для всех атрибутов элемента ТДК, участвующих в расчете интегрального показателя, выполняется свертка частных показателей в интегральный показатель применимости элемента ДК. Для интегральных показателей выполняется перебалансировка в соответствии со структурой связей элементов ТДК и настроек переноса показателей применимости, после чего формируется основная технологическая траектория.

В виду того, что в данном алгоритме расчета основной технологической траектории присутствует значительное количество взаимовлияющих параметров, определяющих порядок получения окончательных результатов, модель является в значительной степени гибкой. Модель может быть ис-

пользована аналитиками предметной области для обработки результатов в интерактивном режиме и проведения сценарного анализа сложных систем со значительным количеством взаимовлияющих объектов. Другим полезным качеством данной модели является возможность проверки большого объема данных на непротиворечивость с учетом ограничений, соответствующих физическому смыслу моделируемых объектов. Например, в случае появления элемента ТДК на временной оси раньше связанного с ним элемента, от которого зависит реализация, аналитик может сделать вывод о несогласованности мнений экспертов и провести повторный сбор сведений с уточнением этого вопроса.

5. Заключение

В работе представлен подход и алгоритмическая реализация способа расчета частных и интегральных показателей применимости элементов технологической дорожной карты и их атрибутов.

В рамках работы предлагается использовать для оценки элементов ТДК многомерные функции применимости. Для формирования интегрального показателя применимости элемента применена многокритериальная теория полезности MAUT, агрегирующая функция представлена в виде аддитивной свертки.

В качестве дополнительного метода повышения точности интегральных показателей применимости предложена перебалансировка показателей применимости с переносом применимости по установленным связям между элементами.

Также в составе работы предложен способ консолидации экспертных сведений, базирующийся на построении частной функции применимости и определении реперных точек, в которых частная функция применимости меняется нелинейно. Описанные научно-технологические решения реализованы в составе программного комплекса «Интерактивная дорожная карта с обратной связью».

Программный комплекс предназначен для формирования долгосрочных стратегий технологического и инновационного развития секторов экономики, корпораций, крупных компаний, а также комплексных сценариев развития рынков, продуктов и продуктовых групп.

Предложенный подход к оценке элементов ТДК является одним из важных конкурентных преимуществ комплекса. ■

Литература

1. Инвестиционная деятельность Роснано – дорожные карты. [Электронный ресурс] // Роснано: [сайт]. URL: <http://www.rusnano.com/investment/roadmap/oil> (дата обращения: 08.05.2013)
2. Albright R.E., Kappel, T.A. Roadmapping in the corporation // *Research Technology Management*. – 2003. – 42 (2). – P. 31-40.
3. Phaah R., Farrukh C., Probert D.R. Technology roadmapping – A planning framework for evolution and revolution // *Technological Forecasting and Social Change*. – 2004. – Vol. 71, Issues 1-2. – P. 5-26.
4. European Industrial Research Management Association (EIRMA). Working Group report, Issue 52. Technology Roadmapping: Delivering Business Vision. – Paris: EIRMA, 1997.
5. Ford L.R., Fulkerson D.R. Flows in networks. – Princeton University Press, 1962.
6. Ceder A. Public transit planning and operation: theory, modeling and practice. – Oxford, UK: Elsevier, Butterworth-Heinemann, 2007.
7. Кини Р., Райфа Х. Принятие решений при многих критериях: замещения и предпочтения / пер. с англ. – М.: Радио и связь, 1981.
8. Фишберн П. Теория полезности для принятия решений. – М.: Наука, 1978.
9. Aleskerov F., Monjardet B. Utility Maximization, Choice and Preference. – Springer Verlag, Heidelberg, 2002.
10. Райфа Г. Анализ решений. Введение в проблему выбора в условиях неопределенности. – М.: Наука, 1977.
11. Leontief W. *Econometrica*. – Issue 15, № 4 (October 1947). – P. 361-373.
12. Debreu G., *Topological Methods in Cardinal Utility Theory / Mathematical Methods in the Social Sciences /* K.J.Arrow, S. Karlin, P.Suppes eds. – Stanford, CA: 1959.
13. Luce R.D., Tukey J.W. Simultaneous conjoint measurement: a new type of fundamental measurement // *Journal of Mathematical Psychology*. – 1964. – №1. – P.1-27.
14. Fishburn P.C. Independence in Utility Theory with Whole Product Sets // *Operations Research*. – January-February 1965. – Issue 13, № 1. – P. 28-45.
15. Schlaifer R.O. *Analysis of Decisions Under Uncertainty*. – NY: McGraw-Hill, 1967.
16. Winterfeld D., Edwards W. *Decision Analysis and Behavioral Research*. – Cambridge, England: Cambridge University Press, 1986.
17. Zeleny M. *Linear Multiobjective Programming*. – Verlag, Berlin: Springer, 1974.
18. Humphreys P.C. Application of multiattribute utility theory // H.Jungerman and G. de Zeeuw (Eds.). *Decision making and change in human affairs*. – Dordrecht: Reidel, 1977.
19. Bauer M., Gmytrasiewicz P., Vassileva J. Proceedings of the UM2001 Workshop on Machine Learning for User Modeling. – Springer, 2001.
20. Bouyssou D., Jacquet-Lagrèze E., Perny P., Slowinski R., Vanderpooten D., Vincke Ph. *Aiding Decisions with Multiple Criteria: Essays in Honour of Bernard Roy*. – Kluwer Academic Publishers, Dordrecht, 2001.
21. Chin D., Porage A. Acquiring User Preferences for Product Customization. // Proceedings of the 8th International Conference, UM, 2001.
22. Colorni A., Paruccini M., Roy B. A-MCD-A: Multiple Criteria Decision Aiding. Joint Research Center. – Luxembourg: The European Commission, 2001.
23. Schütz W., Meyer M. Definition einer Parameter-Hierarchie zur Adaptierung der Benutzer-Interaktion in E-Commerce-Systemen // Proceedings of ABIS 2001. – 2001.
24. Schäfer R. Rules for Using Multi-Attribute Utility Theory for Estimating a User's Interests // CAWICOMS Consortium. Deliverable D01 – Requirements, Application Scenarios, Overall architecture and Test Specification, 2000.
25. Ehrgott M., Gandibleux X. *Multiple Criteria Optimization. State of the Art Annotated Bibliographic Surveys*. – Kluwer Academic Publishers, Dordrecht, 2002.
26. Trzaskalik T., Michnik J. *Multiple Objective and Goal Programming: Recent Developments*. – Physica Verlag, Heidelberg, 2002.

УПРАВЛЕНИЕ ТРЕБОВАНИЯМИ ПРИ РЕАЛИЗАЦИИ ИТ-ПРОЕКТОВ

Т.К. Кравченко,

доктор экономических наук, профессор, заведующий кафедрой бизнес-аналитики
Национального исследовательского университета «Высшая школа экономики»

E-mail: tkravchenko@hse.ru

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

В статье предлагается подход к моделированию процесса управления требованиями при реализации ИТ-проектов. Модель управления требованиями реализуется в три этапа. На первом этапе отбираются бизнес-требования, которые фиксируются в проектном решении и являются рамками проекта. На втором этапе строится модель, связанная с принятием, отклонением, уточнением или отнесением к дополнительным работам каждого поступающего требования пользователей. На третьем этапе для принятых требований пользователей формируются приоритеты системных требований, на основе которых далее составляются планы работ по проекту. При построении моделей принятия решений используются методы анализа иерархий и аналитических сетей Т.Л. Саати и СИПР SuperDecisions.

Ключевые слова: управление требованиями, моделирование процесса управления требованиями, метод анализа иерархий, метод аналитических сетей, система поддержки принятия решений SuperDecisions.

1. Введение

Во многих сферах бизнеса высоки риски убытков по причине прекращения либо приостановки проектов. В рамках проекта важно соблюдать баланс между качеством выполнения, сроками работ и их стоимостью, что ставит задачу поиска компромиссного решения для соблюдения баланса между данными характеристиками.

Аналитическая работа в проектной деятельности в сфере информационных технологий (ИТ) во многом связана с процессом управления требова-

ми, поступающими от заказчика. Соответственно, для того, чтобы цели проекта были достигнуты, стратегически важно правильно построить процесс управления требованиями, включая взаимодействие между членами проектной команды, а также согласовать процедуры отслеживания выполненных работ.

Наиболее характерные подходы в этом плане демонстрирует ИТ-сфера, для которой проектная деятельность является важнейшей составляющей бизнеса. Примерами ИТ-проектов являются: разработка и развитие программного обеспечения,

внедрение информационных систем, инфраструктурные и организационные проекты.

Для таких проектов характерна высокая интенсивность и итерационность выполнения работ с подробным составлением календарно-сетевых графиков. Выработка требований и управление требованиями – важнейшие элементы успеха любого ИТ-проекта. Более половины потерь времени организациями – разработчиками программного обеспечения связаны с неэффективностью подхода к управлению требованиями.

2. Роль бизнес-аналитика при реализации ИТ-проектов

Роль бизнес-аналитика в проекте внедрения информационной системы заключается в разработке непротиворечивой модели требований бизнеса в рамках и терминах внедряемой системы, коммуникация клиентских требований разработчикам и обратно, презентация клиентам реализованной функциональности.

Эффективное управление требованиями подразумевает как их учет на стадии поступления и отслеживание статуса выполнения, так и принятие решений о приоритетности, критичности и сроках выполнения поступающих требований.

Приведем определение процесса управления требованиями: это процесс, включающий идентификацию, выявление, документацию, анализ, отслеживание, приоритизацию требований, достижение соглашений по требованиям и затем управление изменениями и уведомление заинтересованных лиц. Управление требованиями – непрерывный процесс на протяжении всего жизненного цикла продукта [1].

