

БИЗНЕС- ИНФОРМАТИКА

МЕЖДИСЦИПЛИНАРНЫЙ НАУЧНО-ПРАКТИЧЕСКИЙ ЖУРНАЛ

СОДЕРЖАНИЕ

Анализ данных и интеллектуальные системы

В.А. Агиевич, К.Г. Скрипкин
Формализация задачи выбора оптимальной последовательности изменений архитектуры предприятия на основе матрицы изменений Бринйолфссона 7

С.М. Авдошин, А.А. Лифшиц
Формирование портфеля проектов на основе нечеткой модели многокритериальной оптимизации 14

М.А. Хивинцев, А.С. Акопов
Применение многоагентного генетического алгоритма для поиска оптимальных стратегических и оперативных решений 23

Информационные системы и технологии в бизнесе

А.Ф. Моргунов
Внедрение распределенной информационной системы в многофилиальной организации: опыт ФГУП «Почта России» 34

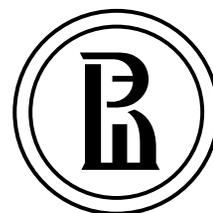
Д.В. Исаев
К вопросу об оценке уровня развития систем управления эффективностью 42

К.В. Нагаев, Е.М. Курбатова
Автоматизация проектирования технологических дорожных карт. Сбор информации и консолидация экспертных мнений 52

Математические методы и алгоритмы бизнес-информатики

С.М. Ямпольский, А.С. Шаламов
Моделирование процессов торгового предприятия 61

А.Г. Дьяконов
Прогноз поведения клиентов супермаркетов с помощью весовых схем оценок вероятностей и плотностей 68



Издатель:

Национальный
исследовательский университет
«Высшая школа экономики»

**Подписной индекс
в каталоге агентства
«Роспечать» –72315**

Выпускается ежеквартально

*Журнал включен в Перечень
российских рецензируемых
научных журналов,
в которых должны быть
опубликованы основные научные
результаты диссертаций
на соискание ученых степеней
доктора и кандидата наук*

*Главный редактор
А.О. Голосов*

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

*Научный редактор
Н.Н. Лычкина*

*Технический редактор
В.И. Осипов*

*Дизайн обложки
С.Н. Борисова*

*Компьютерная верстка
О.А. Богданович*

*Администратор веб-сайта
Д.С. Проценко*

Журнал рекомендован ВАК
для научных публикаций

Адрес редакции:
105187, г. Москва, ул. Кирпичная, д. 33
Тел./факс: +7 (495) 771-32-38
<http://bijournal.hse.ru>
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»

О ЖУРНАЛЕ

«**Б**изнес-информатика» – рецензируемый междисциплинарный научный журнал, выпускаемый с 2007 года Национальным исследовательским университетом «Высшая школа экономики» (НИУ ВШЭ). Администрирование журнала осуществляется факультетом бизнес-информатики НИУ ВШЭ.

Миссия журнала – развитие бизнес-информатики как новой области информационных технологий и менеджмента. Журнал осуществляет распространение последних разработок технологического и методологического характера, способствует развитию соответствующих компетенций, а также обеспечивает возможности для дискуссий в области применения современных информационно-технологических решений в бизнесе, менеджменте и экономике.

Журнал публикует статьи по следующей тематике:

- ◆ корпоративные информационные системы;
- ◆ информационные технологии в бизнесе;
- ◆ организационные и управленческие проблемы создания и внедрения информационных систем;
- ◆ математическое моделирование социально-экономических процессов;
- ◆ методы анализа информации;
- ◆ интеллектуальные системы и управление знаниями в бизнесе;
- ◆ информационные сети и телекоммуникации;
- ◆ программная инженерия;
- ◆ информационная безопасность;
- ◆ электронный бизнес;
- ◆ инновации и бизнес в сфере информационных технологий;
- ◆ стандартизация, сертификация и качество;
- ◆ правовые вопросы бизнес-информатики;
- ◆ обучение в сфере бизнес-информатики.

В соответствии с решением президиума Высшей аттестационной комиссии Российской Федерации с 2010 года журнал включен в Перечень российских рецензируемых научных журналов, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней доктора и кандидата наук.

Журнал выпускается ежеквартально и распространяется как в печатном виде, так и в электронной форме.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

ГЛАВНЫЙ РЕДАКТОР

ГОЛОСОВ Алексей Олегович —
кандидат технических наук, Президент компании
«ФОРС — Центр разработки»

ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА

ГОРБУНОВ Алексей Рэмович —
кандидат экономических наук, старший научный сотрудник,
Института США и Канады РАН

ИСАЕВ Дмитрий Валентинович —
кандидат экономических наук, доцент кафедры бизнес-
анализа, факультет бизнес-информатики, Национальный
исследовательский университет «Высшая школа экономики»

ЧЛЕНЫ РЕДКОЛЛЕГИИ

АБДУЛЬРАБ Абиб —
PhD, профессор департамента математики и программной
инженерии, Национальный институт прикладных наук, Руан,
Франция

АВДОШИН Сергей Михайлович —
кандидат технических наук, заведующий кафедрой управления
разработкой программного обеспечения, заведующий
отделением программной инженерии, Национальный
исследовательский университет «Высшая школа экономики»

АЛЕСКЕРОВ Фуад Тагиевич —
доктор технических наук, профессор, руководитель
департамента математики, факультет экономики, Национальный
исследовательский университет «Высшая школа экономики»

БАБКИН Эдуард Александрович —
кандидат технических наук, PhD, заведующий кафедрой
информационных систем и технологий, факультет бизнес-
информатики и прикладной математики (Нижний Новгород),
Национальный исследовательский университет
«Высшая школа экономики»

БАЙЕР Алекс —
PhD, Директор KAFAN FX Information Services, Нью-Йорк, США

БАРАНОВ Александр Павлович —
доктор технических наук, профессор, заведующий кафедрой
информационной безопасности, факультет бизнес-информа-
тики, Национальный исследовательский университет
«Высшая школа экономики»

БЕККЕР Йорг —
PhD, проректор, профессор, директор Европейского
исследовательского центра в области информационных систем
(ERCIS) Мюнстерского университета, Мюнстер, Германия

БЕЛОВ Владимир Викторович —
доктор технических наук, профессор кафедры вычислительной
и прикладной математики, факультет вычислительной техники,
Рязанский государственный радиотехнический университет.

ГРИБОВ Андрей Юрьевич —
кандидат экономических наук, Генеральный директор
компании «КиберПлат»

ГРОМОВ Александр Игоревич —
кандидат химических наук, профессор, заведующий кафедрой
моделирования и оптимизации бизнес-процессов, факультет
бизнес-информатики, Национальный исследовательский
университет «Высшая школа экономики»

ГУРВИЧ Владимир Александрович —
PhD, приглашенный профессор и исследователь,
Центр исследования операций, Ратгерский университет
(Университет Нью-Джерси), США

ДЖЕЙКОБС Лоренц —
PhD, профессор медицинского факультета, Университет
Цюриха, Швейцария

ЗАНДКУЛЬ Курт —
PhD, заведующий кафедрой информационных систем
для бизнеса, институт информатики, факультет информатики
и электротехники, Университет Ростока, Германия

ИЛЬИН Николай Иванович —
доктор технических наук, профессор, заместитель начальника
Управления специальной связи, Федеральная служба охраны
Российской Федерации (ФСО России)

КАЛЯГИН Валерий Александрович —
доктор физико-математических наук, профессор,
заведующий кафедрой прикладной математики и информатики,
факультет бизнес-информатики и прикладной математики
(Нижний Новгород), Национальный исследовательский
университет «Высшая школа экономики»

КАМЕННОВА Мария Сергеевна —
кандидат технических наук, директор компании «ИДС Шер»,
Россия и страны СНГ

КУЗНЕЦОВ Сергей Олегович —
доктор физико-математических наук, профессор, заведующий
кафедрой анализа данных и искусственного интеллекта,
заведующий отделением прикладной математики, Национальный
исследовательский университет «Высшая школа экономики»

МАЛЬЦЕВА Светлана Валентиновна —
доктор технических наук, профессор, заведующий кафедрой
инноваций и бизнеса в сфере информационных технологий,
декан факультета бизнес-информатики, Национальный
исследовательский университет «Высшая школа экономики»

МЕЙОР Питер —
PhD, заместитель директора консультативной группы
по радиокommunikациям, Международный телекоммуникаци-
онный союз (ITU), заместитель руководителя Комиссии ООН
по науке и технологиям, Женева, Швейцария

МИРКИН Борис Григорьевич —
доктор технических наук, профессор кафедры анализа данных
и искусственного интеллекта, отделение прикладной мате-
матики и информатики, Национальный исследовательский
университет «Высшая школа экономики»

МОТТЛЬ Вадим Вячеславович —
доктор технических наук, профессор кафедры
информационной безопасности, факультет кибернетики,
Тульский государственный университет

ПАЛЬЧУНОВ Дмитрий Евгеньевич —
доктор физико-математических наук, заведующий кафедрой
общей информатики, факультет информационных технологий,
Новосибирский государственный университет

ПАРДАЛОС Панайот (Панос) —
PhD, почетный профессор, директор центра прикладной
оптимизации, департамент промышленной и системной
инженерии, Университет Флориды, США

СИЛАНТЬЕВ Альберт Юрьевич —
доктор технических наук, профессор кафедры
информационных бизнес систем, Институт информационных
бизнес-систем, Национальный исследовательский
технологический университет «МИСиС»

ТАРАТУХИН Виктор Владимирович —
кандидат технических наук, PhD, руководитель научной группы
Европейского исследовательского центра в области
информационных систем (ERCIS) Мюнстерского
университета, Мюнстер, Германия

УЛЬЯНОВ Михаил Васильевич —
доктор технических наук, профессор кафедры управления
разработкой программного обеспечения, отделение
программной инженерии, Национальный исследовательский
университет «Высшая школа экономики»

ШАЛКОВСКИЙ Алексей Геннадьевич —
кандидат технических наук, проректор, директор Института
информационных технологий, Национальный
исследовательский университет «Высшая школа экономики»

ISSN 1998-0663

BUSINESS INFORMATICS

INTERDISCIPLINARY ACADEMIC JOURNAL

CONTENTS

Data analysis and intelligence systems

V. Agievich, K. Skripkin

Formalization of the problem of choosing
the optimal sequence of Enterprise Architecture
changes on basis of Brynjolfsson's Matrix of Change 7

S. Avdoshin, A. Lifshits

Project portfolio formation
based on fuzzy multi-objective model..... 14

M. Khivintcev, A. Akopov

Application of multi-agent genetic algorithm for search
of optimum strategic and operational decisions 23

Information systems and technologies in business

A. Morgunov

Implementation of a distributed information system
in a multi-branch organization:
FSUE «Russian Post» case study 34

D. Isaev

On evaluation of performance
management systems maturity 42

K. Nagaev, E. Kurbatova

Technology roadmapping.
Method for gathering and consolidation
expert opinions 52

Mathematical methods and algorithms of business informatics

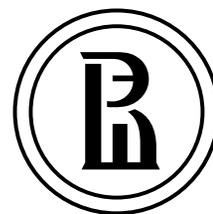
S. Yampolsky, A. Shalamov

Modeling of trade enterprise processes 61

A. D'yakonov

Supermarkets clients behaviour forecasting
by weighted methods of probability
and density estimations 68

№1(27)-2014



Publisher:

National Research University –
Higher School of Economics

**Subscription index
in the «Rospechat» catalogue –
72315**

The journal is published quarterly

*The journal is included
into the list of peer reviewed
scientific editions established
by the Supreme Certification
Commission of the Ministry
of Education and Science
of the Russian Federation*

Editor-in-Chief:

A. Golosov

Deputies Editor-in-Chief:

A. Gorbunov, D. Isaev

Scientific Editor:

N. Lychkina

Technical Editor:

V. Osipov

Design:

S. Borisova

Computer Making-up:

O. Bogdanovich

Website Administration:

D. Protsenko

Address:

33, Kirpichnaya str., Moscow,
105187, Russian Federation

Tel./fax: +7 (495) 771-32-38

<http://bijournal.hse.ru>

E-mail: bijournal@hse.ru

Circulation – 500 copies

Printed in HSE Printing House
3, Kochnovsky proezd, Moscow,
Russian Federation

© National Research University –
Higher School of Economics

ABOUT THE JOURNAL

Business Informatics is a peer reviewed interdisciplinary academic journal published since 2007 by National Research University – Higher School of Economics (HSE), Moscow, Russian Federation. The journal is administered by Faculty of Business Informatics.

The mission of the journal is to develop business informatics as a new field within both information technologies and management. It provides dissemination of latest technical and methodological developments, promotes new competences and provides a framework for discussion in the field of application of modern IT solutions in business, management and economics.

The journal publishes papers in the areas of, but not limited to:

- ◆ Corporate information systems
- ◆ Information technologies in business
- ◆ Organizational and managerial problems of information systems development and implementation
- ◆ Mathematical modeling of economic and social processes
- ◆ Methods of information analysis
- ◆ Intellectual systems and knowledge management in business
- ◆ Information networks and telecommunications
- ◆ Software engineering
- ◆ Information security
- ◆ Electronic business
- ◆ Innovations and business in the sphere of information technologies
- ◆ Standardization, certification and quality
- ◆ Legislation in the field of business informatics
- ◆ Education in the field of business informatics.

Since 2010 the journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Ministry of Education and Science of the Russian Federation.

The journal is published quarterly and distributed both in printed and electronic forms.