В качестве требования рассматривается любое условие, которому должна соответствовать разрабатываемая система. Требованием может быть функциональность системы или ограничение, которому система должна удовлетворять.

В рамках *IEEE Standard Glossary of Software Engineering Terminology* [2] понятие требования трактуется шире:

1. условия или возможности, необходимые пользователю для решения проблем или достижения целей;

2. условия или возможности, которыми должна обладать система или системные компоненты, чтобы выполнить контракт или удовлетворять стандартам, спецификациям или другим формальным документам;

3. документированное представление условий или возможностей для пп. 1 и 2.

Первичные данные, поступающие от клиентов, характеризуются противоречивостью, неполнотой, нечеткостью, изменчивостью. Требования нужны, в частности, для того, чтобы команда исполнителей могла определить и согласовать с заказчиками временные и финансовые перспективы проекта. Поэтому значительная часть требований должна быть собрана и обработана на ранних этапах определения дизайна внедряемой системы.

Однако собрать на ранних стадиях все необходимые данные удается только в исключительных случаях, так как заказчик не всегда способен оценить все нюансы и особенности предстоящей работы в новой системе. На практике процесс сбора, анализа и обработки требований зачастую нетривиален, растянут во времени на протяжении всего проекта и включает доработки функциональности, выведенной на стадию поддержки продуктивной эксплуатации.

Бизнес аналитик, занимаясь управлением требованиями, должен быть способен предоставлять их в виде, необходимом для правильного восприятия аудиторией. Следовательно, требования должны быть сформулированы на языке, понятном как пользователям, так и разработчикам. Для каждого требования необходимо контролировать его статус, сохранять историю изменений, отслеживать смежные требования и использовать уже полученные наработки для реализации новых поступающих требований.

Таким образом, необходимо фиксировать поступающие требования с указанием таких атрибутов как дата получения, автор, а также направление, к которому это требование относится. Также необходимо проверять, соответствует ли требование согласованным рамкам проекта.

Формат, в котором фиксируются требования, должен быть понятен для всех заинтересованных лиц (стейкхолдеров), которым необходимо ознакомиться со списком требований и подтвердить его. Поэтому в задачи бизнес аналитика входит документирование требований в четком, понятном и оправданно детализированном виде, чтобы обеспечить понимание и эффективную коммуникацию между членами команды [3].

Рассмотрим процесс управления требованиями на примере проекта внедрения модуля «Казначейство» системы SAP (SAP Treasury) [4].

3. Классификация требований при внедрении модуля «Казначейство» системы SAP

Основной целью проекта внедрения модуля «Казначейство» системы SAP (полное название – «SAP Treasury and Risk Management, SAP TRM») является создание централизованной системы управления казначейством, которая заменит собственные разработки и будет интегрирована с другими системами, уже используемыми в компании. Для достижения данной цели необходима автоматизация казначейских операций.

При помощи модуля SAP TRM осуществляется управление финансовыми операциями и позициями, а также отношениями с контрагентами (бизнес-партнерами), от процессов ввода и расчета сделок до учета деривативов, операций на денежном и валютном рынках, сделок с кредитными инструментами, операций с ценными бумагами и (внутригрупповыми) займами, а также документарных операций.

В рамках проекта предусматриваются такие этапы работ, как запуск и проектирование, реализация и

заключительная подготовка, внедрение. В ходе планирования очередного этапа работ уточняются цели этапа, состав и контролируемые результаты работ.

Приведем схему классификации требований по различным признакам: по уровням (бизнес требования, требования пользователей, системные требования), по содержанию (функциональные, нефункциональные), по источникам (внутренние или внешние относительно блока функциональности) и по характеру (требования к доработке или сообщения об ошибках) (рис. 1). Данная классификация будет использоваться при построении моделей управления требованиями проекта.

4.1. Классификация по уровням требований

Рассматривая требования в разрезе этапов жизненного цикла разработки системы, можно выделить следующие уровни требований, поступающих от заказчика.

Требования верхнего уровня, – **бизнес требования**, – формулируются руководством компании-

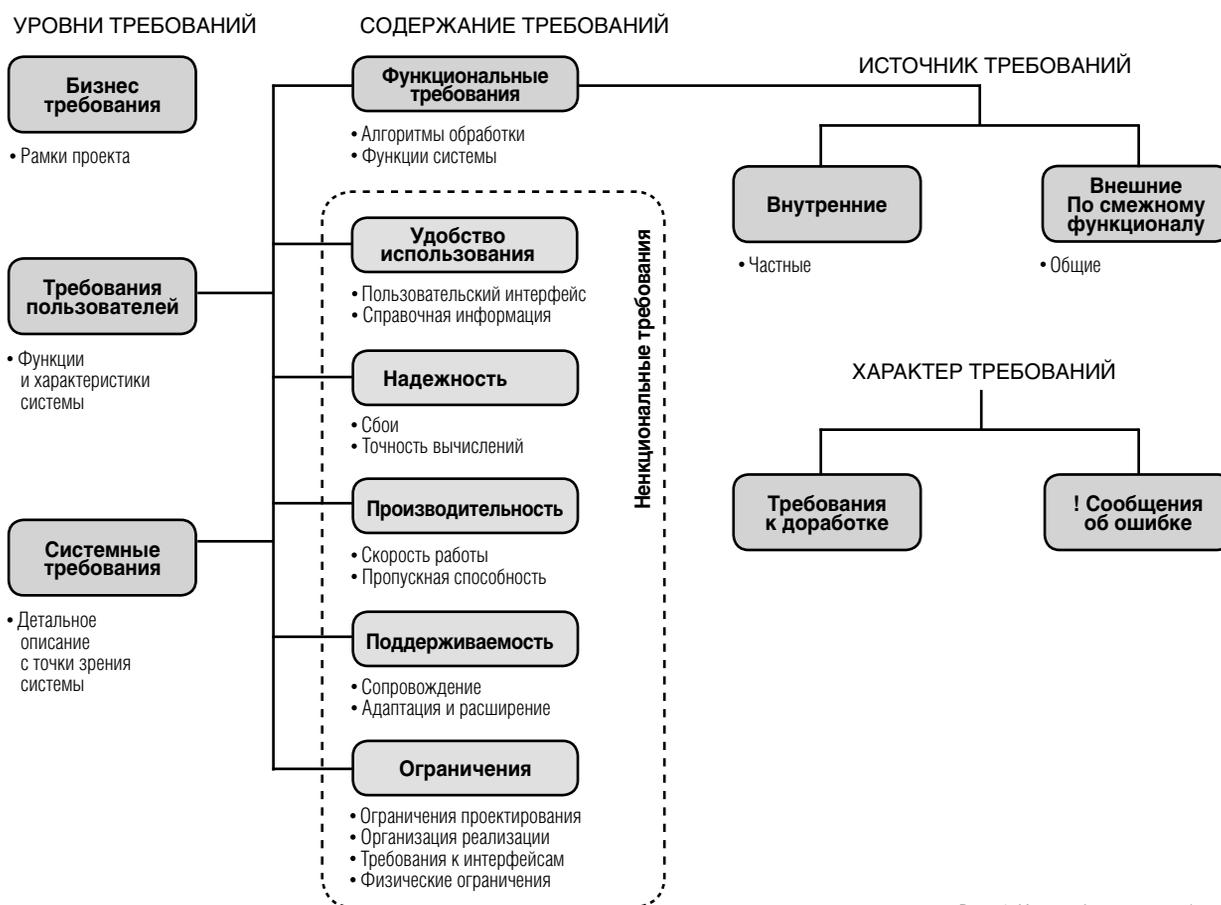


Рис. 1. Классификация требований

заказчика на стадии разработки проектного решения для выбранной системы. Такие требования включают определение укрупненных функциональных модулей, которые должны поддерживать ведение основных бизнес-процессов. Соответственно, этап анализа бизнес-требований закладывает фундамент для проектирования и реализации системы.

Следующий уровень – это уровень **требований пользователей**. Требования пользователей отражают их пожелания к функциям и характеристикам создаваемой системы. Требования могут включать перечисления действий пользователя и откликов системы в альтернативных сценариях, а также содержать характеристики, которым должна удовлетворять система для реализации этих сценариев. Сюда входит организация удобных диалоговых средств и доработка методов расчета.

Требования, формулируемые пользователями, часто бывают недостаточно структурированными, дублирующимися, противоречивыми. Поэтому задача бизнес-аналитика состоит в том, чтобы преобразовать их в требования к функциональности системы, на основе которых будут производиться настройки стандарта и составляться технические задания для разработчиков.

Для внедрения системы важен и третий уровень – **функциональный или системный**, на котором осуществляется формализация требований пользователей. Каждое требование пользователя может стать основой для формулирования нескольких системных требований.

4.2. Классификация по содержанию требований

Поступающие требования могут касаться как алгоритмов работы системы, так и различных нефункциональных характеристик, которыми должна обладать система. Таким образом, по содержанию требования можно разделить на две основные категории: функциональные и нефункциональные.

Функциональные требования относятся к основным свойствам / функциям системы. Ряд требований касается поддержки работы пользователей, это требования к инструментам отслеживания действий пользователей и записи в журнал безопасности конкретных типов событий. К функциональным требованиям также относятся требования к средствам отслеживания, приобретения, установки и контроля использования лицензий

системы; средствам поддержки различных языков; инструментам создания и получения отчетов; средствам контроля доступа к определенным информационным ресурсам; средствам управления приложениями в распределенной среде; поддержки документооборота.