EDITORIAL BOARD

EDITOR-IN-CHIEF

Dr. Alexey GOLOSOV –

President of FORS Development Center, Russian Federation

DEPUTY EDITOR-IN-CHIEF

Dr. Alexey GORBUNOV –

Senior Researcher, Institute of US and Canada Studies

Dr. Dmitry ISAEV –

Associate Professor, Department of Business Analytics, Faculty of Business Informatics, National Research University Higher School of Economics, Russian Federation

EDITORIAL BOARD

Prof. Habib ABDULRAB –

Professor, Mathematical and Software Engineering Department, National Institute of Applied Sciences – Institut national des sciences appliquées de Rouen (INSA de Rouen), Rouen, France

Dr. Sergey AVDOSHIN –

Head of Software Management Department, Head of School of Software Engineering, National Research University Higher School of Economics, Russian Federation

Prof. Fuad ALESKEROV –

Head of Department of Mathematics, Faculty of Economics, National Research University Higher School of Economics, Russian Federation

Dr. Eduard BABKIN –

Head of Department of Information Systems and Technologies, Faculty of Business Informatics and Applied Mathematics (Nizhny Novgorod), National Research University Higher School of Economics, Russian Federation

Dr. Alex BAYER –

Head of KAFAN FX Information Services, New York, USA

Prof. Alexander BARANOV –

Head of Department of Information Security Management, Faculty of Business Informatics, National Research University Higher School of Economics, Russian Federation

Prof. Jorg BECKER –

Vice Rector, Professor, Director of European Research Center for Information Systems (ERCIS), University of Munster, Germany

Prof. Vladimir BELOV –

Professor, Department of Computational and Applied Mathematics, Faculty of Computer Engineering, Ryazan State Radio Engineering University, Russian Federation

Dr. Andrey GRIBOV –

Director General, CyberPlat Company, Russian Federation

Prof. Alexander GROMOV –

Head of Department of Modeling and Business Process Optimization, Faculty of Business Informatics, National Research University Higher School of Economics, Russian Federation

Prof. Vladimir GURVICH –

Invited Professor and Researcher, Rutgers Center for Operations Research, Rutgers, The State University of New Jersey, USA

Prof. Laurence JACOBS –

Professor, Medical School, University of Zurich, Switzerland

Prof. Kurt SANDKUHL –

Head of Department of Business Information Systems, Institute of Computer Science, Faculty of Computer Science and Electrical Engineering, University of Rostock, Germany

Dr. Nikolay ILYIN –

Deputy Head, Administration of Special Communication, Federal Security Guard, Russian Federation

Prof. Valery KALYAGIN –

Head of Department of Applied Mathematics and Informatics, Faculty of Business Informatics and Applied Mathematics (Nizhny Novgorod), National Research University Higher School of Economics, Russian Federation

Dr. Maria KAMENNOVA –

director, IDS Scheer Russia and CIS Countries, Russian Federation

Prof. Sergey KUZNETSOV –

Head of Department of Data Analysis and Artificial Intelligence, Head of School of Applied Mathematics and Information Science, National Research University Higher School of Economics, Russian Federation

Prof. Svetlana MALTSEVA –

Head of Department of Innovation and Business in Information Technologies, Acting Dean of Faculty of Business Informatics, National Research University Higher School of Economics, Russian Federation

Dr. Peter MAJOR –

Vice-chairman, Radiocommunication Advisory Group of International Telecommunication Union (ITU), vice-chairman of the UN Commission on Science and Technology for Development (CSTD), Geneva, Switzerland

Prof. Boris MIRKIN –

Professor of Department of Data Analysis and Artificial Intelligence, School of Applied Mathematics and Information Science, National Research University Higher School of Economics, Russian Federation

Prof. Vadim MOTIL –

Professor, Department of Information Security Management, Faculty of Cybernetics, Tula State University, Russian Federation

Prof. Dmitry PALCHUNOV –

Head of Department of General Informatics, Faculty of Information Technologies, Novosibirsk State University, Russian Federation

Prof. Panagote (Panos) PARDALOS –

Distinguished Professor and University of Florida Research Foundation Professor, Director of Center for Applied Optimization, Department of Industrial and Systems Engineering, University of Florida, USA

Dr. Albert SILANTYEV –

Professor, Department of Information Business Systems, Institute of Information Business Systems, National University of Science and Technology «MISIS», Russian Federation

Dr. Victor TARATOUKHIN –

Managing Director European Research Center for Information Systems (ERCIS) Competence Center ERP, Head of ERCIS Lab. Russia, University of Munster, Germany

Prof. Mikhail ULYANOV –

Professor of Software Management Department, School of Software Engineering, National Research University Higher School of Economics, Russian Federation

Dr. Alexey SHALCOVSKY –

Vice Rector, Director of Institute of Information Technologies, National Research University Higher School of Economics, Russian Federation

ФОРМАЛИЗАЦИЯ ЗАДАЧИ ВЫБОРА ОПТИМАЛЬНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ИЗМЕНЕНИЙ АРХИТЕКТУРЫ ПРЕДПРИЯТИЯ НА ОСНОВЕ МАТРИЦЫ ИЗМЕНЕНИЙ БРИНЬОЛФССОНА

В.А. Агиевич,

аспирант кафедры инноваций и бизнеса в сфере информационных технологий,
факультет бизнес-информатики, Национальный исследовательский университет
«Высшая школа экономики»

К.Г. Скрипкин,

кандидат экономических наук, доцент кафедры экономической информатики,
экономический факультет, Московский государственный университет
имени М.В. Ломоносова

E-mail: vagievich@hse.ru, k.skripkin@gmail.com

Адрес: г. Москва, ул. Мясницкая, 20

Несмотря на большое разнообразие методов и подходов к построению архитектуры предприятия, при их практическом применении проявляется ряд недостатков. Одним из наиболее существенных пробелов в этой области знаний является недостаточная проработка и слабая формализованность методов планирования перехода от текущего состояния архитектуры предприятия к целевому. При этом зачастую планирование перехода является творческим процессом, успех которого сильно зависит от опыта, интуиции, знания корпоративной культуры, истории предприятия. Кроме того, в крупных организациях процесс усложняется большим числом элементов архитектурных моделей, что делает затруднительным применение описанных в литературе методов.

В литературе по архитектуре предприятия также отмечается важность принятия во внимание взаимодействия между элементами архитектуры предприятия во время планирования миграции, но не приводятся методы, позволяющие учесть это взаимодействие. Аналогичная проблема решается в теории комплементарных активов.

Матрица изменений Э.Бриньолфссона – эффективный инструмент управления изменениями организации на основе теории комплементарных активов. Однако этот инструмент может применяться только в небольших проектах и для оценки отдельных крупных изменений. Причина – ограничение размера матрицы. В статье описывается математическая модель и соответствующая постановка задачи дискретной оптимизации, решение которой позволит снять это ограничение за счет использования математического аппарата вместо визуальной оценки при поиске оптимальной последовательности изменений.

Ключевые слова: управление изменениями, матрица изменений, комплементарные активы, дискретная оптимизация, архитектура предприятия.

Введение

Значение архитектуры предприятия в современных условиях постоянно увеличивается за счет обеспечения возможностей эффективного использования существующих технологий и эволюционного перехода к новейшим технологиям. Однако, несмотря на большое разнообразие методов и подходов к построению архитектуры предприятия, при их практическом применении проявляется ряд недостатков. Одним из наиболее существенных пробелов в этой области знаний является недостаточная проработка и слабая формализованность методов планирования перехода от текущего состояния архитектуры предприятия к целевому.

Анализ существующих методологий архитектуры предприятия и соответствующей специальной литературы [1] показывает, что описание процессов внедрения целевой архитектуры (миграции) представлено, как правило, общими рекомендациями и перечислением факторов, которые необходимо учитывать при планировании миграции. При этом в качестве одного из наиболее значимых факторов планирования архитектурных изменений указывается необходимость принятия во внимание разного рода зависимостей, однако не приводится достаточно строгих методов, позволяющих эффективно реализовать это указание на практике.

В [1] для решения данной проблемы авторами было предложено использовать матрицу изменений – инструмент управления изменениями, учитывающий комплементарные взаимосвязи между базовыми и целевыми практиками предприятия.

1. Теория комплементарных активов и матрица изменений

Необходимость учитывать взаимосвязи при планировании изменений на предприятии подтверждена многими исследованиями. В частности, еще в 1992 году они явились центральным предметом исследования теории комплементарных активов П.Милгрота и Дж.Робертса [2].

Согласно [2] комплементарность приводит к образованию предсказуемых связей между отдельными видами деятельности. Милгром и Робертс дают следующее определение комплементарности [2]: «Несколько видов деятельности считаются комплементарными, если увеличение объема любого из них увеличивает (или, по крайней мере, не уменьшает) предельную прибыльность каждого из всех остальных видов деятельности».

В исследованиях Бринйолфсона и более поздних источниках вместо понятия «актив» используется понятие «практика» или «организационная практика», определяемая как определенный способ решения задачи, стоящей перед организацией [3].

Как можно описать и проанализировать взаимосвязи различных организационных практик компании друг с другом, с ИТ-сервисами и со свойствами человеческого капитала? Для описания бизнес-процессов, организационной структуры, архитектуры информационных систем и данных в настоящее время разработан целый ряд моделей, широко применяемых как самостоятельно, так и при формировании архитектуры предприятия. Однако, как отмечают авторы [4], на сегодняшний день существует лишь одна модель, решающая задачу описания взаимосвязей между практиками – матрица изменений. Она была предложена Э.Бринйолфссоном и соавторами в работе «Матрица изменений» [5], [6] (рис. 1).

Матрица изменений состоит из двух таблиц, наложенных друг на друга. Каждая состоит из прямоугольной части – списка организационных практик и треугольной, содержащей данные о взаимосвязях между практиками. Знак «+» в ячейках треугольника означает комплементарность двух практик, знак «-» – что эти практики выступают по отношению друг к другу как конкуренты. Горизонтальная таблица описывает существующие практики, вертикальная – практики, внедряемые в ходе проекта. Данные о комплементарности организационных практик за-

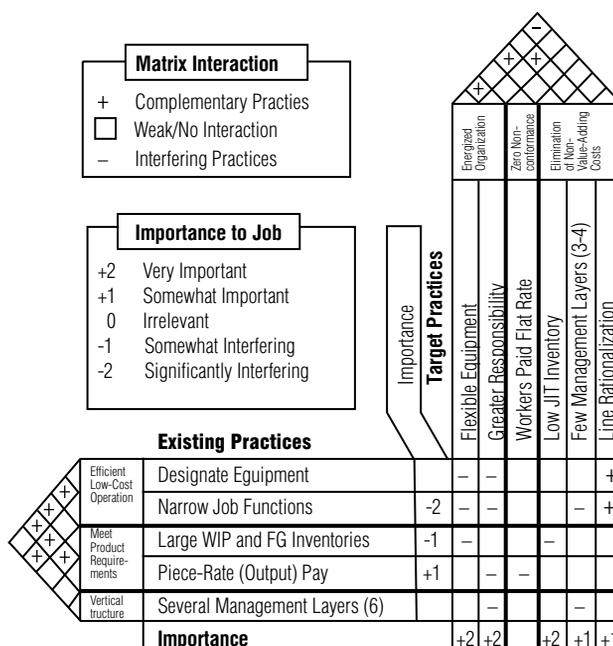


Рис. 1. Матрица изменений [5]

полняются на основе экспертных оценок сотрудников организации. Строка и столбец «Importance» описывают важность внедряемых практик по шкале Ликерта (от -2 – очень мешает до +2 – очень важна). Наконец, прямоугольник на пересечении двух таблиц описывает сочетаемость существующих и внедряемых практик и, соответственно, трудности перехода от «как есть» к «как будет» [3].

Матрица изменений – полезный инструмент для ответа на следующие типы вопросов [7].

- ◆ Выполнимость изменений.
- ◆ Последовательность выполнения изменений.
- ◆ Темп и характер изменений.
- ◆ Оценка изменений организаторами процесса.

Несмотря на всю простоту и эффективность подхода, заложенного в описанном инструменте, его применение для планирования изменений, связанных с переходом к целевой архитектуре предприятия осложняется большим объемом данных. Как показывает практика применения матрицы изменений, максимальная размерность матрицы, позволяющая работать с ней непосредственно, составляет 10x10.

Далее для решения проблемы размерности матрицы изменений описана формальная модель матрицы изменений и поставлена задача дискретной оптимизации, решение которой позволит работать с системами практик больших размерностей.

2. Построение математической модели матрицы изменений

Для построения формальной модели матрицы изменений опишем, прежде всего, этот инструмент в аспекте содержащихся в нем данных и ограничений, а также приемов, которые используются при его практическом применении.

Фактически при построении матрицы изменений задается два множества практик: базовые практики, заданные вертикальной треугольной матрицей, и целевые практики, заданные горизонтальной треугольной матрицей [5], рис. 2.

В матрице изменений, предложенной Э. Бринойлфссоном, взаимодействия между практиками описываются следующими отношениями:

◆ «+» между элементами множества практик обозначает, что соответствующие им практики взаимно усиливают друг друга; соответствующие практики называются комплементарными.

◆ «-» между элементами множества практик обозначает, что соответствующие им практики взаим-

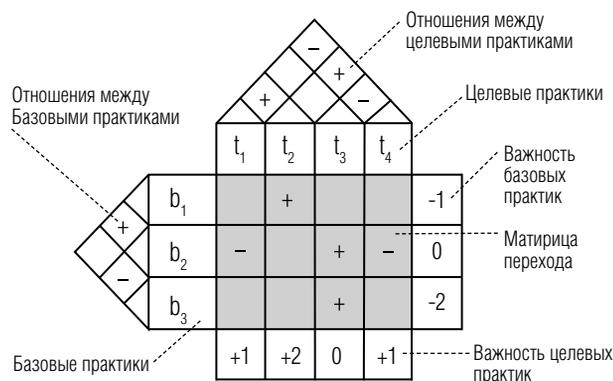


Рис. 2. Макет матрицы изменений

но ослабляют друг друга; соответствующие практики называются конкурирующими.

Если взаимодействия между практиками не определены, такие практики называются независимыми. Обозначим множество базовых практик $B = \{b_1, \dots, b_n\}$, где n – количество базовых практик, а множество целевых практик – $T = \{t_1, \dots, t_m\}$, где m – количество целевых практик.

Множество всех практик, входящих в матрицу, будем обозначать X :

$$X = B \cup T = \{x_1, \dots, x_{n+m}\}$$

Для элементов множества B отношения заданы горизонтальной треугольной матрицей (на рис. 2 слева). Для элементов множества T отношения заданы вертикальной треугольной матрицей (на рис. 2 сверху).

Аналогичные отношения установлены между некоторыми элементами множества T и некоторыми элементами множества B . Эти отношения заданы прямоугольной матрицей (на рис. 2 в центре).

Каждой базовой и целевой практике в матрице приписывается степень важности по шкале Ликерта (от -2 до +2).

Пусть степени важности практик множеств B и T заданы множествами V_b и V_t соответственно:

$$V_b = \{v_1, \dots, v_n\};$$

$$V_t = \{v_1, \dots, v_m\}.$$

При анализе подходов, используемых при работе с матрицей изменений, описанных в [5], выявлено, что некоторые решения, принимаемые на ее основе, требуют наличия дополнительных данных, которые отсутствуют в матрице. Такими дополнительными данными являются:

◆ порядок выполнения изменений, существующий для некоторых практик (примеры: «переходить к запасам точно в срок необходимо до переключения

на гибкое оборудование», «невозможно делегировать рабочим больше ответственности, когда надзор за ними осуществляется на всех уровнях»);

♦ явное указание целевой практики, на которую заменяется базовая практика (может существовать для некоторых базовых практик);

♦ степень взаимодействия между практиками (для выделения в дальнейшем блоков практик, которые необходимо внедрять или удалять одновременно).

Как отмечается в [4], одним из недостатков матрицы изменений, предложенной Э.Бринйолфссоном, является тот факт, что взаимосвязи между практиками описываются только качественно, хотя степень их взаимного влияния, а также влияния на результат может сильно отличаться. Для устранения этого недостатка в [4] предложена идея расширенной матрицы изменений, в которой вместо шкалы «+» – практики комплементарны», «-» – практики конкурируют» вводится следующая шкала:

- 2 при совместном использовании практики неработоспособны;
- 1 совместное использование практик снижает эффективность;
- 0 связи практик нет;
- +1 совместное использование практик повышает эффективность;
- +2 по отдельности практики неработоспособны.

Данная шкала предлагается к использованию в матрице изменений как в треугольных матрицах, создаваемых для базовых и целевых практик, так и в матрице перехода. При этом для матрицы перехода отношение «+2» является нерелевантным, поскольку рассматриваются взаимосвязи двух разных систем.

Явное соответствие между практикой, используемой в настоящее время, и практикой, на которую она будет заменена, может быть задано при помощи значения «-2» в переходной матрице.

Таким образом, вместо матрицы изменений Э.Бринйолфссона для исследования была принята расширенная матрица изменений (рис. 3), дополнен-

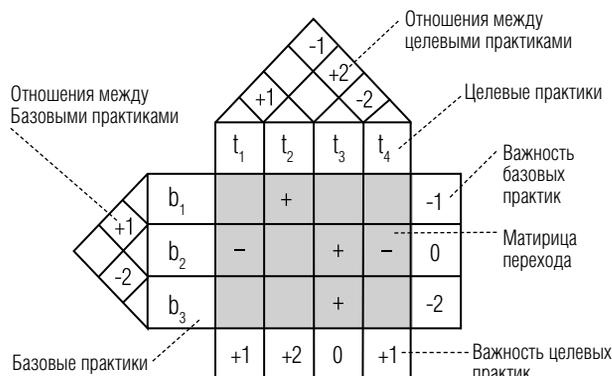


Рис. 3. Макет расширенной матрицы изменений

ная информацией о порядке выполнения изменений.

Определим шкалу степеней взаимосвязи практик как множество $R = \{r_1, \dots, r_5\} = \{-2, -1, 0, +1, +2\}$. Взаимосвязи практик, представленные в расширенной матрице изменений, также могут быть заданы в виде графа или при помощи функций:

$$r(b_i, b_j) = r_i, r(t_i, t_j) = r_i, r(b_i, t_j) = r_i,$$

$$\text{или } r(x_i, x_j) = r_l, \text{ где } l \in [1, 5].$$

Порядок выполнения изменений, существующий для некоторых практик, может быть задан дополнительно к матрице изменений в виде отношений частичного порядка:

$$b_i \leq b_j, \text{ где } i, j \in [1, n];$$

$$t_i \leq t_j, \text{ где } i, j \in [1, m];$$

$$b_i \leq t_j, \text{ где } i \in [1, n], j \in [1, m].$$

Модель расширенной матрицы изменений можно представить в виде взвешенного неориентированного раскрашенного графа с весами ребер от -2 до +2 («0» не используется) и весами вершин от -2 до +2 (рис. 4).

При этом базовое состояние системы практик представлено подграфом G_b , а целевое состояние системы практик представлено подграфом G_t .

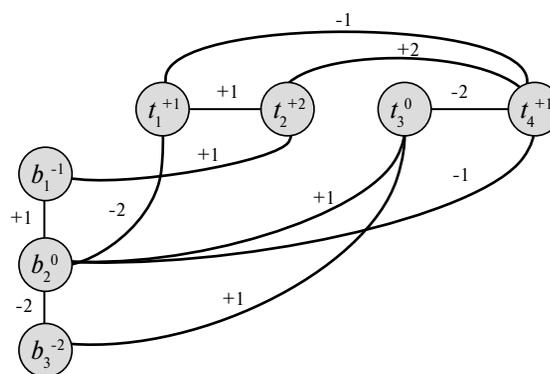


Рис. 4. Представление расширенной матрицы изменений в виде графа

3. Постановка задачи дискретной оптимизации

Определение. Элементарным преобразованием графа системы практик (или просто элементарным преобразованием) c_i будем называть внедрение целевой практики $t \in T$ или удаление базовой практики $b \in B$. Элемент множества X , соответствующий элементарному преобразованию c_i будем определять при помощи функции $X(c_i)$.

Определение. Последовательность элементарных преобразований будем называть траекторией $Tr = \{c_1, \dots, c_{n+m}\}$, $c_i \succ c_{i+1}$. Множество всех возмож-

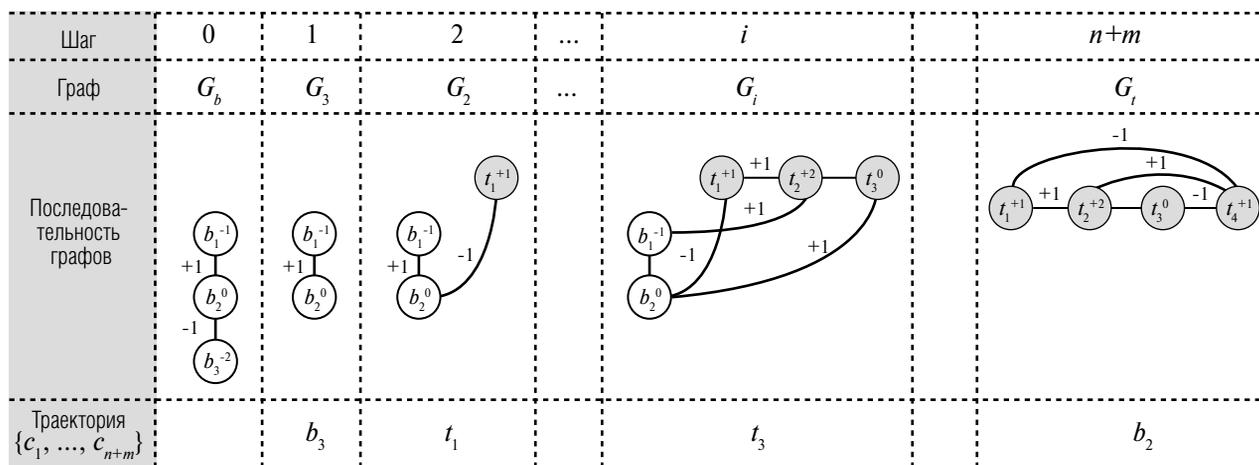


Рис. 5. Преобразование графа базовых практик в граф целевых практик

ных траекторий от графа G_b к графу G_t обозначим через $Tr(G_b, G_t)$.

Согласно [5] целевая функция должна отражать «легкость» изменений: «Легче всего ликвидируются практики, конкурирующие с другими базовыми практиками. Легче внедряются те целевые практики, которые комплементарны базовым практикам».

«Легкость» внедрения или удаления практик определяется на основе теории комплементарных активов при помощи суммирования показателей взаимосвязи между внедряемой или удаляемой практикой и всеми практиками текущей системы практик. Для внедряемой целевой практики как сумма величины взаимосвязей между внедряемой практикой и всеми практиками текущей системы практик.

«Легкость» изменения c_i будем измерять относительно текущего состояния системы практик, представленного графом $G_i = (X, V)$, соответствующим графу G_b с выполненными изменениями $c_1 \dots c_{i-1}$. При этом «легкости» внедрения новой целевой практики будет соответствовать функция

$$l(c_i) = \sum_{j=1}^n r(X(c_i), b_j) + \sum_{j=1}^{i-1} \begin{cases} -r(X(c_i), X(c_j)), & \text{если } X(c_j) \in B \\ r(X(c_i), X(c_j)), & \text{если } X(c_j) \in T \end{cases}$$

а «легкости» исключения базовой практики — та же функция с противоположным знаком. То есть функция «легкости» для изменения c_i есть сумма значений взаимодействий между практикой $X(c_i)$ и всеми практиками текущей системы практик G_i (рис. 5). Следовательно, суммарный показатель «легкости» изменений Tr_i на траектории может быть вычислен как

$$L(Tr_i) = \sum_{j=1}^{n+m} \begin{cases} l(c_j), & \text{если } X(c_j) \in B \\ -l(c_j), & \text{если } X(c_j) \in T \end{cases}$$

Таким образом, для решения задачи построения оптимальной последовательности изменений, содержательно описанной в [1], необходимо найти такую последовательность преобразований от графа G_b к графу G_t (траекторию $Tr(G_b, G_t)$), чтобы суммарная «легкость» перехода $L(Tr_i)$ была максимальной, а максимальное значение «трудности» отдельного преобразования $(-l(c_i))$ было минимальным. Минимаксный критерий вводится для исключения скачков трудности отдельных преобразований.

Определим теперь ограничения задачи.

Как было отмечено выше, для некоторых практик существуют отношения частичного порядка, определяемые в общем виде как $x_i \leq x_j$, где $i, j \in [1, n+m]$.

Также, в соответствии с [5], необходимо обеспечить удаление базовой практики и внедрение целевой, если целевая практика должна заменить базовую (в расширенной матрице изменений эта связь показана как «-2»). Значит, мы должны выстроить соответствующие изменения так, чтобы между ними не выполнялись никакие другие. Таким образом, для всех пар практик b, t , таких что $r(b, t) = -2$, должно быть задано ограничение

$$\begin{aligned} X(c_i) = b \text{ и } X(c_{i+1}) = t \\ \text{или } X(c_i) = t \text{ и } X(c_{i+1}) = b. \end{aligned}$$

В итоге оптимизационная задача формулируется в следующем виде:

Найти такую траекторию $Tr_i \in Tr(G_b, G_t)$, что

$$L(Tr_i) \rightarrow \max,$$

$$\max_{1 \leq j \leq n+m} (-l(c_j)) \rightarrow \min, c_j \in Tr_i$$

при ограничениях $x_i \leq x_j$, где $i, j \in [1, n+m]$

$$X(c_i) = b \text{ и } X(c_{i+1}) = t (b, t: r(b, t) = -2).$$

Заключение

Авторами статьи описана математическая модель матрицы изменений и поставлена задача дискретной оптимизации, решение которой позволит формализовать процесс нахождения оптимальной по-

следовательности изменений на основе матрицы.

На следующем этапе исследования планируется решение данной задачи методами комбинаторной оптимизации с апробацией результатов на реальном бизнес-кейсе построения архитектуры предприятия. ■

Литература

1. Агиевич В. А., Скрипкин К. Г., Гимранов Р. Д. Матрица изменений Бринйолфссона как инструмент планирования архитектуры предприятия // Современные информационные технологии и ИТ-образование / Сборник избранных трудов VIII международной научно-практической конференции. Москва: ИНТУИТ.ру, 2013. С. 785-794.
2. Roberts J., Milgrom P. Economics, Organization, and Management. EnglewoodCliffs, NJ: PrenticeHall. 1992.
3. Скрипкин К. Организационный капитал российских предприятий: проблема разнообразия // Инновационное развитие экономики России: региональное разнообразие: Шестая международная научная конференция. Сборник статей. Москва: ТЕИС, 2013. Т. 1. С. 286-294.
4. Лугачев М., Скрипкин К., Ананьин В., Зимин К. Эффективность инвестиций в ИТ. Альманах лучших работ. Москва: СоДИТ, 2013. С. 178-186.
5. Brynjolfsson E., Renshaw A., Alstynе M. The Matrix of Change // Sloan Management Review. 1997. Vol. 38.No. 2.
6. Хромов-Борисов С. Управление сложностью. Операционная система бизнеса. Москва: ИД Гребенникова, 2012. С. 143-183.
7. Щегельская О. «Матрица изменений» – инструмент управления трансформацией компании // Управление компанией. 1999. № 5.