К **нефункциональным требованиям** относятся требования к удобству использования, надежности и производительности системы, возможности поддержки. К различным ограничениям относятся ограничения проектирования, реализации, разработки, написания программного кода и др.

4.3. Классификация по источникам требований

В процессе реализации различных блоков проекта, поступающие требования можно разделять на **частные** (полученные по доработкам, касающимся конкретного блока, и не влияющие на смежную функциональность) и **общие** (требования, которые касаются обработки нескольких блоков функциональности системы). При выполнении общих требований необходимо учитывать, насколько они совместимы с частными требованиями каждого из отдельных функциональных блоков, для которых реализуется общее требование.

Подобное разделение можно также описать с точки зрения источника получения требования. Для отдельно взятого функционального блока требования могут быть **внутренними** (они же частные), полученными от пользователей в процессе разбора вопросов касающихся данного блока. Также требования могли поступить из смежных блоков (**внешние**), если в процессе работ по смежному блоку было получено какое-либо общее требование.

4.4. Классификация по характеру требований

Приоритет обработки требований зависит от того, является ли полученное требование замечанием к доработке или сообщением об ошибке. Сообщения об ошибках на этапах интеграционного теста и поддержки продуктивной эксплуатации имеют наивысший приоритет и должны устраняться в максимально короткий срок.

Большое количество ошибок требует дополнительного использования ресурсов и может задерживать основной ход работ. Поэтому необходимо учитывать потенциальные риски возникновения

ошибок в работе, когда требование принимается к исполнению. Например, риск дальнейших ошибок будет выше, если требование касается общих доработок по нескольким функциональным блокам, по которым ведутся параллельные работы. Изменение алгоритмов обработки имеет более высокий риск, чем доработка пользовательских интерфейсов и других нефункциональных требований.

5. Разработка модели управления требованиями при реализации финансового инструмента «Депозит на расчетном счете»

Модель управления требованиями реализуется в три этапа. На первом этапе отбираются бизнес требования, которые фиксируются в проектном решении и являются рамками проекта (табл. 1).

Решение о принятии или отклонении бизнес требования необходимо принимать для каждого бизнес требования, поступающего на этапе запуска/проектирования. Бизнес требование оценивается с позиций следующих признаков: выполнимость в рамках функциональности, гарантии успешного выполнения, критичность в бизнес-процессе заказчика и соответствие стандартам организации. Признаки для оценки бизнес требований не связаны между собой, поэтому на первом этапе исполь-

Таблица 1.

Список бизнес требований для инструмента «Депозит на расчетном счете»

Код	Формулировка
B1	ВНЕСЕНИЕ ДЕПОЗИТА НА РАСЧЕТНЫЙ СЧЕТ В СПИСОК ОБЯЗАТЕЛЬНЫХ ДЛЯ РЕАЛИЗАЦИИ ФИНАНСОВЫХ ИНСТРУМЕНТОВ В целевой системе необходимо реализовать финансовый инструмент «Депозит на расчетном счете» как один из основных инструментов используемых казначейством.
B2	РАСЧЕТ ПРОЦЕНТОВ В ЗАВИСИМОСТИ ОТ СРЕДНЕМЕСЯЧНОЙ СУММЫ ОСТАТКА Расчет начисляемого процента осуществляется ежемесячно на средний внутримесячный остаток на расчетном счете. Процентная ставка различается в зависимости от размера внутримесячного среднего остатка на расчетном счете.
B3	ВЫДЕЛЕНИЕ ДЕПОЗИТА СВИП КАК ОТДЕЛЬНОГО ИНСТРУМЕНТА Выделение депозита СВИП как отдельного финансового инструмента, поскольку суммы остатков по счетам СВИП, в отличие от расчетных счетов, содержатся в банковской выписке.
B4	АВТОМАТИЧЕСКИЙ РАЗБОР БАНКОВСКОЙ ВЫПИСКИ Для инструментов депозитов, депозитов на расчетном счете, депозитов СВИП должен быть реализован автоматический разбор банковской выписки, как для основных использующихся казначейством инструментов.

зуется метод анализа иерархий (МАИ) [5] и система SuperDecisions (www.superdecisions.com). Приведем пример построения модели для обработки бизнес требования B2 «Расчет процентов в зависимости от среднемесячной суммы остатка» (рис. 2).

В результате проведенных исследований бизнес требования B1, B2, B4 принимаются и должны быть отражены в проектном решении, для того, чтобы оценивать дальнейшие требования, поступающие от пользователей C1, C2, C3, C4, C5, C6 (табл. 2).

Таблица 2.

Список требований пользователей для инструмента «Депозит на расчетном счете»

№	Формулировка
C1	СОЗДАНИЕ ТАБЛИЦЫ ПРОЦЕНТНЫХ СТАВОК / ВОЗМОЖНОСТЬ ВЕДЕНИЯ ТАБЛИЦЫ ПОЛЬЗОВАТЕЛЯМИ Необходимо вести соответствие значений среднемесячной суммы остатка и процентных ставок. Ведение данных значений должно производиться пользователями системы. Установленные значения процентных ставок могут изменяться, необходимо указывать сроки действия.
C2	РАЗДЕЛЕНИЕ ОБРАБОТКИ НА ДВЕ ПРОГРАММЫ В бизнес-процессе различаются депозиты на расчетных счетах и депозиты на счетах СВИП. Так как остатки по счетам СВИП содержатся в банковской выписке, обработку сделок СВИП необходимо проводить при разборе выписки. Для депозитов на расчетных счетах в выписке содержится только значение начисленных процентов в конце месяца, поэтому ежедневная обработка сумм остатков должна проводиться отдельной программой.
C3	ПОВТОРНАЯ ОБРАБОТКА ПОТОКОВ Значения остатков на расчетных счетах может изменяться. Необходимо предусмотреть возможность запуска программы пользователями для изменения сумм потоков в сделках депозитов на расчетных счетах, а также изменение сумм потоков для депозитов СВИП при повторном разборе позиций по выписке.
C4	АВТОМАТИЧЕСКОЕ СОЗДАНИЕ СДЕЛКИ С УЧЕТОМ УСЛОВИЙ ДОГОВОРА Даты начала и конца среднемесячных сделок фиксируются в договоре с банком и могут быть разными для различных договоров. Необходимо предусмотреть возможности настройки вариантов (начисление на последний/первый день месяца или сдвиг на первый рабочий день) для договоров депозитов на расчетных счетах. Создание новых сделок следует осуществлять путем добавления ежедневных сумм при наступлении даты конца сделки по условиям договора.
C5	ОТРАЖЕНИЕ ФАКТИЧЕСКИХ ВЫПЛАТ В ОТЧЕТЕ ПО ДЕПОЗИТАМ БЕЗ УЧЕТА РЕЗУЛЬТАТА РАЗБОРА ВЫПИСКИ Для отчета по доходности депозитов реализовать отражение суммы фактических выплат за период в соответствии с указанными в условиях сделок планами погашения процентов, без учета результатов разбора банковской выписки.
C6	УДАЛЕНИЕ СТОРНИРОВАННЫХ СДЕЛОК ИЗ ДОГОВОРА В случае сторнирования ошибочно созданных сделок следует также удалять привязку сделок к договорам.

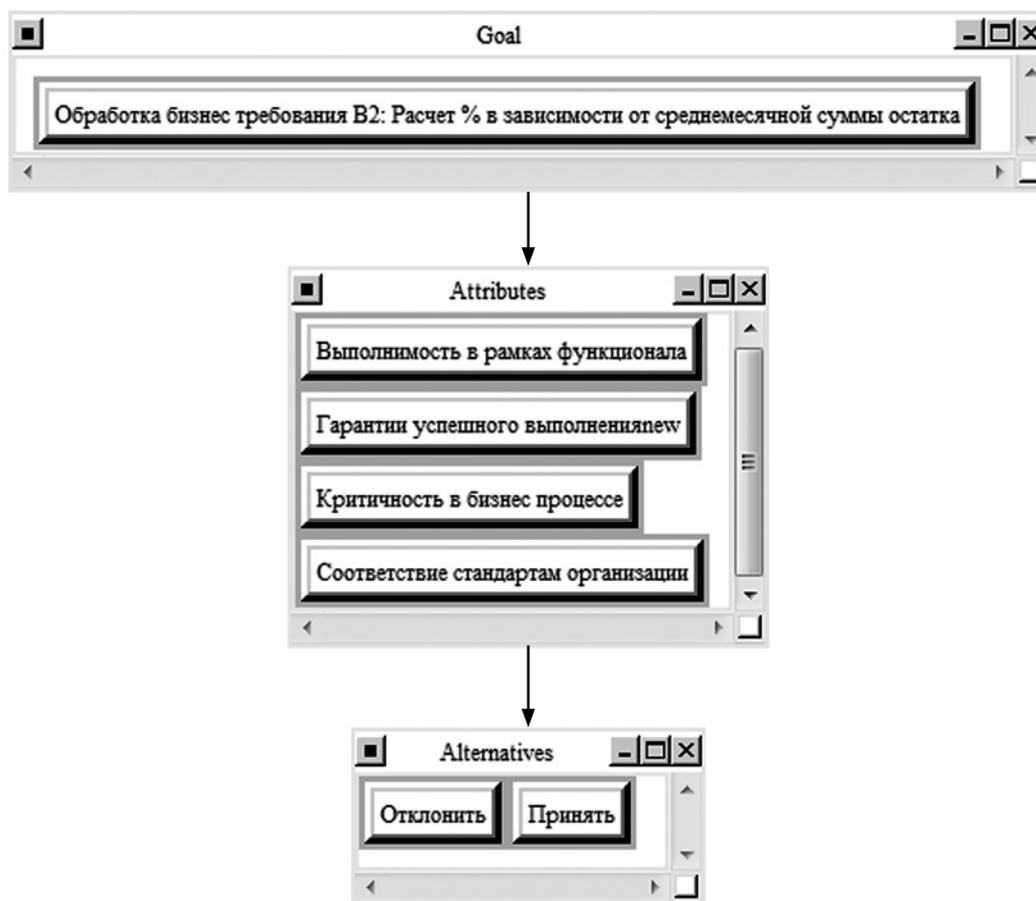


Рис. 2. Модель принятия решения относительно бизнес требования В2

На втором этапе строится модель, позволяющая принимать решения о принятии, отклонении, уточнении или отнесении требования к дополнительным работам, для каждого поступающего требования пользователей в отдельности.

Следует отметить, что не все бизнес требования могут оказывать влияние на принятие решения относительно конкретного требования пользователя. Так требование С5 «Отражение фактических выплат в отчете без учета разбора выписки» связано лишь с бизнес требованиями В2 и В4. Поэтому на данном этапе учитывается взаимосвязь требования пользователя С5 с соответствующими бизнес требованиями. Аналогично устанавливаются связи всех остальных требований пользователей с теми или иными бизнес требованиями.

В процессе принятия решения относительно каждого конкретного требования пользователя находятся приоритеты соответствующих бизнес требований с учетом следующих признаков: выполнимость в рамках функциональности, гарантии

успешного выполнения, критичность в бизнес-процессе, соответствие стандартам организации.

Данные приоритеты учитываются при обработке требования пользователя в соответствии со следующими признаками: важность автора требования, дата поступления, было ли требование заявлено пользователями при проектировании, критичность для работы в системе, риск возникновения дальнейших ошибок.

Признаки для оценки бизнес требований связаны с признаками требований пользователей, последние также связаны между собой. Поэтому для построения модели принятия решения на втором этапе в системе SuperDecisions используется метод аналитических сетей (МАС) [6]. Приведем пример построения модели для обработки требования пользователя С5 «Отражение фактических выплат в отчете без учета разбора выписки» (рис. 3).

Таким образом, в результате проведенного исследования требования пользователей С1, С3, С4 принимаются. На их основе будут сформулированы системные требования.

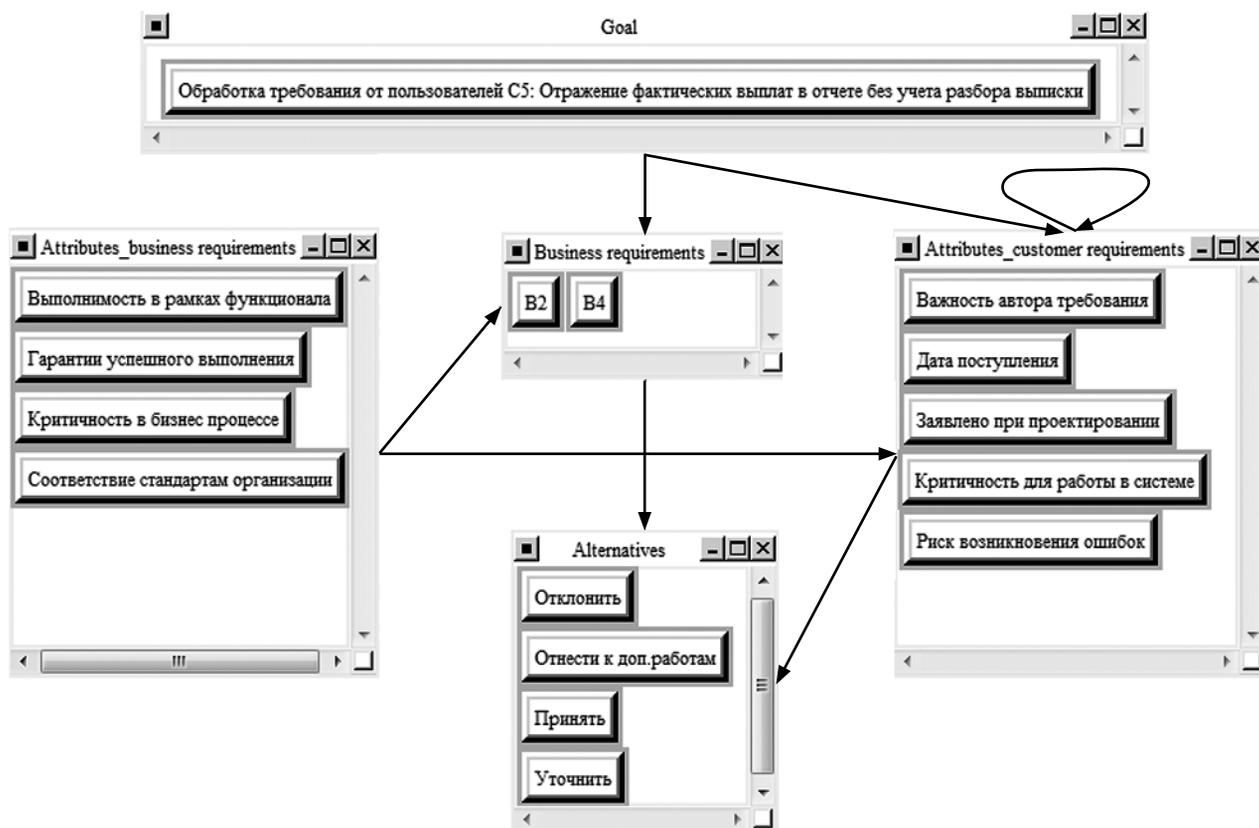


Рис. 3. Модель принятия решения относительно требования пользователя С5

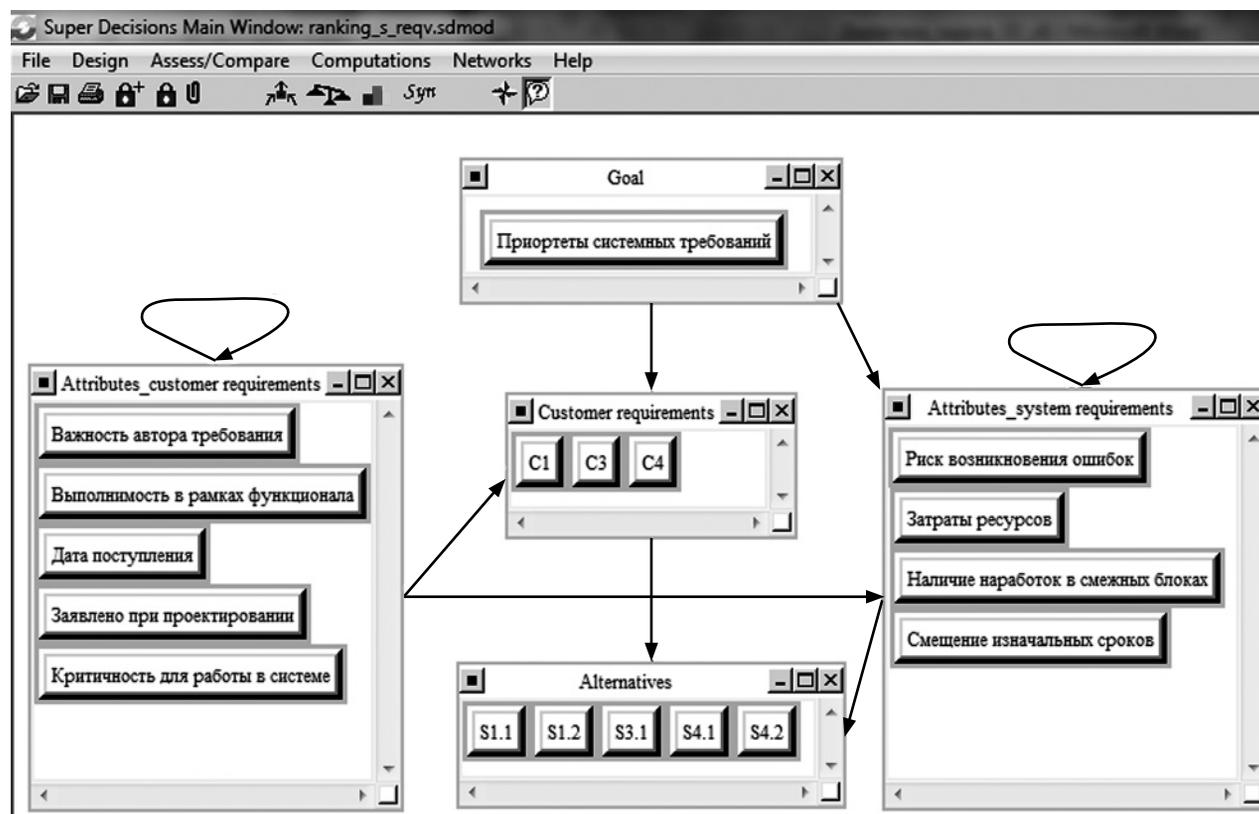


Рис. 4. Модель принятия решения относительно системных требований

Таблица 3.