FORMALIZATION OF THE PROBLEM OF CHOOSING THE OPTIMAL SEQUENCE OF ENTERPRISE ARCHITECTURE CHANGES ON BASIS OF BRYNJOLFSSON'S MATRIX OF CHANGE

Vadim AGIEVICH,

Postgraduate Student, Department of Innovation and Business in Information Technologies,
Faculty of Business Informatics, National Research University Higher School of Economics

Address: 20, Myasnitskaya str., Moscow, 101000, Russian Federation

E-mail: vagievich@hse.ru

Kirill SKRIPKIN,

Associate Professor, Department of Economic Informatics, Faculty of Economics,
Lomonosov Moscow State University

Address: 1, Leninskie Gory, Moscow 119991, Russian Federation

E-mail: k.skripkin@gmail.com

Despite the great variety of methods and approaches to Enterprise Architecture development, their application in practice reveals a number of shortcomings. One of the most significant gaps in this area of knowledge is insufficient study and weak formalization of planning of transition from the current state to the target Enterprise Architecture. That is why transition planning is often a creative process, its success depending much on experience, intuition, knowledge of corporate culture, the history of a company. Besides, in big companies the process is complicated by a great number of elements of architecture models, which makes it more demanding to implement the methods described in literature.

EA literature describes also the importance of taking into consideration the interactions between EA elements during

migration planning, but do not express the methods of this. Similar problem is solved by the theory of complementary assets.

Brynjolfsson's Matrix of Change is an effective tool for managing organizational change based on the theory of complementary assets. However this tool can be used only in small projects or for assessment of individual consolidated changes. The reason is the limited size of the matrix. The paper describes the mathematical model and the corresponding discrete optimization problem formulation, the solution of which will overcome this restriction by using the mathematical apparatus instead of visual assessment when searching for the optimal sequence of change.

Key words: change management, matrix of change, complementary assets, discrete optimization, Enterprise Architecture.

References

1. Agievich V. A., Skripkin K. G., Gimranov R. D. (2013) Matrica izmenenij Brinjolfssona kak instrument planirovanija arhitektury predpriyatija [The Brynjolfsson's Matrix of Change as a Tool for Enterprise Architecture Planning]. *Sovremennye informacionnye tehnologii i IT-obrazovanie. Sbornik izbrannyh trudov VIII mezhdunarodnoj nauchno-prakticheskoj konferencii* [Modern Information Technology and IT Education / Collection of Selected Works of the VIII International Scientific and Practical Conference], Moscow: INTUIT.ru, pp. 785-794. (in Russian)
2. Roberts J., Milgrom P. (1992). *Economics, Organization, and Management*. Englewood Cliffs, NJ: Prentice Hall.
3. Skripkin K. (2013) Organizacionnyj kapital rossijskih predpriyatij: problema raznobrazija [Organizational Capital of Russian Companies: The Problem of Diversity]. *Innovacionnoe razvitie ekonomiki Rossii: regional'noe raznobrazije: Shestaja mezhdunarodnaja nauchnaja konferencija. Sbornik statej* [Innovative Development of the Russian Economy: Regional Diversity. Sixth International Scientific Conference. Collection of articles], Moscow: TEIS, vol. 1, pp. 286-294. (in Russian)
4. Lugachev M., Skripkin K., Anan'in V., Zimin K. (2013) *Effektivnost' investicij v IT. Al'manah luchshih rabot* [Effectiveness of IT Investments. Almanac of the Best Works], Moscow: SoDIT, pp. 178-186. (in Russian)
5. Brynjolfsson E., Renshaw A., Alstynne M. (1997) The Matrix of Change. *Sloan Management Review*, vol. 38, no 2.
6. Hromov-Borisov S. (2012) *Upravlenie slozhnost'ju. Operacionnaja sistema biznesa* [Complexity Management. The Operation System of Business], Moscow: ID Grebennikova, pp. 143-183. (in Russian)
7. Schegel'skaya O. (1999) «Matrica izmenenij» – instrument upravlenija transformaciej kompanii [The Matrix of Change – a Tool for Company Transformation Management]. *Company Management*, no 5. (in Russian)

ФОРМИРОВАНИЕ ПОРТФЕЛЯ ПРОЕКТОВ НА ОСНОВЕ НЕЧЕТКОЙ МОДЕЛИ МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ

С.М. Авдошин,

*профессор, руководитель отделения программной инженерии,
факультет бизнес-информатики, Национальный исследовательский
университет «Высшая школа экономики»*

А.А. Лифшиц,

*студент магистратуры отделения программной инженерии,
факультет бизнес-информатики, Национальный исследовательский
университет «Высшая школа экономики»*

Адрес: 101000, г. Москва, ул. Мясницкая, д. 20

E-mail: savdoshin@hse.ru, alexeus1992@yandex.ru

Компании, являющиеся лидерами IT-индустрии, ведут от нескольких десятков до нескольких сотен проектов одновременно. Отбор соответствующих стратегическим целям компании и удовлетворяющих ресурсным ограничениям проектов является важной задачей процесса управления портфелями проектов. Таким образом, задачей формирования портфеля проектов является выбор множества проектов, которые лучшим образом отвечают целям компании в условиях ресурсных ограничений компании. В представленной работе предложена многокритериальная математическая модель формирования портфеля проектов в терминах нечетких множеств.

Приводится обзор существующих методов решения многокритериальных детерминированных задач формирования портфеля проектов. Обосновывается выбор методов муравьиной оптимизации и генетического алгоритма в качестве основных для обобщения на случай нечетких множеств. Описывается реализация муравьиной оптимизации, основанной на минимаксной системе с одной структурой феромонов и одной колонией. Рассматриваются вариации с бинарной турнирной и ранговой функциями селекции алгоритма SPEA II применительно к данной задаче. Предлагается модификация алгоритма, основанная на генерации части начальной популяции неслучайным образом.

Приводятся данные численных экспериментов для алгоритма муравьиной оптимизации и вариаций генетического алгоритма. В качестве параметров сравнения взяты скорость выполнения и С-метрика. Результаты показали превосходство алгоритма, использующего неслучайный механизм генерации начальной популяции. Таким образом, для решения задачи формирования портфеля проектов предлагается использовать данный алгоритм.

Ключевые слова: портфель проектов, многокритериальная модель, нечеткие числа, генетический алгоритм, муравьиная оптимизация.

Введение

В настоящее время компании, являющиеся лидерами IT-индустрии, ведут от нескольких десятков до нескольких сотен проектов одновременно. Причем проекты могут отвечать различным стратегическим целям компании. Следуя [1], под проектом мы будем понимать уникальный набор процессов, включающий в себя управляемые задачи, даты начала и завершения, предпринятых для достижения конкретной цели.

Часто при реализации проектов не учитываются миссия компании и ее стратегические цели, которые могут быть прямо не связаны с основной деятельностью компании, такие как увеличение конкурентоспособности на рынке, развитие новых технологий, и т.д. В результате этого может возникнуть ситуация, когда прибыльные проекты, не отвечающие целям компании, превосходят менее прибыльные проекты, пусть и в полной мере отвечающие стратегическим целям компании. Согласно статистике [2], лишь около 20% инициатив руководства, призванных проследить за выполнением стратегических целей, реализуются. Для решения проблемы учета стратегических целей компании вводится процесс управления портфелем проектов. Здесь под портфелем проектов мы понимаем набор проектов (не обязательно технологически зависимых), реализуемых организацией в условиях ресурсных ограничений и обеспечивающих достижение стратегических целей организации [2].

Цель данной работы – решение проблемы формирования портфеля проектов, являющейся первым этапом в процессе управления портфелем проектов. При этом, во-первых, нужно формализовать математическую модель задачи формирования портфеля проектов и, во-вторых, нужно выбрать наиболее эффективные методы решения этой задачи.

Математическая модель должна учитывать различные стратегические цели компании, а также оперировать с различными ресурсами.

Для решения вопроса о включении проекта в портфель требуется первоначальная оценка его соответствия различным целям компании и оценка требуемых в процессе реализации ресурсов. Однако на ранних этапах фактически невозможно определить точные числовые значения и параметры конкретных проектов, поскольку точной информации о финансовых потоках и ресурсных затратах орга-

низации не имеют. Для прогнозирования показателей соответствия целям могут использоваться разнообразные методы оценивания, позволяющие обеспечить некоторую точность задания данных проекта. При этом, часто заключительная оценка финансовых показателей состоит из трех чисел: минимальной оценки, максимальной оценки и наиболее вероятной. В случае, когда оцениваются лингвистические (нефинансовые) индикаторы проекта (например, степень технологической новизны), чаще всего используются экспертные оценки, состоящие также из трех показателей: минимального (наименее благоприятного), максимального (наиболее благоприятного) и наиболее вероятного. Таким образом, представление данных в виде трех оценок является удобным и позволяет учитывать граничные сценарии развития проекта. В случае если в задаче формирования портфеля проектов будут использованы стандартные данные, где невозможно указать наличие трех оценок по каждому показателю, можно взять наиболее вероятную оценку или некоторую агрегированную величину, учитывающую все три оценки. Отметим, что и при выборе наиболее вероятной оценки, и при замене ее агрегированной величиной часть данных теряется, и это обстоятельство может негативно сказаться на точности получаемого решения.

Присутствие многокритериального аспекта целеполагания, позволяет сформулировать задачу формирования сбалансированного портфеля с учетом различных целей организации. Решением будет являться множество портфелей, окончательный выбор из которых может быть сделан экспертами компании.

1. Модели задачи

А. Обзор существующих моделей

В работе [3] предложена многокритериальная модель формирования портфеля с одним ограничением. На первом шаге, авторы предполагают экспертную оценку проектов, в результате которой каждому из проектов ставится в соответствие одна из трех категорий: приоритет, удовлетворительность или приемлемость, также внутри каждой из групп проекты сортируются согласно мере соответствия своей категории. Для модели авторы предлагают использовать 6 целевых функций: количество проектов в возможном решении для каждой из трех групп, суммарное количество проектов, и для категорий приоритета и удовлетворительности – оцен-

ки, основанные на мере соответствия. Также автора учитывают бюджетное ограничение компании.

Lean Yu рассматривает в своей работе [4] одно-критериальную модель формирования портфеля проектов. Целевая функция состоит из двух частей: первая часть представляет собой оценку независимого эффекта проектов (для каждого проекта идет расчет взвешенной суммы меры соответствия проекта целям организации), а вторая основывается на эффекте от совместной реализации нескольких проектов. Также количество проектов в портфеле задается в качестве ограничения модели.

В. Предложенная математическая модель задачи

Введем обозначения:

- ◆ N – множество доступных проектов, $|N| = n$
- ◆ M – множество ресурсов, $|M| = m$
- ◆ K – множество критериев (стратегических целей компании), $|K| = k$
- ◆ Q – портфель проектов, $|Q| = q, Q \subseteq N$.

В рамках данной работы мы будем отождествлять множество проектов, ресурсов и критериев с множеством их номеров.

Проект i из множества доступных проектов N характеризуется:

- ◆ вектором соответствия критериям

$$a_i = (a_{1i}, a_{2i}, \dots, a_{ki}), \text{ где}$$

a_{il} – численная характеристика соответствия проекта i критерию l

- ◆ вектором требуемых ресурсов

$$c_i = (c_{1i}, c_{2i}, \dots, c_{mi}), \text{ где}$$

c_{ji} – потребность в ресурсе j для реализации проекта i

Портфель проектов Q характеризуется:

- ◆ вектором соответствия критериям

$$a_Q = (a_{1Q}, a_{2Q}, \dots, a_{kQ}),$$

где $a_{lQ} = \sum_{i=1}^q a_{li}, l \in K$

- ◆ вектором требуемых ресурсов

$$c_Q = (c_{1Q}, c_{2Q}, \dots, c_{mQ}),$$

где $c_{jQ} = \sum_{i=1}^q c_{ji}, j \in M$.

Здесь мы предполагаем наличие принципа аддитивности значений проектов, согласно которому вектор соответствия критериям компании и вектор требуемых ресурсов равны сумме соответствующих векторов проектов, входящих в портфель. Также в данной модели не рассматриваются зависимости между проектами.

Ресурсные ограничения компании будем задавать вектором ограничений по ресурсам, который обозначим через R , где $R = (R_1, R_2, \dots, R_m)$. При этом портфель проектов должен удовлетворять неравенству $c_Q \leq R$ покомпонентно.

Требуется максимизировать степени соответствия стратегическим целям компании.

$$\begin{cases} \sum_{i=1}^N a_{1i} x_i \rightarrow \max \\ \dots \\ \sum_{i=1}^N a_{li} x_i \rightarrow \max \\ \dots \\ \sum_{i=1}^N a_{ki} x_i \rightarrow \max \end{cases} \quad (1)$$

$$\sum_{i=1}^N c_{ji} x_i \leq R_j, j \in M$$

Здесь $x_i \in \{0, 1\}$, $x_i = 1$, если проект включен в портфель и $x_i = 0$, в противном случае.

Компоненты векторов ресурсных ограничений, соответствия критериям и требуемых ресурсов будем задавать с помощью нечетких трапециевидных чисел. Это позволит учесть недостаток информации на ранних этапах.

Нечеткое трапециевидное число A будем задавать в виде функции принадлежности $\mu_A: R \rightarrow [0, 1]$ определяемой следующей формулой:

$$\mu_A(x) = \begin{cases} 0, & x < a_1 \text{ или } x > a_4 \\ \frac{x - a_1}{a_2 - a_1}, & a_1 \leq x < a_2 \\ 1, & a_2 \leq x \leq a_3 \\ \frac{a_4 - x}{a_4 - a_3}, & a_3 < x \leq a_4 \end{cases} \quad (2)$$

Таким образом, нечеткое трапециевидное число задается с помощью четырех упорядоченных по возрастанию параметров a_1, a_2, a_3, a_4 .

Заметим, что операции сложения и вычитания являются внутренними бинарными операциями, определенными на множестве параметрически заданных нечетких трапециевидных чисел. Таким образом, можно проводить вычисления над такими числами, не меняя их представление [5].

С. Сравнение моделей

Модель [3] основана на экспертной оценке и сортировке проектов относительно друг друга и не учитывает возможные количественные результаты проекта (чистый приведенный доход, внутренняя

Таблица 1.

Сравнительный анализ моделей формирования портфеля проектов

Модель	Количество целевых функций	Количество ограничений	Учет нечетких оценок	Формат решения	Учет зависимостей между проектами
Модель [2]	6	1	нет	множество возможных портфелей проектов	нет
Модель [3]	1	1	нет	портфель проектов	да
Предложенная модель	Не ограничено	Не ограничено	да	множество возможных портфелей проектов	нет

норма доходности) и оценки экспертов, выраженные в количественной форме, что может негативно повлиять на результирующие портфели проектов.

В качестве недостатка модели [4] можно отметить однокритериальность, что делает результат сильно зависимым от выбранных весов целей и отсутствие учета ресурсных ограничений компании.

Приведем сравнительную характеристику результатов анализа, см. табл. 1.

На основании сравнительного анализа можно заключить, что предложенная математическая модель учитывает ряд важных факторов проблемы формирования портфеля проектов, что доказывает ее практическую значимость.

2. Методы решения

А. Обзор существующих подходов

Рассматриваемые в данном разделе подходы применялись к решению задачи без использования нечетких чисел.

Метод перебора портфелей [2] является адаптацией метода ветвей и границ. Асимптотическая временная сложность данного метода $O(2^n)$, где $n = |N|$.

Адаптивный метод с эpsilon вариацией ограничений [6] позволяет определять оптимальные по Парето решения на основе результатов полученных от решения задач однокритериальной оптимизации. На каждой итерации определяет оптимальное ресурсное ограничение. Асимптотическая временная сложность метода $O(k^{m-1} \cdot T)$, где k – количество найденных Парето-оптимальных решений, m – количество целей, а T – время решения задачи однокритериальной оптимизации.

Муравьиная оптимизация. Данный метод инспирирован поведением колонии муравьев, а именно тем как они ищут пищу, выбирая определенный путь, муравей оставляет феромоны, чем больше муравьев выбрало данный путь, тем больше количество оставленных там феромонов, причем со

временем их количество уменьшается. Возможные пути представляются структурой феромонов, в которой вершины являются проектами, а значения на ребрах пути являются количеством феромонов. Ключевой особенностью является обновление в конце итерации ребер, которые входили в состав Парето-оптимальных решений поколения [7].

Генетический алгоритм. Алгоритм основан на эволюционных методиках, соответственно популяция состоит из особей, которые являются возможными решениями. Причем каждая особь хранит данные в генотипе и в фенотипе. В данном случае генотипом особи будет являться битовая строка, состоящая из нулей и единиц, показывающих входит тот или иной проект в текущее решение. Фенотип будет интерпретироваться двумя векторами портфеля проектов: вектором ресурсных ограничений и вектором соответствия целям. При использовании модификации SPEA II, которая является одной из лучших среди генетических методов для многокритериальных задач [8], асимптотическая временная сложность для каждой итерации составляет $O(m^3)$, где m – сумма особей в популяции и в архиве.

В. Выбор методов

Для решения задачи формирования портфеля проектов были выбраны методы муравьиной оптимизации и модификация генетического алгоритма SPEA II. Выбор в пользу этих методов сделан, поскольку они имеют лучшую временную сложность, чем метод перебора портфелей и адаптивный метод с эpsilon вариацией ограничений.

С. Алгоритм муравьиной оптимизации

В данной работе для решения задачи формирования портфеля проектов предлагается алгоритм, реализующий методы муравьиной оптимизации, изложенные в [7]. Общая структура данного алгоритма выглядит следующим образом:

1. Инициализация структуры феромонов.

Каждое ребро структуры феромонов инициализируется заранее заданным значением τ_{max} .

2. Для каждого муравья строим решение.
3. Обновление структуры феромонов.

В полученном множестве возможных решений удаляем доминируемые по Парето и обновляем ребра на структуре феромонов, которые входят в оставшиеся возможные решения. Значения всех ребер между выбранными в таком муравье вершинами увеличиваются на фиксированное значение τ_{upd} .

4. Проверка удовлетворения ограничений на ребрах структуры феромонов

$$\begin{aligned} \text{Если } \tau_{cur} > \tau_{max}, \tau_{cur} &= \tau_{max}. \\ \text{Если } \tau_{cur} < \tau_{min}, \tau_{cur} &= \tau_{min}. \end{aligned}$$

5. Если заданное количество итераций не выполнено, то переходим к пункту 2, иначе заканчиваем работу алгоритма.

Определим отношение доминирования по Парето: портфель A доминирует портфель B , если хотя бы по одному критерию он лучше, а по остальным критериям не хуже, чем B .

Остановимся на пункте 2 и покажем, каким образом создается решение для каждого агента (муравья) применительно к задаче формирования портфеля проектов.

1. Инициализация множеств:

$$S = \emptyset;$$

$$Cand = V;$$

$R^* = R$, где V – множество проектов, R – ресурсные ограничения компании, а S – представляет путь муравья.

2. Проверка на завершение:

Если $Cand = \emptyset$ то решение построено.

3. Добавление решения

Выбираем $\{v_i\} \in Cand$ с вероятностью $p_s(v_i)$;

$$S = S + \{v_i\}.$$

4. Обновление параметров:

$$Cand = Cand - \{v_i\};$$

$$R_j^* = R_j^* - c_{jv_i}, j = 1, \dots, m,$$

где c_{jv_i} – требования к ресурсу типа j проекта v_i .

5. Проверка на нарушение ресурсных ограничений R_j^* :

Проекты из множества $Cand$, которые нарушают текущие ресурсные ограничения, удаляются.

6. Переход к пункту 2.

Далее рассмотрим метод расчета вероятности, с которой муравей будет выбирать следующий проект из числа возможных.

Вероятность $p_s(v_i)$ для каждого агента рассчитывается следующим образом:

$$p_s(v_i) = \frac{[\tau_S^{R^*}(v_i)]^\alpha \cdot [\sigma_S^{R^*}(v_i)]^\beta}{\sum_{v_j \in Cand} [\tau_S^{R^*}(v_j)]^\alpha \cdot [\sigma_S^{R^*}(v_j)]^\beta}, \quad (3)$$

где $\tau_S^{R^*}(v_i)$ – фактор феромонов, а $\sigma_S^{R^*}(v_i)$ – эвристический фактор.

Фактор феромонов вычисляется на основе структуры феромонов:

$$\tau_S^{R^*}(v_i) = \sum_{v_j \in S} \tau(v_i, v_j).$$

Т.е., это сумма по количеству феромонов, лежащих на ребрах соединения данного проекта с уже выбранными текущим муравьем проектами.

Эвристический фактор задачи формирования портфеля проектов будем вычислять следующим образом:

1. Коэффициент мощности оставшихся ресурсов:

$$h_S^{R^*}(v_i) = \sum_{j=1}^m \frac{c_{ji}}{R_j^*}, \quad (4)$$

где R_j^* – оставшееся количество ресурса j , а c_{ij} – количество требуемого ресурса j проекта i .

2. Итоговый эвристический фактор:

$$\sigma_S^{R^*}(v_i) = \frac{a_i}{h_S^{R^*}(v_i)}, \quad (5)$$

где a_i – случайно выбранная цель проекта i .

D. Генетический алгоритм

Предлагаемый альтернативный алгоритм для решения задачи формирования портфеля проектов основан на общей схеме генетического алгоритма SPEA II [8].

1. Инициализация популяции и архива.
2. Генерация начальной популяции.
3. Расчет функции приспособленности для текущей популяции и архива.
4. Отбор лучших решений в архив (усечение архива при необходимости).
5. Останов при выполнении заданного количества итераций.
6. Селекция (отбор решений) для последующего выполнения генетических операторов.

7. Формирование нового поколения и переход к пункту 3.

Алгоритм для расчета функции состоит из нескольких шагов:

1. Расчет «силы» решения, сколько решений данная хромосома доминирует из всей популяции плюс архива:

$$S_i = |\{j | j \in P_i + P_i^- \cap i >_{\text{Парето}} j\}|, \quad (6)$$

где $i = 1, \dots, n$, $+$ — объединение множеств и $>_{\text{Парето}}$ показывает превосходство по отношению Парето.

2. На основе S_i рассчитывается следующий показатель приспособленности, показывающий количество решений, которые доминируют данное решение.

$$R_i = \sum_{j \in P_i + P_i^-, j >_{\text{Парето}} i} S_j, \quad (7)$$

где $i = 1, \dots, n$.

3. Расчет δ_i (дистанции до k соседа, где $k = \sqrt{N + N^-}$, где N — размер популяции, а N^- — размер архива.

4. Расчет плотности решения:

$$D_i = \frac{1}{\delta_i + 2}. \quad (8)$$

5. Итоговая функция приспособленности $F_i = R_i + D_i$, в данном варианте целевая функция минимизируется.

Т.к. каждая особь представляет собой возможный портфель проектов, определим алгоритм расчета фенотипа применительно к задаче формирования портфеля проектов.

1. Инициализируем вектора требуемых ресурсов и соответствия целям пустыми значениями:

$Res = \emptyset$, где Res — вектор требуемых ресурсов для особи (возможного портфеля);

$Goal = \emptyset$, где $Goal$ — вектор соответствия целям для особи (возможного портфеля).

2. Для всех проектов $i, i \in N$.

3. Проверим, входит ли выбранный проект в генотип особи, если да, то перейдем к пункту 4, иначе выбираем следующий проект.

4. Выполним обновление параметров:

$$Res = Res + a_i;$$

$$Goal = Goal + c_i.$$

5. Проверим удовлетворение ограничениям компании R . Если $Res \leq R$, то выбираем следующий проект, иначе в пункт 6.

6. Будем выбирать случайным образом проект из

списка всех возможных, пока не выберем проект, входящий в данную особь. Удалим его из генотипа особи и переходим к пункту 1.

Рассмотрим следующие механизмы отбора решений: ранговый, и бинарный турнирный.

Ранговая селекция. Здесь применяется следующая последовательность действий:

1. Все решения сортируются.

2. Худшему решению ставится в соответствие ранг 1, следующему 2 и так далее.

3. Особь выбирается с фиксированной вероятностью $pi/total$, где pi — ранг данной особи, а $total$ — сумма рангов по всем особям.

Бинарная турнирная селекция. Данный вид состоит из следующих шагов:

1. Выбор двух особей из популяции.

2. Они сравниваются согласно значениям функции приспособленности.

3. Особь с меньшим значением (лучшая в контексте задачи формирования портфелей проектов) выбирается с фиксированной вероятностью p ($p > 80\%$). Другая особь выбирается соответственно с вероятностью $1 - p$.

В стандартном варианте генерации начальной популяции, генотип особей в ней задается случайным образом. Это приводит к тому, что решения, полученные на первых итерациях алгоритма, обладают малым значением фенотипа. Мы предлагаем иной способ генерации начальной популяции, при котором часть особей задается неслучайным образом. Решив однокритериальную задачу о формировании портфеля проектов для каждого из критериев и каждого из ограничений, мы получаем особи, которые Парето-доминируют большую часть полученных случайным образом. В дополнение к этому мы используем жадный алгоритм решения однокритериальной задачи со многими ограничениями [9].

3. Численные эксперименты

Численные эксперименты для выявления наиболее эффективного алгоритма проводились над четкими моделями, поскольку использование нечеткости одинаково влияет на каждый из методов.

Приведем сравнение скорости выполнения алгоритмов в зависимости от числа проектов. В экспериментах использовались тестовые объекты из 100, 250 и 500 проектов при трех целях и пяти ограничениях.

Таблица 2.

Время работы алгоритмов

Название алгоритма	Время работы алгоритма, % (относительно SPEA-II) (бинарная турнирная селекция)		
	100 проектов	250 проектов	500 проектов
Алгоритм муравьиной оптимизации	200	1800	15900
SPEA-II (бинарная турнирная селекция)	100	100	100
SPEA-II (ранговая селекция)	102	101	100
SPEA-II (сильная начальная популяция)*	104	102	103

* В данной реализации использовался ранговый механизм селекции

На основании данных *табл. 2* можно заключить, что алгоритм муравьиной оптимизации демонстрирует сильную зависимость от количества проектов и при значительном числе проектов является неэффективным, требуя значительных временных затрат, уже при 250 проектах серьезно проигрывая модификациям генетического алгоритма. Они, в свою очередь, слабо зависят от общего числа проектов, показывая примерно одинаковое время.

Также произведем сравнение алгоритмов попарно на основе рассчитанной *C*-метрики [7]. Для двух множеств решений X' и X'' , *C*-метрика рассчитывается следующим образом:

$$C(X', X'') = \frac{|\{a'' \in X'' : \exists a' \in X', a' >_{\text{Парето}} a''\}|}{|X''|}$$

Таблица 3.

C-метрика

Сравниваемые алгоритмы	Средняя C-метрика		
	100 проектов	250 про- ектов	500 про- ектов
SPEA-II (бинарная турнирная селекция), SPEA-II (ранговая селекция)	(0,21;0,19)	(0,23;0,18)	(0,31;0,54)
SPEA-II (бинарная турнирная селекция), SPEA-II (сильная начальная популяция)	(0,2;0,21)	(0,18;0,21)	(0,04;0,67)
SPEA-II (ранговая селекция), SPEA-II (сильная начальная популяция)	(0,21;0,23)	(0,17;0,21)	(0;0,97)
Алгоритм муравьиной оптимизации, SPEA-II (сильная начальная популяция)	(0;1)	(0;1)	(0;1)

Т.е. если $C(X', X'') = 1$, то все решения из множества X'' доминируются решениями из множества X' .

Сравним попарно алгоритмы на основе средней *C*-метрики за 10 запусков. В скобках в ячейках таблицы указаны $C(X', X'')$ и $C(X'', X')$, где X' – множество решений алгоритма, указанного первым в соответствующем столбце таблицы, а X'' – указанного вторым.