Список возможных системных требований для принятых требований пользователей для инструмента «Депозит на расчетном счете»

№	Формулировка требования пользователя	№	Формулировка системного требования
С1	СОЗДАНИЕ ТАБЛИЦЫ % СТАВОК/ВОЗМОЖНОСТЬ ВЕДЕНИЯ ТАБЛИЦЫ ПОЛЬЗОВАТЕЛЯМИ Необходимо вести соответствие значений средне-месячной суммы остатка и процентных ставок. Ведение данных значений должно производиться пользователями системы. Установленные значения % ставок могут изменяться, необходимо указывать сроки действия	S1.1	СОЗДАНИЕ ТАБЛИЦЫ В СИСТЕМЕ Создать таблицу ZTTD_EXT_008_PRC для хранения значения % ставок в разрезе БЕ, ID карточки договора, срока действия и границ средне-месячной суммы
		S1.2	СОЗДАНИЕ ВЫЗОВА ТАБЛИЦЫ ИЗ КАРТОЧКИ ДОГОВОРА Таблицу ZTTD_EXT_008_PRC раскрывать, основываясь на карточке договора Депозита на Расчетный счет с позиционированием на строках, относящихся к текущей карточке и настоящему промежутку времени
С3	ПОВТОРНАЯ ОБРАБОТКА ПОТОКОВ Значения остатков на расчетных счетах может изменяться. Необходимо предусмотреть возможность запуска программы ZTRM_CRT_DEAL_F_MEMO пользователями для изменения сумм потоков в сделке депозита на Расчетный счет, а также предусмотреть изменение сумм потоков для депозита Свип при повторном разборе позиций по выписке	S3.1	ОБНОВЛЕНИЕ ПОТОКОВ В СДЕЛКАХ, ПРИ ПОВТОРНОМ ПРОГОНЕ ОБРАБОТКИ При обработке сумм остатков необходимо анализировать сделку на наличие уже добавленных потоков на дату. Если поток размещения на дату обработки уже присутствует в сделке, необходимо обновлять сумму и не создавать новый поток
С4	АВТОМАТИЧЕСКОЕ СОЗДАНИЕ СДЕЛКИ С УЧЕТОМ УСЛОВИЙ ДОГОВОРА Даты начала и конца среднемесячных сделок прописываются в договоре с банком и могут быть разными для различных договоров. Необходимо предусмотреть возможности настройки вариантов (начисление на последний/первый день месяца/сдвиг на первый рабочий день) для договоров Депозитов на Расчетный счет. Создание новых сделок производить обработкой добавления ежедневных сумм при наступлении даты конца сделки по условиям договора.	S4.1	ДОБАВЛЕНИЕ ПАРАМЕТРОВ СМЕЩЕНИЯ ДАТ НАЧАЛА (КОНЦА) СДЕЛОК НА КАРТОЧКУ ДОГОВОРА Для вида договора Депозит на Расчетный счет и Депозит Свип необходимо добавить параметры запуска: смещение даты окончания сделки (1 – нет смещения/ 2 – на следующий рабочий/ 3 – на предыдущий рабочий день)
		S4.2	СОЗДАНИЕ НОВОЙ СДЕЛКИ ПРИ ОБРАБОТКЕ ПОСЛЕДНЕГО ДНЯ СДЕЛКИ ТЕКУЩЕГО МЕСЯЦА При обработке сумм остатков необходимо анализировать дату обработки - если она приходится на дату окончания сделки, в соответствии с параметрами смещения, указанными в карточке договора необходимо создавать сделку на следующий месяц в привязке к карточке договора. Также при создании новой сделки должна происходить фиксация процента - т.е. расчет среднемесячной суммы и добавления значения процента в структуру

Рассмотрим системные требования, сформулированные на основе принятых требований пользователей (см. табл. 3). Для одного требования пользователя может быть сформулировано одно и более системных требований.

Ранжироваться будут все сформулированные системные требования, полученные из принятых на данный момент требований пользователей, они же будут выступать в роли альтернатив на третьем этапе решения задачи, в данном случае S1.1, S1.2, S3.1, S4.1, S4.2.

При ранжировании системных требований (рис. 4) учитывается относительная значимость требований пользователей, на основе которых они были сформулированы. Поэтому на данном этапе находятся приоритеты требований пользователей с использованием указанных выше признаков их сравнения. Для нахождения приоритетов системных требований

учитываются следующие признаки: затраты ресурсов, необходимых для реализации; наличие наработок в смежных функциональных блоках; риск возникновения ошибок в работе системы; были ли смещены сроки реализации, относительно изначальных, если какое либо требование рассматривается повторно.

Признаки для оценки требований пользователей и системных требований связаны друг с другом и между собой, поэтому для построения модели принятия решения на третьем этапе также используется метод аналитических сетей (МАС).

В результате получается следующее ранжирование системных требований S1.1>S4.1>S1.2> S4.2>S3.1.

Таким образом, модели принятия решений строятся с уровня бизнес требований и далее детализируются для следующих уровней. При этом наряду

с признаками для оценки альтернатив рассматриваемого уровня, учитывается относительная значимость требований, находящихся выше в иерархии. В результате получаются приоритеты системных требований (нижнего уровня в иерархии), на основе которых далее составляются планы работ.

В данном примере рассмотрены требования, касающиеся реализации одного из функциональных блоков – «Депозит на расчетном счете». Полученные приоритеты системных требований используются при составлении плана работ по реализации в системе данного финансового инструмента, выделении ресурсов и определении сроков завершения доработок.

6. Заключение

Таким образом, процесс управления требованиями является одной из ведущих функций бизнес-аналитика при внедрении ИТ-проектов. Для достижения целей проекта стратегически важно

правильно построить процесс управления требованиями между членами проектной команды, согласовать с клиентом рамки релевантности поступающих требований и процедуру отслеживания выполненных работ.

В статье рассмотрены требования, касающиеся реализации одного из функциональных блоков – «Депозит на расчетном счете» при внедрении модуля «Казначейство» системы SAP. В результате полученные приоритеты системных требований можно использовать при составлении плана работ по реализации в системе данного финансового инструмента, выделять ресурсы и определять сроки завершения доработок.

Если таким же образом рассматривать все требования, полученные при реализации различных функциональных блоков проекта, то полученную информацию можно использовать для составления плана работ по проекту в целом. ■

Литература

1. Технология Rational Unified Process (IBM Rational Software). <http://citforum.ru/programming/application/program/3.shtml>
2. IEEE Standard Glossary of Software Engineering Terminology. <http://web.ecs.baylor.edu/faculty/grabow/Fall2011/csi3374/secure/Standards/IEEE610.12.pdf>
3. A Guide to the Business Analysis Body of Knowledge (BABOK Guide). Version 2.0. – Toronto: International Institute of Business Analysis, 2009.
4. SAP Treasury and Risk Management (TRM). http://help.sap.com/saphelp_erp60_sp/helpdata/en/08/a9e139709ba63be10000000a114084/content.htm
5. Саати Т.Л. Принятие решений: Метод анализа иерархий. – М.: Радио и Связь, 1993.
6. Саати Т.Л. Принятие решений при зависимостях и обратных связях. Аналитические сети. – М.: Изд. ЛКИ., 2008.

СОЗДАНИЕ ИНФРАСТРУКТУРЫ СМАРТ-РЕГИОНА НА ОСНОВЕ РАЗВИТИЯ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И ЭЛЕКТРОННОГО ОБРАЗОВАНИЯ

Б.Б. Славин,

кандидат физико-математических наук,
руководитель Центра ИТ экспертизы Союза ИТ-директоров России,
директор по исследованиям и инновациям компании АйТи

И.У. Ямалов,

доктор технических наук, руководитель Агентства по информационным
технологиям Республики Башкортостан

E-mail: bbslavin@gmail.com, yamalov.i@minsvyazrb.ru

Адрес: г. Москва, ул. Ленинская Слобода, д. 19, стр. 6

В статье изложены результаты исследования и моделирования роста индустрии информационных технологий (ИТ) в российском регионе за счет дополнительных инвестиций в электронное образование. Показывается, что такие инвестиции формируют инфраструктуру, характерную для СМАРТ региона, увеличивают долю высокопрофессиональных рабочих мест, создают инструменты развития для высокотехнологичного бизнеса.

Ключевые слова: электронное образование, дистанционное обучение, СМАРТ общество, региональная информатизация.

1. Введение

В последнее время по мере построения основ информационного общества большое внимание уделяется созданию инфраструктуры следующей эпохи, называемой СМАРТ обществом или обществом знаний. Такая инфраструктура подразумевает внедрение интеллектуальных технологий, снижающих рутинную составляющую во всех сферах человеческой жизни, а также ставку в развитии экономики на творческую и интеллектуальную деятельность населения. Возрастание доли знаний

в товарах и услугах позволяет тиражировать продукцию темпами, которые были невозможны даже на заре первоначального накопления капитала. Иллюстрацией к зарождению новой экономики знаний могут служить истории успеха таких крупнейших и одновременно «молодых» компании как Apple, Microsoft, Google, Facebook и др. В России к числу таких успешных проектов можно отнести Mail.Ru, Yandex, «ВКонтакте» и др.

В лидеры современного наукоемкого бизнеса сегодня вырвалась отрасль информационно-коммуникационных технологий (ИКТ). Все перечис-

ленные выше компании относятся именно к этой отрасли. Авангардная роль ИКТ неслучайна: социальные сети, поисковые системы, интернет-магазины, дистанционное обучение – все эти ресурсы тиражируют информацию и знания, и позволяют получать уникальные дивиденды за счет массового потребления. ИКТ отрасль не требует высоких инвестиционных затрат: основным ресурсом является не оборудование, а человек, его умение создавать программные (логические) продукты. Именно поэтому развитие ИКТ отрасли представляется перспективным для развития регионов, в которых имеется хорошая образовательная база (школы, вузы, институты повышения квалификации и т.д.), избыток молодого трудоспособного населения.