Рассмотрев *табл. 3*, мы видим, что генетические алгоритмы демонстрируют схожие результаты для 100 и 250 проектов с небольшим преимуществом алгоритма, основанного на сильной начальной популяции. Однако при 500 проектах он почти полностью доминирует алгоритмы с начальной популяцией, заданной полностью случайно. Также он превосходит алгоритм муравьиной оптимизации. Таким образом, для задачи формирования портфеля проектов оптимальным является генетический алгоритм с сильной начальной популяцией.

Заключение

В статье описана задача формирования портфеля проектов и обоснована ее текущая актуальность. Представлена многокритериальная математическая постановка задачи. Также описано расширение модели с помощью нечетких множеств, что позволяет учесть неточность ранних оценок проектов. Мы провели обзор существующих методов решения задач многокритериальной оптимизации, и в результате выбрали метод муравьиной оптимизации и генетический алгоритм для реализации, т.к. они обладают лучшей временной зависимостью. Реализация генетического алгоритма основана на методе SPEA II и мы представляем три ее модификации, различающиеся по варианту механизма селекции и по генерации начальной популяции. Проведенные эксперименты показывают, что время выполнения алгоритма муравьиной оптимизации сильно возрастает с увеличением числа проектов, что делает его значительно медленнее при числе проектов больше 250 по сравнению с модификациями генетического алгоритма. Наиболее эффективным алгоритмом для решения задачи формирования портфеля проектов является генетический алгоритм, основанный на сильной начальной популяции, который превосходит остальные алгоритмы. ■

Литература

1. Управление проектами в соответствии с ISO 21500 // Международная организация по стандартизации. [Электронный ресурс]: <http://iso21500.ru/>. (дата обращения: 15.07.13)
2. Матвеев А.А., Новиков Д.А., Цветков А.В. Модели и методы управления портфелями проектов. М.: ПМСОФТ, 2005. 206 с.
3. Bastiani S.S., Cruz L., Fernandez E., Gómez C., Ruiz V. Project Ranking-Based Portfolio Selection Using Evolutionary Multiobjective Optimization of a Vector Proxy Impact Measure // Proceedings of the Eureka Fourth International Workshop, 2013. Mazatlan, Mexico, November 6-8, 2013.
4. Yu L., Wang S., Wen F., Lai K.K. Genetic Algorithm-Based Multi-Criteria Project Portfolio Selection // Annals of Operations Research. 2012. 197(1). P. 71-86.
5. Аньшин В.М., Демкин И.В., Царьков И.Н., Никонов И.М. Применение теории нечетких множеств к задаче формирования портфеля проектов // Проблемы анализа риска. 2008. Сер. 3. № 5. С. 8-21.
6. Zitzler E., Laumanns M., Thiele L. An Efficient, Adaptive Parameter Variation Scheme for Metaheuristics Based on the Epsilon-Constraint Method // European Journal of Operational Research. 2006. 169(3). P. 932-942.
7. Alaya I., Solnon C., Ghedira K. Ant Colony Optimization for Multi-objective Optimization Problems // ICTAI '07. Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence. Patras, 2007. Vol. 01. P. 450-457.
8. Zitzler E., Laumanns M., Thiele L. SPEA2: Improving the Strength Pareto Evolutionary Algorithm: Technical Report 103 / Computer Engineering and Communication Networks Lab (TIK). Swiss Federal Institute of Technology (ETH) Zurich. 2001.
9. Akca Y., Li H., Xu S. Greedy Algorithm for the General Multidimensional Knapsack Problem // Annals of Operations Research. 2007. 150 (1). P. 17-29.

PROJECT PORTFOLIO FORMATION BASED ON FUZZY MULTI-OBJECTIVE MODEL

Sergey AVDOSHHIN,

Professor, Head of School of Software Engineering, Faculty of Business Informatics,
National Research University Higher School of Economics

Alexey LIFSHITS,

MSc Program Student, School of Software Engineering, Faculty of Business Informatics,
National Research University Higher School of Economics

Address: 20, Myasnitskaya str., Moscow, 101000, Russian Federation

E-mail: savdoshin@hse.ru, alexeus1992@yandex.ru

Leading IT companies run simultaneously several dozens or even several hundreds of projects. One of the major objectives is to decide whether a project meets the current strategic goals and resource limits of a company or not. This leads firms to the issue of a project portfolio formation, where the challenge is to choose a subset of projects which meet the strategic objectives of a company in the best way. In this present article we propose a multi-objective mathematical model of the project portfolio formation problem, defined on the fuzzy trapezoidal numbers.

We provide an overview of methods for solving this problem, which are a Branch and bound approach, an adaptive parameter variation scheme based on the epsilon-constraint method, ant colony optimization method and genetic algorithm. After our analysis, we choose the ant colony optimization method and SPEA II method, which is a modification of genetic algorithm. We describe the implementation of these methods applied to the project portfolio formation problem.

The ant colony optimization is based on the max min ant system with one pheromone structure and one ant colony. Three modifications of our SPEA II implementation have been considered. The first adaptation uses the binary tournament selection, while the second requires the rank selection method. The last one is based on another variant of generating initial population. Part of the population is generated by a non-random manner on the basis of solving a one-criterion optimization problem. This fact makes the population stronger than the initial one which is generated completely at random.

We compare the ant colony optimization algorithm and the three modifications of a genetic algorithm on the basis of the following parameters: speed of execution and the C-metric between each pair of algorithms. Genetic algorithm with non-random initial population show better results than other methods. Thus, we propose using this algorithm for solving project portfolio formation problem.

Key words: project portfolio, multi-objective model, fuzzy numbers, genetic algorithm, ant colony optimization.

References

1. International Standardization Organization. *Upravlenie proektami v sootvetstvii s ISO 21500* [Project Management in Accordance with ISO 21500]. Available at: <http://iso21500.ru/> (accessed 15 July 2013). (in Russian)
2. Matveev A., Novikov D., Tsvetkov A. (2005) *Modeli i metody upravleniya portfeljami proektov* [Models and Methods of Project Portfolio Management]. Moscow: PMSOFT. (in Russian)
3. Bastiani S., Cruz L., Fernandez E., Gómez C., Ruiz V. Project Ranking-Based Portfolio Selection Using Evolutionary Multiobjective Optimization of a Vector Proxy Impact Measure. Proceedings of the *Eureka Fourth International Workshop, 2013 (Mazatlan, Mexico, November 6-8, 2013)*.
4. Yu L., Wang S., Wen F., Lai K. (2012) Genetic Algorithm-Based Multi-Criteria Project Portfolio Selection. *Annals of Operations Research*, vol. 197, issue 1, pp. 71-86.
5. Anshin V., Dyomkin I., Tsarkov I., Nikonov I. (2008) Primenenie teorii nechetkih mnozhestv k zadache formirovaniya portfelja proektov [On Application of Fuzzy Set Theory to the Problem of Project Portfolio Selection]. *Issues of Risk Analysis*, vol. 3, no. 5, pp 8-21. (in Russian)
6. Zitzler E., Laumanns M., Thiele L. (2006) An Efficient, Adaptive Parameter Variation Scheme for Metaheuristics Based on the Epsilon-Constraint Method. *European Journal of Operational Research*, vol. 169, no 3, pp. 932-942.
7. Alaya I., Solnon C., Ghedira K. (2007) Ant Colony Optimization for Multi-objective Optimization Problems. Proceedings of the *19th IEEE International Conference on Tools with Artificial Intelligence, 2007 (Patras, Greece, October 29-31, 2007)*, vol. 1, pp. 450-457.
8. Zitzler E., Laumanns M., Thiele L. (2001) *SPEA2: Improving the Strength Pareto Evolutionary Algorithm. Technical Report 103*. Computer Engineering and Communication Networks Lab (TIK). Swiss Federal Institute of Technology (ETH) Zurich. Zurich: ETH.
9. Akcay Y., Li H., Xu S. (2007) Greedy Algorithm for the General Multidimensional Knapsack Problem. *Annals of Operations Research*, vol. 150, no. 1, pp.17-29.

ПРИМЕНЕНИЕ МНОГОАГЕНТНОГО ГЕНЕТИЧЕСКОГО АЛГОРИТМА ДЛЯ ПОИСКА ОПТИМАЛЬНЫХ СТРАТЕГИЧЕСКИХ И ОПЕРАТИВНЫХ РЕШЕНИЙ

М.А. Хивинцев,
аспирант кафедры бизнес-аналитики, факультет бизнес-информатики,
Национальный исследовательский университет «Высшая школа экономики»

А.С. Акопов,
доктор технических наук, профессор кафедры бизнес-аналитики, факультет
бизнес-информатики, Национальный исследовательский университет
«Высшая школа экономики»

Адрес: 101000, Москва, Мясницкая ул., 20
E-mail: mkhivintsev@hse.ru, aakopov@hse.ru

В статье представлен новый подход к применению многоагентного генетического алгоритма (MAGAMO) для поиска оптимальных стратегических и оперативных решений в имитационных моделях большой размерности.

Цель работы – разработка с использованием методов системной динамики имитационной модели типового Интернет-магазина и применение многоагентного генетического алгоритма MAGAMO для решения многокритериальной оптимизационной задачи стратегического и оперативного управления, относящейся к классу задач сверхбольшой размерности. Для реализации математической модели типового Интернет-магазина используется система имитационного моделирования Powersim Studio.

Объектом исследования являются многокритериальные оптимизационные задачи большой размерности, реализуемые в системах имитационного моделирования.

Для решения подобных задач предложен многоагентный генетический алгоритм MAGAMO. Особенностью данного алгоритма является распределение набора управляющих параметров системы между агентами на основе предварительного кластерного анализа. Каждый агент представляет собой независимый генетический алгоритм с собственной эволюцией решений, соответствующих заданным управляющим параметрам. Информационный обмен между агентами, функционирующими в параллельных процессах, осуществляется через разделяемую память системы (многомерную базу данных). При этом центральный процесс отвечает за отбор решений наивысшего ранга Парето. С использованием специального программного средства ParetoFrontViewer обеспечивается визуализация фронт Парето.

Разработанная имитационная модель интегрирована с алгоритмом MAGAMO, системой визуализации границы Парето и многомерной базой данных.

В результате проведенных численных экспериментов, осуществленных на реальных данных Интернет-магазина, продемонстрирована высокая эффективность разработанного многоагентного генетического алгоритма для поиска оптимальных решений в системах имитационного моделирования большой размерности.

Ключевые слова: многокритериальная оптимизация, системная динамика, параллельные генетические алгоритмы, проблема большой размерности, граница Парето.

Введение

Эффективное стратегическое и оперативное управление сложными организационными структурами обуславливает необходимость решения многокритериальных оптимизационных задач большой размерности. При этом для проектирования интеллектуальных систем управления подобными системами, как правило, используются методы и инструменты имитационного моделирования.

Так, с помощью методов системной динамики [1] – [4], можно описать важнейшие взаимосвязи между финансовыми и материальными потоками крупной компании и исследовать динамику ключевых показателей деятельности (KPI) в зависимости от сценарных условий. Следующим этапом является интеграция имитационной модели предприятия с генетическими оптимизационными алгоритмами для поддержки механизма оптимального управления. Так, например, в работах [5] – [7] описана реализация генетического алгоритма с угасающей селекцией, интегрированного с разработанной имитационной моделью крупной нефтяной компании (ВИНК). В результате решается стратегическая задача максимизации акционерной стоимости ВИНК при различных ограничениях.

Применение генетических алгоритмов в системах имитационного моделирования оправдано невозможностью использования точных аналитических методов градиентного типа. Значение целевой функции является результатом «прогона» имитационной модели.

Однако, с ростом размерности задачи и количества исходных данных эффективность известных генетических алгоритмов (ГА) существенно падает, так как для работы ГА требуется многократный пересчет фитнес-функции, основанный на соответствующих «прогонах» имитационной модели предприятия. При этом, время «прогона» имитационной модели сверхбольшой размерности даже с использованием суперкомпьютерных технологий, как правило, велико (исчисляется минутами). Поэтому для подобных систем наблюдается экспоненциальный рост вычислительной сложности с увеличением числа размерностей оптимизационной задачи.

Поэтому был разработан многоагентный генетический алгоритм для многоцелевой оптимизации (MAGAMO), описанный в работе [8]. Данный алгоритм использует в своей основе принципы

известного генетического алгоритма SPEA2 [9], параллельные вычисления для оценки значений фитнес-функций и приспособленности популяции [10] и др. Вместе с тем, имеется существенное отличие MAGAMO от известных параллельных ГА, в том числе от так называемой «островной» модели ГА [11], заключающееся в том, что в MAGAMO имеются интеллектуальные агенты, представляющие собой независимые ГА (исполнимые на отдельных вычислительных кластерах или в потоках), между которыми осуществляется распределение пространства искомых переменных. При этом центральный процесс отвечает за отбор решений наивысшего ранга Парето и формирование границы Парето.

В настоящей статье рассмотрен пример применения MAGAMO для решения многокритериальной оптимизационной задачи в разработанной имитационной модели типового Интернет-магазина, относящейся к классу задач сверхбольшой размерности. Для реализации математической модели типового Интернет-магазина используется система имитационного моделирования PowersimStudio. Для формирования множества Парето-оптимальных решений применяется генетический алгоритм MAGAMO. Для визуализации границы Парето используется основанный на методах аппроксимации программный продукт ParetoFrontViewer, разработанный в ВЦ РАН [12].

1. Многокритериальная оптимизационная задача типового Интернет-магазина

С использованием методов системной динамики разработана типовая (референтная) имитационная модель для крупного Интернет-магазина.

Рассматриваемая условная компания осуществляет дистанционную торговлю несколькими товарными категориями в различных регионах (городах). Следует отметить, что компания является дистрибьютором готовой продукции, т.е. в имитационной модели предприятия нет производственной составляющей.

Разработанная модель предназначена для поддержки принятия решений по динамическому управлению ассортиментом, ценообразованием, качеством обслуживания, оборачиваемостью запасов, маркетинговой активностью, расходами, географическим покрытием (присутствием в регионах

и т.д.). При этом поиск оптимальных управленческих решений должен осуществляться в разрезе регионов, товарных категорий, сегментов потребительских предпочтений и др.

Следует отметить, что в модели выделяются следующие *оперативные управляющие параметры*, оптимальные значения которых формируются весьма часто (в частности, еженедельно):

- ◆ уровень цен на товарные категории;
- ◆ комиссия с 1 заказа за доставку по регионам;
- ◆ интенсивность маркетинговой активности.

В модели имеются *стратегические управляющие параметры*, оптимальные значения которых формируются довольно редко (в частности, по годам):

- ◇ коэффициент маркетинговой активности по регионам;
- ◇ коэффициент маркетинговой активности по категориям товаров;
- ◇ коэффициент маркетинговой активности по сегментам потребительских предпочтений;
- ◇ качество обработки заказа (обслуживания);
- ◇ доступность товаров на складе по товарным категориям.

Основная задача торгового предприятия заключается в поиске оптимальных значений стратегических и оперативных решений, обеспечивающих сбалансированное динамическое развитие компании. В частности, выбраны следующие целевые показатели:

- накопленная EBITDA (прибыль до вычета расходов по уплате налогов, процентов по кредиту и начисленной амортизации);
- размер активной клиентской базы (количество клиентов, совершивших хотя бы одну покупку за предшествующие 12 месяцев);
- средняя оборачиваемость товарных запасов.

Для решения оптимизационной задачи требуется максимизировать на конец периода моделирования накопленную *EBITDA* и *размер активной клиентской базы*, минимизировав при этом среднюю за периоды моделирования *оборотчиваемость товарных запасов*. Эти цели являются стратегически важными для типового Интернет-магазина, при этом они являются конкурентными (например, рост клиентской базы за счет снижения цен приводит к уменьшению прибыли). При этом важность каждой из целей не поддается весовой оценке, так как имеется

актуальная задача максимально возможного приближения к достижению каждой из них. Поэтому необходимо нахождение Парето-оптимальных решений и построение Парето-фронта (т.е. искомые управленческие решения *должны лежать на границе множества Парето*).

Следует отметить, что в каждый момент быстрого времени должны выполняться стратегические ограничения, имеющие понятный экономический смысл:

- при определении ценовой политики средняя маржинальность продаж каждой товарной категории в каждом периоде должна быть в диапазоне от заданного минимального уровня маржи до максимально допустимого;

- стоимость доставки в каждом периоде должна составлять от 0 до максимального заданного уровня в процентах от стоимости товара;

- качество обслуживания должно быть более 0 и меньше либо равно максимально допустимому уровню;

- доступность товаров на складе по каждой категории должно быть от минимально допустимого уровня до 100%;

- интенсивность маркетинговой активности в каждом периоде должна быть от 0 до максимально допустимого уровня;

- сумма коэффициентов маркетинговой активности по регионам, категориям, сегментам потребительских предпочтений равна 1;

- оборачиваемость запасов в каждом периоде не должна превышать заданный максимальный уровень в днях;

- доля рынка в количественном выражении в каждом периоде на каждом региональном уровне должна быть не ниже минимального уровня;

- доля рынка в количественном выражении в каждом периоде по каждой категории должна быть не ниже минимального уровня.

Разработка модели осуществлялась с использованием методов системной динамики. Была выполнена исследовательская работа по объекту моделирования – проведен маркетинговый анализ рынка и выявлены различные зависимости, что позволило основать прогноз продаж на коэффициенте трансформации активной клиентской базы (готовность лояльных клиентов к совершению повторной покупки) и суммарной емкости рынка, а также выделить влияние различных

факторов на вероятность покупки в конкретном Интернет-магазине для каждого из трех сегментов потребительских предпочтений. К первому сегменту относятся экономные клиенты, которым важна цена и они реагируют на рекламу. Во второй сегмент вошли клиенты, активно реагирующие на рекламу – ее имиджевую составляющую – и повышенное требование они, как правило, предъявляют к качеству (надежности). В третий сегмент попали наиболее рациональные клиенты, которые ищут выгодную цену и хорошее качество (надежность), стараясь игнорировать рекламу.

Факторами, влияющими на вероятность покупки, являются: общий размер клиентской базы, потребительская активность периода (сезонность), сила конкуренции, разница между установленной итоговой ценой товара (с учетом комиссии за доставку) и среднерыночной ценой, лояльность к другим Интернет-магазинам (фактор естественной конкуренции), накопленный имидж магазина, накопленное маркетинговое покрытие, доступность товаров на складе. Для каждого из сегментов потребительских предпочтений эти факторы имеют разную силу воздействия на вероятность покупки. Перечисленные факторы входят в экспоненциальное уравнение логистической регрессии, при помощи которого вычисляется вероятность совершения покупки различными группами клиентов.

Кроме того, данные факторы имеют различное воздействие на клиентов, совершавших покупку в данном магазине и не совершавших. Поэтому прогноз спроса строится отдельно для повторных клиентов (имевших хотя бы одну покупку за последний год, составляющих активную клиентскую базу) и новых клиентов.

Для повторных клиентов прогноз спроса строится на основе коэффициентов трансформации клиентской базы, предварительно разбитой по периодам совершения клиентами своей последней покупки, в число намеревающихся совершить в данном периоде очередную покупку. На окончательное их решение о покупке влияют перечисленные выше факторы, но с другими коэффициентами, отличными от применяемых для новых клиентов, ввиду уже однажды проявленной лояльности к магазину. К тому же, добавляется в расчет фактор качества обслуживания при их последней покупке.

Для новых клиентов прогноз спроса основан на определении максимального потенциала роста клиентской базы (емкость рынка за вычетом имеющихся человек в клиентской базе) в каждом периоде и доли рынка, которую удастся получить благодаря определяющим вероятность покупки факторам.

В модели присутствуют нелинейные зависимости, зависимости переменных от их значения в прошлом периоде, стохастические функции, присваивание значений переменным массива в циклах. Так, например, прогнозирование изменений накопленного маркетингового покрытия и имиджа осуществлено через подход дисконтирования, согласно которому значение прошлого периода учитывается с определенным дисконтом в новом периоде. Перечисленные особенности модели делают невозможным представление целевых функций в аналитическом виде, что приводит к неприменимости классических методов (основанных на вычислении производных функции) для решения данной оптимизационной задачи.

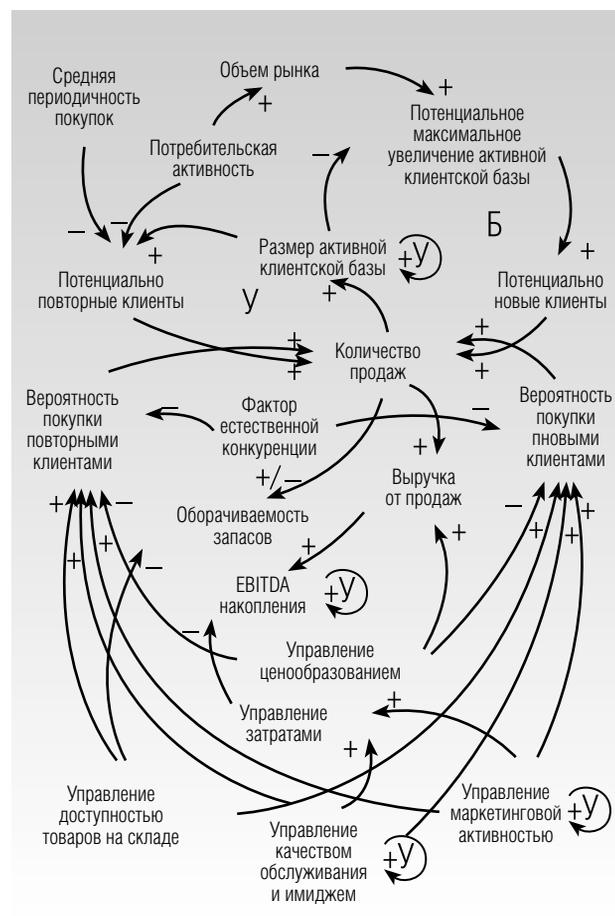


Рис. 1. Диаграмма причинно-следственных связей

В качестве горизонта моделирования рассматривается один год, а шаг моделирования равен 1 неделе (итого 52 недели, 52 шага). Неделя является наилучшим периодом для принятия и оценки оперативных управленческих решений при управлении Интернет-магазином. Горизонт моделирования может быть продлен на несколько лет.

На *рис. 1* представлена диаграмма, демонстрирующая причинно-следственные связи модели.

Введем следующие обозначения:

Индексы, используемые в модели:

$t = t_0, t_0 + 1, \dots, T$ – индекс периода (быстрое время: 1 период = 1 неделя), T – горизонт планирования (52 недели \approx год);

$i = 1, 2, \dots, I$ – индекс регионов (городов);

$j = 1, 2, \dots, J$ – индекс товарных категорий;

$k = 1, 2, \dots, K$ – индекс сегментов потребительских предпочтений;

$w = 1, 2, \dots, W$ – индекс периода последней покупки (номера недели).

Экзогенные переменные:

$s_j(t)$ – сезонная потребительская активность данного периода по j -ым товарным категориям (%);

$v_{i,k,j}$ – исходный объем рынка по i -ым регионам, k -ым сегментам потребительских предпочтений, j -ым товарным категориям (шт);

$cb_{i,k,w}$ – исходная клиентская база, сегментированная по i -ым регионам, k -ым сегментам потребительских предпочтений, w -ым периодам последней покупки (чел);

$r_w(t)$ – степень готовности повторных клиентов, разделенных по периодам последней покупки, к следующей покупке (%);

av – среднее число покупок клиентами за период;

$\varepsilon_{k,j}$ – базовая эластичность по цене по k -ым сегментам потребительских предпочтений и j -ым категориям;

l_i – длительность исполнения заказа по i -ым регионам (дней);

σ_j – длительность оборачиваемости товарных запасов по j -ым категориям (дней);

ab_j – уровень базовой доступности товаров на складе по категориям (%);

b – коэффициент нелинейного роста длительности оборачиваемости склада при повышении доступности товаров на складе.

β_j – доля возвратов после продажи по j -ым категориям товаров (%);

u – доля неликвидного в возвратном потоке товаров (%);

cp_j – исходная средняя себестоимость товаров по j -ым категориям (руб);

DC_i – расходы на доставку 1 кг в i -ый регион (руб/кг);

Wt_j – средний вес товара в j -ой категории (кг);

mb – базовый маркетинговый бюджет (руб);

oc_0, oc_1, oc_2 – коэффициенты, используемые для определения функции операционных расходов на обработку 1 заказа (oc_0, oc_1 – руб);

mc_0, mc_1, mc_2 – коэффициенты, используемые для определения функции управленческих расходов (mc_0, mc_1 – руб);

c – уровень комиссии за прием платежей (%).

Стратегические управляющие параметры модели:

q – уровень качества обработки заказов;

a_j – доступность товаров на складе по товарным категориям (%);

$m1_i$ – маркетинговая активность по i -ым регионам (%);

$m2_j$ – маркетинговая активность по j -ым товарным категориям (%);

$m3_k$ – маркетинговая активность по k -ым сегментам потребительских предпочтений (%).

Оперативные управляющие параметры модели:

$p_j(t)$ – средние цены на j -ые товарные категории в период времени t (руб);

$d_i(t)$ – соотношение стоимости доставки к стоимости товаров в i -ых регионах в период времени t (%);

$m(t)$ – коэффициент интенсивности маркетинговой активности в период времени t .

Динамика основных показателей модели

(эндогенных переменных) в момент времени t .

Суммарный объем рынка:

$$V_{i,k,j}(t) = v_{i,k,j} \cdot ss_j(t), \quad (1)$$

где $ss_j(t)$ – это накопленная потребительская активность:

$$ss_j(t) = \begin{cases} 1, & t = t_0 \\ ss_j(t-1) \cdot (1 + s_j(t)), & t > t_0. \end{cases} \quad (2)$$

Число клиентов в активной клиентской базе:

$$CB_{i,k,w}(t) = \begin{cases} cb_{i,k,w}, & t = t_0 \\ \frac{Q_{ij}(t)}{av}, & t > t_0 \text{ и } w = 1, \\ CB_{i,k,w}(t-1), & t > t_0 \text{ и } w > 1 \end{cases} \quad (3)$$

где $Q_{i,j}(t)$ обозначает динамику продаж в штуках и будет определено ниже.

Максимальный потенциал продаж повторным (лояльным) клиентам:

$$N_{i,k,w,j}^{(нов)}(t) = CB_{i,k,w} \cdot r_w(t) \cdot ss_j(t) \cdot av. \quad (4)$$

Максимальный потенциал продаж новым клиентам:

$$N_{i,k,j}^{(нов)}(t) = V_{i,k,j}(t) - \sum_{w=1}^W N_{i,k,w,j}^{(нов)}(t). \quad (5)$$

Вероятность покупки повторными клиентами:

$$Pb_{i,k,w,j}^{(нов)}(t) = \frac{1}{1 + e^{-z_{i,k,w,j}^{(нов)}(t)}}, \quad (6)$$

где $z_{i,k,w,j}^{(нов)}(t)$ определяется суммой факторов, оказывающих влияние на спрос для повторных клиентов (конкуренции, цены, доступности товаров, имиджа, маркетинга, качества оказанного обслуживания в прошлый раз).

Вероятность покупки новыми клиентами:

$$Pb_{i,k,j}^{(нов)}(t) = \frac{1}{1 + e^{-z_{i,k,j}^{(нов)}(t)}} \quad (7)$$

где $z_{i,k,w,j}^{(нов)}(t)$ определяется суммой факторов, оказывающих влияние на спрос для новых клиентов (конкуренции, цены, доступности товаров, имиджа, маркетинга).