В статье представлены результаты моделирования ускоренного развития ИКТ отрасли в регионе за счет целевых инвестиций со стороны регионального правительства. Описанные ниже прогнозные расчеты легли в основу создания Дорожной карты развития ИТ индустрии в Республике Башкортостан до 2022 года, что говорит о практической ценности представленных результатов, о возможности использования аналогичных моделей и в других субъектах федерации России. При этом решаемая задача не сводилась только к созданию экономической модели роста отдельной отрасли, целью исследования являлось определение стратегии развития ИКТ инфраструктуры, необходимой для построения СМАРТ общества в регионе, для увеличения доли интеллектуального труда и повышения уровня образованности населения.

2. Перспективы развития ИКТ отрасли в регионах

Информационные технологии востребованы во всех отраслях экономики, поскольку являются основой автоматизации бизнес-процессов в коммерции и на производстве, в сфере государственной деятельности и социального обеспечения. При этом внедрение ИТ приводит к увеличению потребности в более квалифицированной рабочей силе. За счет роста производительности труда автоматизация ведет к сокращению ручного труда и к созданию новых рабочих мест в смежных отраслях, которые для сферы знаний требуют более интеллектуальные кадры. Профессор экономики из Университета Беркли в Калифорнии Энрико Моретти пишет: «Мои оценки, основанные на анализе данных 8 миллионов рабочих мест в 320 мегаполисах, показывают, что для каждого нового высокотехнологичного рабочего ме-

ста в городе возникает еще пять рабочих мест в области услуг. Эти пять дополнительных рабочих мест требуют различных работников: два с высшим образованием, необходимым для профессиональной деятельности, и три работника с низким уровнем образования для неквалифицированной работы» [1]. Таким образом, ИКТ приводит не столько к увеличению занятости, сколько к увеличению доли высокоинтеллектуального труда.

ИКТ отрасль представляет собой достаточно широкую область: это и производство вычислительного и коммуникационного оборудования, и создание программных продуктов для корпоративного сектора, и разработка приложений для массовых пользователей вычислительной техники, и оказание аутсорсинговых услуг поддержки и программирования, и создание контента для электронного обучения, и другие направления. Не стоит пытаться инвестировать во все эти направления сразу, эффект может оказаться достаточно низким – ИКТ имеют специфику реализации в индустриальном масштабе. Необходимо сосредоточить усилия на одном из направлений и при этом учесть возможности региона, наиболее подходящие для развития той или иной компетенции.

Например, разработка и производство сетевого и вычислительного оборудования требует сильной инженерной школы в области микроэлектроники, сетевых технологий и системного программирования. Без развития научно-исследовательской базы инвестиции в электронику обернутся созданием примитивных сборочных предприятий, добавочная стоимость знаний в которых не будет велика. Ставка на оказание аутсорсинговых услуг целесообразна в тех регионах, где стоимость рабочей силы при надлежащем качестве ниже рыночной. Например, Индия, которая изначально пошла по этому пути, имела достаточно большое число англоязычных трудовых ресурсов, стоимость которых была шокирующе низкой. Интересен опыт Ирландии, где международные ИТ компании льготами мотивируются к созданию центров локализации своих программных продуктов для Европы и всего мира.

Создание рабочих мест в ИКТ отрасли не дорого (меньше, чем в металлургии, машиностроении или на химическом производстве – в 10-15 раз), отрасль может развиваться очень быстрыми темпами (с характерным временем возврата инвестиций – от года до двух). Чтобы понять какого порядка может быть такой рост в целом для экономики, стоит обратиться к опыту развивающихся стран, сделавших ставку на развитие ИКТ. На *рис. 1* приведен рост объема ИКТ

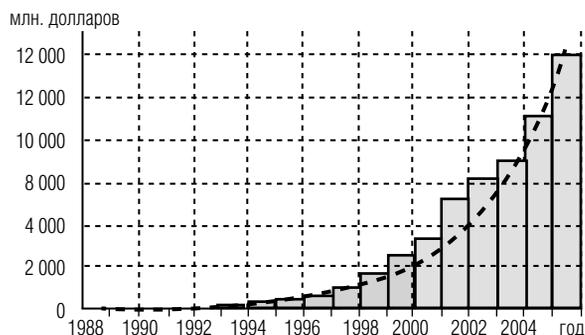


Рис. 1. Рост ИТ экспорта Индии

услуг (см. [2]) в Индии, правительство которой приняло долгосрочную национальную программу поддержки отрасли информационных технологий и разработки программного обеспечения¹. Пунктирная кривая, хорошо аппроксимирующая приведенные показатели, является степенной функцией с характерным пятикратным ростом каждые пять лет. Высокие темпы развития на протяжении почти двух десятилетий свидетельствуют о долгосрочности программы и высоком потенциале ИКТ отрасли, опирающейся на растущий международный рынок.

Отдельно надо отметить необходимость поддержки научных исследований в области ИКТ. Не случайно в расчете рейтинга конкурентоспособности ИТ индустрии по версии международной некоммерческой торговой ассоциации в области программного обеспечения BSA² вес показателя развитости исследовательской составляющей в отрасли является самым высоким – 25% (из шести индикаторов). Это же подтверждает и то, что многие западные ИТ компании (например, Google) возникли в результате научных изысканий. Кстати, такие примеры можно найти и на российском ИКТ рынке – Лаборатория Касперского, АБВУУ. Особенно продуктивным следует считать сочетание науки, образования и высокотехнологичного бизнеса. Именно таким примером является Кремниевая долина, где Стэнфордский Университет играет роль поставщика кадров для бизнеса и одновременно поддерживает большое число исследовательских лабораторий и центров, идеи которых подхватываются бизнесом.

3. Электронное образование как драйвер развития СМАРТ-региона

Особым направлением развития ИКТ отрасли является электронное образование. Уникальность это-

го направления состоит в том, что оно не только требует профессиональные кадры, как и любое другое направление в ИКТ, но и само формирует такие кадры за счет увеличения эффективности образования с использованием средств интерактивного и дистанционного обучения. Аналогичный путь развития ИКТ индустрии в мире был выбран Южной Кореей [3]: инвестировав в электронное образование, южнокорейское правительство поддержало ИТ компании и отрасль электроники (компьютерное обеспечение, интерактивные доски и т.п.), и одновременно сформировало спрос на ИКТ в стране. Двойной эффект от поддержки электронного образования делает это направление наиболее интересным при формировании стратегии развития ИКТ в отдельном регионе.

Особенно интересной «ставка» на электронное обучение в регионе становится в свете перехода к информационной и, далее, к СМАРТ («умной») стадии развития общества, опирающейся на знаниевый потенциал населения. В современную эпоху базовое образование уже неспособно обеспечить поддержание необходимого для экономики уровня компетенции людей. Дополнительное обучение в различных его формах (курсы повышения квалификации, бизнес-образование, обучающие семинары и мастер-классы, и т.д.) приближается по объему к среднему и высшему образованию. Процесс обучения становится непрерывным, сопровождающим человека на протяжении всей его жизни. Колоссальный объем и широкий спектр необходимых для освоения современных знаний невозможно передавать только в рамках традиционных форм обучения, необходимо использовать разнообразные электронные средства: дистанционные и интерактивные. Рынок электронного обучения будет только расти, и инвестиции в него окупятся многократно в будущем.

Электронное обучение напрямую связано с информационно-коммуникационными технологиями: создание интерактивного контента требует программных разработок, а дистанционное обучение – развитие каналов связи, систем производства и хранения видео и мультимедиа информации. В отличие от обычного образования, стоимость которого в основном определяется зарплатой профессорско-преподавательского состава и затратами на поддержание инфраструктуры учебных зданий, электронное обучение – это в первую очередь инвестиции в труд инженеров-программистов. Дистанционные технологии снижают потребность в преподавателях за счет тиражирования выступлений лучших квалифицированных педагогов и внедрения интерактивных методов, делающих про-

¹ <http://www.it-taskforce.nic.in/infplan.htm>

² См. исследование Business Software Alliance конкурентоспособности стран в области ИТ за 2011: <http://globalindex11.bsa.org/>

цесс обучения самостоятельным и порой более увлекательным, чем с учителем. Безусловно, электронное обучение не может полностью заменить обычное преподавание, особенно в школе и вузах, где необходимо личное общение с педагогом. Но в дополнительном образовании современные цифровые технологии вполне могут стать основной формой обучения.

В развитии СМАРТ общества (в России в рамках субъекта федерации такое развитие называют развитием СМАРТ-региона) образование начинает играть более существенную роль, чем это было ранее, становясь средой, сопровождающей человека на протяжении всей его жизни. Профессор Тихомиров считает, что «происходит смена образовательной парадигмы с традиционной модели обучения к электронному обучению (e-learning), и далее к Smart education» [3, с.7]. Однако СМАРТ образование не отменяет традиционных моделей обучения, оно дополняет их новыми моделями, причем не только за счет современных технологий (которые, безусловно, используются гораздо шире), но и за счет новых форм организации процесса обучения. Достаточно упомянуть уже получивший международную известность проект создания электронной образовательной платформы Coursera [4], который инициировали профессор Стэнфордского Университета Дафни Коллер и Андре Нг. Десятки крупнейших университетов мира бесплатно размещают на этой платформе свои электронные лекции, к которым получают доступ сотни тысяч студентов со всех уголков Земли. Массовые обсуждения и продвижение электронных ресурсов в социальной среде превращает процесс образования одновременно и в инструмент отбора лучших образовательных программ, и в инструмент поиска новых идей, и в коллективное самообучение.

4. Принципы и каналы развития индустрии ИТ и электронного обучения

Стимулирование развития той или иной отрасли со стороны власти должно строиться на принципе не прямых инвестиций в отрасль. Непосредственно созданием рабочих мест и разработкой услуг, расширением числа выпускников и программ обучения ИТ должен заниматься бизнес и учреждения образования. А власть должна создавать для этого условия. Одной из форм создания таких условий является частно-государственное партнерство, которое предполагает совместное финансирование проектов со стороны государства и бизнеса. Но совместное — не означает просто доленое. Необходимо, чтобы инвестиции государства создавали условия, которые бы

способствовали росту бизнеса. И паритет в таком случае означает, что финансирование со стороны государства в развитие бизнес среды привлекает примерно столько же инвестиций со стороны бизнеса на организацию рабочих мест. Именно такой принцип был заложен в модель развития индустрии электронного образования в Республике Башкортостан (РБ).

Инвестиции регионального Правительства в бизнес среду целесообразно распределять по нескольким каналам, необходимым для мотивации бизнеса и системы образования к росту. Такими каналами могут быть:

- ◆ финансирование программы электронного образования населения и служащих, реализация которой подтолкнет бизнес и учреждения образования к расширению услуг;
- ◆ софинансирование фондов поддержки ИТ-индустрии, которые призваны привлекать инвестиции из бизнеса в ИТ отрасль;
- ◆ финансирование инфраструктурных и медиа проектов, которые дают преимущества ИТ компаниям, чьи разработчики проживают в регионе.

Без реализации первой задачи (финансирования долгосрочной программы электронного обучения) не удастся в достаточной мере увеличить рост ИТ отрасли. Должна быть разработана долгосрочная программа электронного обучения (например, рассчитанная на 10 лет), которая станет важным организационным условием перехода к информационному, и далее к СМАРТ обществу в регионе. В долгосрочной программе необходимо учесть предоставление образовательных услуг всех уровней — от краткосрочного повышения квалификации до высшего профессионального образования. При этом должны использоваться все современные электронные образовательные ресурсы (тренажеры, обучающие программы, электронные учебные пособия, информационные киоски и др.), все имеющиеся средства медиа-обучения (видеоматериалы, вебинары, онлайн-конференции, круглые столы и др.).

Без реализации второй задачи (создания частно-государственных инвестиционных фондов в области ИТ) высоки риски, что по истечении 10-ти лет развитие ИТ-отрасли пойдет на убыль. Основное условие деятельности инвестиционных фондов поддержки ИТ-индустрии — возвратность инвестиций. Выполнение этого условия обеспечит участие в инвестициях на паритетных началах бизнеса и стабильный рост накоплений в фондах. В свою очередь накопленные активы станут гарантией функционирования созданных инвестиционных институтов и

после окончания программы стимулирования развития ИТ-отрасли. Реализация третьей задачи (поддержка со стороны регионального Правительства инфраструктурных проектов, медийной и научной деятельности в области ИТ) необходима для того, чтобы инвестиции остались в регионе.

5. Модель развития индустрии электронного обучения

Экономическая модель развития электронного обучения строится на целевых показателях роста отрасли, которые трудно предсказать — их надо задавать. Проще всего пойти по пути сравнения с имеющимися практиками. Например, за «сильные» показатели можно взять показатели роста ИТ индустрии в Индии (см. рис. 1), и выстроить постепенный переход от текущих показателей к «сильным» за определенный срок. На рис. 2 показан такой переход от умеренного текущего роста ИТ отрасли (на примере Республики Башкортостан) к высокому росту, аналогичному Индии. Переходный период рассчитан на 10 лет — срок, за который можно существенно изменить характер развития отрасли. Введение переходного периода позволяет плавно наращивать ресурсы для поддержки роста ИТ индустрии. В частности, при темпах, изображенных на рис. 2 рост отрасли за 10 лет составит 11,5 раз, в отличие от Индии, где рост за 10 лет составил 25 раз.

На рис. 3 также для 10-летнего периода показаны стратегические показатели (пунктиром) для республики Башкортостан, которые выходят на «сильный» темп (131,4% роста в год — обозначен сплошной линией) к 10-му году. Такие показатели закладывались в экономическую модель развития ИТ отрасли РБ. Даже при такой «слабой» динамике (увеличения объема ИТ индустрии с 0,9 млрд. руб. в год до 10,3 млрд. руб.) инвестиции в организацию рабочих мест³ с учетом инфляции составят 6,31 млрд. руб. Эта величина является рассчитываемой из целевых показателей. Поскольку одним из принципов поддержки

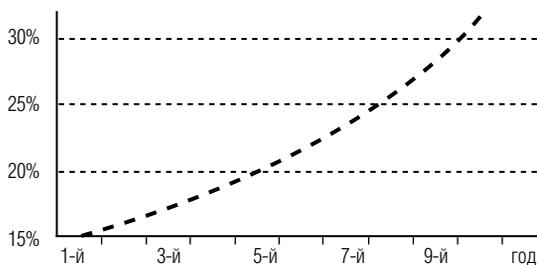


Рис. 2. Динамика роста ИТ отрасли РБ за 10 лет

³ Согласно оценке экспертов, инвестиции на одно рабочее место в ИТ области составляли в 2012 г. 1 млн. руб.

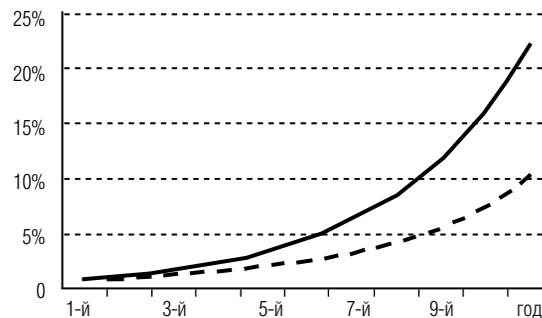


Рис. 3. Рост объема ИТ отрасли РБ в млрд. руб.

ИТ-отрасли (см. предыдущий раздел) является паритетное финансирование со стороны бизнеса и государства, общие расходы региона за 10 лет также должны составить 6,31 млрд. руб. Т.е. общие инвестиции в ИТ отрасль в этой модели за 10 лет превысят 12,6 млрд. долл. При этом прямых инвестиций со стороны государства в создание новых рабочих мест не предполагается, рост занятости должен стать результатом стимулирования и регулирования рынка: бизнес сам создаст необходимое количество рабочих мест.

Принцип соинвестирования должен сделать рост отрасли необратимым, поскольку обеспечит ее развитие частным интересом. Общие расходы со стороны регионального правительства целесообразно разделить на три составляющие, соответственно каналам развития ИТ индустрии (см. выше). Безусловно, разделение на каналы не может быть арифметически ровным, и должно учитывать специфику региона или даже определенного времени. Например, в случае Республики Башкортостан было учтено, что на третьем году реализации программы в регионе будет проведен международный саммит, на который отводилось существенное финансирование. Понятно, что к саммиту необходимо было реализовать ряд первостепенных программ (создание необходимой инфраструктуры отрасли, обучение населения английскому языку и т.п.), и именно на это время имеет смысл планировать максимальным финансирование.

Все три составляющие взаимосвязаны между собой, и их динамика самосогласована. Например, выделение со стороны государства средств на создание инфраструктуры в регионе (коворкинг-центр, студии разработки электронного контента и т.д.), позволит профинансировать строительство или реконструкцию зданий ИТ-парков из частно-государственных фондов поддержки ИТ-индустрии (в этом случае такие инвестиции будут привлекательными для бизнеса). Долгосрочная программа электронного обучения в регионе также является залогом успешности инве-

стиций со стороны Фонда в ИТ-проекты, связанные с созданием контента для такого обучения. Проведение конференций, участие в различных российских и международных выставках в свою очередь будет способствовать росту популярности ИТ-компаний региона, обеспечивая их конкурентоспособность в тендерах на электронное обучение.

Расходы на электронное (дистанционное и интерактивное) обучение граждан и государственных служащих — это основа программы развития ИТ-индустрии, желательно наращивать в течение всего периода: например, с менее чем 25% в первый год реализации до 2/3 в последний год. На рис. 4 показано, что первая составляющая расходов бюджета (электронное образование — закрашено равномерными точками), достигнет за 10 лет 2,7 млрд. руб. (более 40% от всех расходов). Такая динамика связана с необходимостью постепенного развития мощности региональных ИТ компаний, создающих контент для электронного образования. Чрезмерное ускорение может привести к нехватке кадров и «уходу» инвестиций за пределы региона.

Вторая составляющая связана с фондами поддержки ИТ-индустрии (закрашено черным на рис. 4). Фонды не являются посевными (посевное финансирование может быть реализовано в третьем канале), носят частно-государственный характер и должны обеспечивать в целом возвратность инвестиций. Поэтому даже при нулевой прибыли они должны расти и для приведенного примера целевых показателей за десять лет смогут аккумулировать более 1,5 млрд. руб., а с учетом частных инвестиций более 3 млрд. руб. При доходности фондов в 20%⁴ их совокупный размер к концу периода может превысить 9 млрд. руб., что сделает такие фонды мощнейшим инструментом для развития ИТ-индустрии и после реализации программы поддержки. Динамика дополнительных инвестиций со стороны государства в фонды должна быть ниспадающей по мере их формирования, и к последнему году периода, охватываемого программой, может составлять символические 50 млн. руб. в год. Третья составляющая связана с поддержкой ИТ отрасли через финансирование инфраструктуры и медиа поддержку (ИТ-парки, создание системы инфраструктуры,

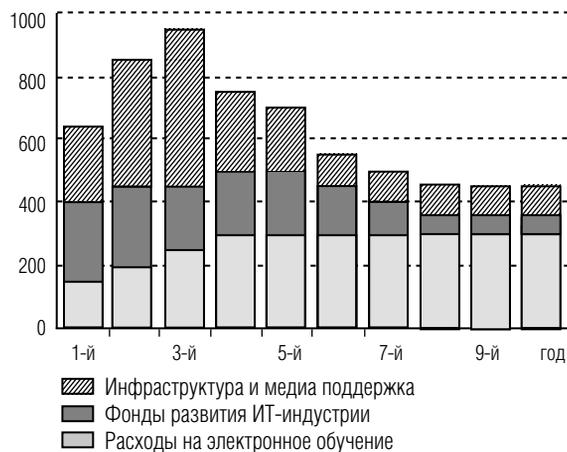


Рис. 4. Ежегодные расходы по каналам, млн. руб.

продвижение ИТ отрасли в других российских регионах и за рубежом, конференции и т.д.). Эти расходы необходимы для придания заданных программой темпов развития отрасли.

6. Заключение

В заключение заметим, что формирование СМАРТ-региона должно строиться, прежде всего, на росте объема интеллектуального труда в экономике. Именно увеличение доли творческого труда в производстве товаров и услуг позволяет получить конкурентное преимущество благодаря тиражируемости знаний. Модернизация производства создаст условия для перехода к экономике знаний (за счет снижения рутинного труда), но не формирует среду развития СМАРТ-региона, которая должна поддерживать научные и творческие виды деятельности. Как раз ставка на электронное образование является одним из возможных сценариев создания экономики знаний. Электронное образование обладает двойным эффектом «смартвости»: стимулирует развитие ИТ отрасли и повышает общий интеллектуальный уровень населения. Поддержка формирования полномасштабной системы электронного образования региона должна включать в себя инвестиции в образование, поддержку фондов развития и создание региональной инфраструктуры, необходимой для работы в бизнесе местных кадров.■

⁴ По данным РАВИ (Российской Ассоциации Венчурного Инвестирования), средняя внутренняя норма доходности инвестиционных проектов венчурных фондов, действующих в России составляет 35%, что совпадает с заявляемым самими фондами типичным минимумом доходности интересующих их проектов (30–40%), а среднеевропейский показатель IRR (1) венчурных фондов составляет 12–14%. 20% - средняя оценка с учетом того, что фонды поддержки ИТ индустрии в будущем выполнять не только коммерческие, но и государственные задачи.

⁵ <http://www.reuters.com/article/2012/04/18/net-us-usa-college-online-idUSBRE83H0PC20120418>

Литература

1. Moretti E. The high-tech boom and the rest of us. — SFgate.com, November 18, 2012.
2. Technology, Adaptation, and Exports. How some developing countries got it right / Edited by V.Chandra. — Washington, DC: The World Bank, 2006.
3. Россия на пути к Smart обществу / Под ред. Н.В.Тихомировой, В.П.Тихомирова. — М.: НП «Центр развития современных образовательных технологий», 2012.
4. Top U.S. colleges to offer free classes online. Reuters5 // Stephanie Simon. — Apr 18, 2012.

**ЖУРНАЛ «БИЗНЕС-ИНФОРМАТИКА»
ОСУЩЕСТВЛЯЕТ РАЗМЕЩЕНИЕ РЕКЛАМНЫХ
И РЕКЛАМНО-ИНФОРМАЦИОННЫХ МАТЕРИАЛОВ**

Расценки:

Обложка:

2, 3, 4 страница обложки, полноцветная печать,
полоса 210 X 290 мм (A4) — 40 тыс. руб.

Текстовый блок, черно-белая печать:

- ♦ полоса — 20 тыс. руб.;
- ♦ 1/2 полосы — 15 тыс. руб.;
- ♦ 1/4 полосы — 10 тыс. руб.;
- ♦ меньший объем — 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 учетом следующих параметров:

- ♦ дообрезной формат — 215 X 300 мм;
- ♦ обрезной формат — 210 X 290 мм;
- ♦ поле набора полосной рекламы — 190 X 270 мм — с отступом от границ обрезного формата по 10 мм с каждой стороны;
- ♦ файл TIF, EPS, PDF— разрешение не менее 300 dpi.

METHODS OF PATTERN ANALYSIS IN STATICS AND DYNAMICS, PART 1: LITERATURE REVIEW AND CLARIFICATION OF THE TERM

F. Aleskerov, V. Belousova, L. Egorova, B. Mirkin

Annotation

Pattern analysis is a new area in data analysis related with the search for relationships between the objects, the construction of the objects classification and study of object's changes over time. In the first part of the article the notion of 'pattern' is introduced, and a survey of methods of cluster analysis and pattern analysis is presented.

Key words: data patterns, dynamic pattern analysis, cluster analysis.

CROSSLEXICA: A UNIVERSE OF LINKS BETWEEN RUSSIAN WORDS

I. Bolshakov

Annotation

A superlarge Russian computer dictionary is created with a vocabulary of 290,000 elements and as many as 8 million links of all possible types between them. It is highly multidisciplinary and is designed for any audience applying to it with requests in Russian or English. In interactive mode, it facilitates text editing and language learning via linguistic and encyclopedic information. In non-interactive mode, it is accessible for any external software.

Key words: Russian language, computer dictionary, links between words, multidisciplinary, interactive and programmed access, requests in Russian or English.

APPLICATION OF ERGONOMIC SEMIOTICS PRINCIPLES FOR USER INTERFACE DESIGN IN THE POLY CULTURAL CONTEXT

J. Taratukhina, D. Aldunin

Annotation

Nowadays there is an active implementation of information technology in the learning process. This is expressed in e-textbooks and educational environments as well as web sites and Internet communities. Very often such e-resources imply multicultural audience. There are different approaches to creation of ergonomic user interface design in different cultures, though. This particular paper is devoted to analysis of approaches' differentiation and is also aimed at working out a list of recommendations on the ways to improve ergonomic user interface design of electronic learning materials oriented for multicultural audience.

Key words: ergonomic semiotics, ergonomic design, elec-tronic learning materials, multicultural audience.

DISTRIBUTED EVOLUTIONARY NETWORK FOR THE SOLUTION OF MULTI-OBJECTIVE OPTIMIZING PROBLEMS IN SIMULATION SYSTEMS

M. Hivintsev, N. Akopov

Annotation

In the article is presented a novel approach to the solution of multi-objective optimizing problems of large-scale dimension systems realized, in particular, in the simulation systems of the class AnyLogic through distributed calculations. A new concept of creation of the distributed evolutionary network is suggested, based on splitting of space of required variables into clusters and assignments is offered to each computing element of a network of the cluster on which search of intermediate results by means of interacting genetic algorithms is carried out.

Key words: simulation modeling, genetic algorithms, distributed calculations, multi-objective optimization.

WEIGHTED SUM METHOD IN THE ANALYSIS OF MULTICRITERIAL DECISIONS: PRO ET CONTRA

V. Podinovski, M. Potapov

Annotation

The paper presents results of the analyzing the popular weighted sum method (WSM).

Key words: multiple criteria decision making problems, weighted sum method, scales of criteria, weights of criteria, criteria importance, normalization of criteria.

METHOD OF EMPIRICAL PROBABILITIES: AUTOMATIC SYSTEM TO RECOMMEND THE FOLLOWING TEN LECTIONS AFTER VIEWING THREE GIVEN LECTIONS

V. Nikulin, S. Palesheva, D. Zubareva

Annotation

In this paper we present an algorithm and the corresponding experimental results, which were obtained online during the VideoLectures.Net ECML/PKDD 2011 Discovery Challenge (Track N2), where we were awarded a prize for the third best result. We propose to use two lectures (out of the given three lectures) in order to define a direction of the prediction. The relevance of the whole predicted set is calculated according to the remaining third lecture.

Key words: recommender system, collaborative filtering, online learning, ensembling, bagging, resampling.

AUTOMATION OF THE TECHNOLOGY ROADMAP DEVELOPMENT. CALCULATION OF THE INTEGRAL INDICATORS OF APPLICABILITY

O. Ena, K. Nagaev

Annotation

The methods of calculating of particular and integral indicators of applicability for the technological roadmap elements and their attributes are considered. These indicators are applied to the automated tools for the foresight researches, the long-term scientific and technological forecasting, identifying innovative trajectories of the emerging domains evolution.

Key words: technology roadmap, foresight, indicator of applicability, technology trajectory, knowledge formalization, expert survey, MAUT.

REQUIREMENTS MANAGEMENT IN IT PROJECTS

T. Kravchenko

Annotation

In the paper an approach to modeling of requirements management process associated with IT projects is considered. The requirements management model includes three stages. The first stage is related with the choice of business requirements, which are described in the project solution and define the project scope. The second stage includes development of a model that is associated with accepting, rejecting, clarification or classification as 'additional task' for every of incoming user requirements. On the third stage for all the accepted user requirements priorities of system requirements are formulated; these priorities are subsequently used for project planning. The decision making models are based on the methods of analytic hierarchy process (AHP) and analytic network process (ANP), as well as on SuperDecisions decision support system.

Key words: requirements management, modeling of requirements management process, analytic hierarchy process (AHP), analytic network process (ANP), SuperDecisions decision support system.

CREATING SMART-REGION'S INFRASTRUCTURE THROUGH THE DEVELOPMENT OF INFORMATION TECHNOLOGY AND E-LEARNING

B. Slavin, I. Yamalov

Annotation

The article presents the results of research and modeling of IT-industry growth in the Russian region through additional investments in e-learning. It is shown that such investments form the infrastructure typical for SMART region; increase the proportion of highly skilled jobs; create tools for the development of high-tech businesses.

Key words: electronic education, distance learning, SMART society, regional informatization.