Так, например, фактор влияния цены вычисляется через отклонение установленной итоговой цены на товар ($p_j(t)$ с учетом цены за доставку $d_j(t)$) от среднерыночной с использованием подхода, основанном на эластичности. При этом коэффициент эластичности ($\varepsilon_{k,j}$), в свою очередь, имеет степенную зависимость от отклонения от среднерыночной цены. А динамика среднерыночной цены определяется инфляцией ($inf_j(t)$) «прогнозом уровня конкуренции» и случайным отклонением, заданным через нормальное распределение.

Динамика объема продаж в штуках:

$$Q_{ij}(t) = \sum_{k=1}^K (N_{i,k,j}^{(нов)}(t) \cdot Pb_{i,k,j}^{(нов)}(t)) + \sum_{k=1}^K \sum_{w=1}^W (N_{i,k,w,j}^{(нов)}(t) \cdot Pb_{i,k,w,j}^{(нов)}(t)). \quad (8)$$

Динамика объема продаж в деньгах:

$$G(t) = \sum_{i=1}^I \sum_{j=1}^J (Q_{i,j}(t) \cdot P_{i,j}(t)). \quad (9)$$

Динамика себестоимости в денежном выражении:

$$CP_j(t) = \left(\sum_{i=1}^I Q_{i,j}(t) \right) \cdot cp_j \cdot PI_j(t). \quad (10)$$

где $PI_j(t)$ – это накопленный индекс цен по товарным категориям, который меняется каждый период под воздействием инфляции, конкуренции и сезонного регулирования среднерыночных цен.

Динамика общих расходов:

$$E(t) = \sum_{i=1}^I \sum_{j=1}^J \{ CP_j(t) \cdot (1 + \beta_j \cdot u) + Q_{i,j}(t) \cdot (OC + DC_i \cdot Wt_i) \} + G(t) \cdot c + MB(t) + MC(t) \quad (11)$$

где $OC = oc_0 + oc_1 \cdot e^{oc_2 \cdot q}$ – норма операционных затрат на 1 заказ;

$MB(t) = mb \cdot m(t)$ – маркетинговый бюджет данного периода;

$MC(t) = \text{MAX}(mc_0 + mc_1 \cdot \ln((Q_{i,j}(t))^{mc_2}); MC(t-1) \cdot 0,9)$ – зависимость управленческих расходов от объема продаж в штуках, причем они не могут более чем на 10% ниже, чем в прошлом периоде.

Средневзвешенный срок исполнения заказов:

$$L1(t) = \frac{\sum_{i=1}^I \sum_{j=1}^J \{ CP_j(t) \cdot l_i / (1 - \beta_j) \}}{\sum_{j=1}^J CP_j(t)}. \quad (12)$$

Средняя оборачиваемость складских запасов:

$$L2(t) = \frac{\sum_{j=1}^J \{ CP_j(t) \cdot \sigma_j \cdot (a_j / ab_j)^b \}}{\sum_{j=1}^J CP_j(t)}. \quad (13)$$

Накопленная прибыль (EBITDA) рассчитывается по формуле:

$$EBITDA = \sum_{t=t_0}^T (G(t) - E(t)). \quad (14)$$

Размер активной клиентской базы рассчитывается по формуле:

$$CBSUM(t) = \sum_{i=1}^I \sum_{k=1}^K \sum_{w=1}^W CB_{i,k,w}(t). \quad (15)$$

Динамика средней оборачиваемости запасов в течение периода моделирования рассчитывается по формуле:

$$\bar{L} = \left\{ \sum_{t=t_0}^T (L1(t) + L2(t)) \right\} / T. \quad (16)$$

Далее можно сформулировать задачу поиска оптимального стратегического и оперативного управления для типового Интернет-магазина.

Задача 1.

Необходимо вычислить оптимальные значения набора оперативных и стратегических управляющих параметров $\{q, a, m1_i, m2_j, m3_k, p_j(t), d_i(t), m(t)\}$, обеспечивающих максимальные значения прибыли и размера

клиентской базы при минимальном времени оборачиваемости запасов:

$$\begin{cases} \max_{\{q, a, m_1, m_2, m_3, p_i(t), d_i(t), m(t)\}} \{EBITDA\} \\ \max_{\{q, a, m_1, m_2, m_3, p_i(t), d_i(t), m(t)\}} \{CB(t)\} \\ \min_{\{q, a, m_1, m_2, m_3, p_i(t), d_i(t), m(t)\}} \{\bar{L}\} \end{cases} \quad (17)$$

при выполнении следующих ограничений в каждый момент времени t :

ограничение на уровень маржинальности

$$M_j^{min} < \frac{p_j(t)}{cp_j \cdot PI_j(t)} - 1 \leq M_j^{max}, \quad (18)$$

ограничения на долю рынка по городам

$$MS1_j^{min} \leq \sum_{i=1}^I Q_{i,j}(t) / \sum_{i=1}^I \sum_{k=1}^K V_{i,k,j}(t), \quad (19)$$

ограничения на долю рынка по товарным категориям

$$MS2_i^{min} \leq \frac{\sum_{j=1}^J Q_{i,j}(t)}{\sum_{k=1}^K \sum_{j=1}^J V_{i,k,j}(t)}, \quad (20)$$

ограничения на скорость оборачиваемости и пр.

$$L1(t) + L2(t) \leq L^{max}. \quad (21)$$

В рассматриваемой модели, содержащей 5 товарных категорий, 6 городов, 3 клиентских сегмента, 52 недели, 3 целевых функции, с учетом ограничений, размерность решений достигает примерно 380000, что характеризует размерность задачи как сверхбольшую.

2. Многоагентный генетический алгоритм MAGAMO

Для эффективного решения многокритериальной оптимизационной задачи сверхбольшой размерности был разработан многоагентный генетический алгоритм MAGAMO [8]. В его основе лежит динамическое взаимодействие синхронизированных интеллектуальных агентов, каждый из которых выполняет свой эволюционный генетический алгоритм по методу SPEA2. При этом каждый агент имеет собственную фитнес-функцию, содержащую свой набор переменных. Следовательно, каждый агент занят поиском собственных субоптимальных решений. Как только агент находит субоптимальное решение, он обменивается лучшим генетическим материалом с другими агентами. В то же время, агенты активны и они могут инициировать обмен в случае необходимости (когда цель улучшена). Это способствует улучшению разделенных эволюционных процессов соединенных агентов. В результате во всей системе происходит нахождение глобальных Парето-оптимальных решений.

Преимущества MAGAMO в задачах сверхбольшой размерности достигается благодаря снижению количества расчетов фитнес-функций, что, в свою очередь, происходит за счет снижения размерности пространства решений для каждого из агентов и, как следствие, снижения требований к количеству особей и расчетов фитнес-функций в результате их скрещивания.

Входные параметры алгоритма:

\bar{N} – набор решений

\bar{N}_j – архивный размер j -агента ($j = 1, 2, \dots, J$).

$P_{q,j}$ – популяция j -агента на q -итерации.

Q – максимальное число итераций эволюции агентов

$\bar{P}_{q,j}$ – архивы агентов с недоминируемыми особями на q -итерации ($P_{0,j}$)

Выходные значения алгоритма:

A – набор недоминируемых решений.

Функционал алгоритма MAGAMO может быть описан в следующем виде:

Шаг 1 – Статистический анализ задачи: проведение исследования целевых функций и кластеризация всех решений для формирования фитнес-функций для каждого агента. Распределение управляющих переменных между агентами.

Шаг 2 – Инициализация: инициализация набора решений и распределение их по n -размерностям между агентами. Создание исходной популяции для каждого j -агента и создание пустых архивов. Установка $q = 0$.

Шаг 3. Запуск основного процесса и выбор агентов: копирование всех недоминируемых решений, сформированных агентами и сохранённых в разделяемой памяти (базе данных) наилучших особей из $\bar{P}_{q,j}$ в $\bar{P}_{q+1,j}$. Если размер $\bar{P}_{q+1,j}$ достигает \bar{N}_j , тогда снижается посредством оператора усечения. Иначе, если размер $\bar{P}_{q+1,j}$ меньше, чем \bar{N}_j , то происходит заполнение $\bar{P}_{q+1,j}$ доминируемыми особями $P_{q,j}$.

Шаг 4. Запуск j -агентов в потоках: запуск модели в PowerSimStudio с набором лучших особей, который переходят между агентами для оценки целевых функций. Происходит движение вперед на шаг эволюции агентов для формирования лучших особей агентом $q_j = q_j + 1$. Эволюция агентов основана на операторах кроссовера и мутации. Периодически (раз в несколько итераций) осуществляется сохранение наилучших особей в разделяемую память системы (базу данных).

Шаг 5. Завершение: если $q_j \geq Q$ или достигаются другие критерии останова, тогда множество A – это набор решений в виде вектора, состоящий из недоминируемых решений в $\bar{P}_{q+1,j}$. Завершение работы алгоритма.



Рис. 2. Архитектура системы

Программное приложение для управления распределенной системой агентов согласно принципу MAGAMO было разработано в Microsoft Visual Studio 2012 на языке программирования C#. Оно также было интегрировано с использованием библиотеки SDK со средством имитационного моделирования PowerSim Studio, в котором была построена описанная референтная модель типового Интернет-магазина.

Архитектура разработанного программного комплекса представлена на рис. 2.

3. Построение и визуализация границы Парето

Для демонстрационного решения поставленной оптимизационной задачи на основе MAGAMO была создана распределенная система, состоящая из центрального процесса и 4 агентов. Исходное распределение пространства решений по агентам представлено на табл. 1.

Таблица 1.

Распределение пространства решений между агентами

	Пространство решений
Агент 1	$q, m1_{i/2}, m2_{j/2}, m3_{k/2}, m(t)$
Агент 2	$q, m1_{i/2}, m2_{j/2}, m3_{k/2}, m(t)$
Агент 3	$a_{j/2}, p_{j/2}(t), d_{i/2}(t)$
Агент 4	$a_{j/2}, p_{j/2}(t), d_{i/2}(t)$

В результате работы системы на приближенных к реальности тестовых данных были найдены 3 тысячи Парето-оптимальных решений за 2,5 часа. Каждое решение приводит к комбинации из значений 3 целевых функций. Эти комбинации были экспортированы в специальный программный продукт Pareto Front Viewer, разработанный в ВЦ РАН [12], что позволило графически изобразить фронт Парето (рис. 3).

Для визуализации границы Парето выбрано представление на плоскости с использованием цветовой шкалы для третьего показателя: ось Oх – накопленная EBITDA в тыс. руб., ось Oy – размер активной клиентской базы в тыс. чел., ось Oz – средняя оборачиваемость товарных запасов в днях. График представлен на рис. 3.

Диапазон светлого цвета содержит решения, которые позволяют достичь относительно низких показателей накопленной прибыли и размера клиентской базы при минимальной оборачиваемости запасов. Диапазон черного цвета, напротив, позволяет добиться одновременного увеличения первых двух целевых показателей и ухудшение третьего. Решение оптимизационной задачи и визуализация границы Парето позволяет выбрать решение, исходя из выбранной стратегии, которая, в качестве примера, может заключаться в балансировке всех целевых показателей (сценарий 1 на рис. 3) или в максимизации прибыли при минимально допустимом размере клиентской базы на уровне 1,28 млн. человек и при оборачиваемости запасов не более 32 дней (сценарий 2 на рис. 3).

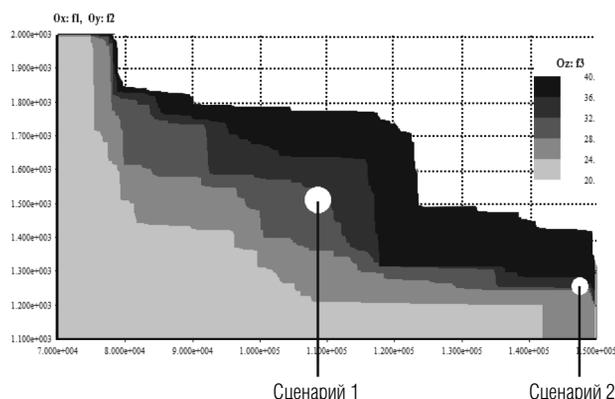


Рис. 3. Визуализация границы Парето в Pareto Front Viewer

Для демонстрационного примера анализ границы Парето позволяет также сделать вывод, что добиться максимальной клиентской базы можно при различном уровне оборачиваемости запасов, но прибыль по всем случаям будет предельно низкой.

Заключение

С использованием методов системной динамики разработана и описана имитационная модель типового Интернет-магазина. Выполнена реализация имитационной модели в системе имитационного моделирования PowerSim Studio. Сформулирована задача оптимального стратегического и оперативного управления, относящаяся к классу многокритериальных оптимизационных задач сверхбольшой размерности.

Осуществлен поиск Парето-оптимальных решений с использованием разработанного многоагентного генетического алгоритма (MAGAMO). Особенностью данного алгоритма является распределение набора управляющих параметров системы между агентами на основе предварительного кластерного анализа. Информационный обмен между

агентами, функционирующими в параллельных процессах, осуществляется через разделяемую память системы (многомерную базу данных). При этом центральный процесс отвечает за отбор решений наивысшего ранга Парето. С использованием специального программного средства Pareto Front Viewer визуализирован фронт Парето.

Разработан программный комплекс, представляющий собой информационно-аналитическую систему поддержки принятия стратегических и оперативных решений, в рамках которой, реализована интеграция системы имитационного моделирования PowersimStudio с многомерным информационным хранилищем и многоагентным генетическим алгоритмом (MAGAMO). Экспериментально показана высокая эффективность разработанной системы для решения многокритериальных оптимизационных задач большой размерности. ■

Литература

1. Forrester J.W. Industrial Dynamics. MIT Press, 1961.
2. Sterman J. Business Dynamics. Irwin McGraw-Hill, 2000. 982pp.
3. Горбунов А.Р. Управление финансовыми потоками. Проект «сборка холдинга». М.: Глобус, 2005. 224 с.
4. Исаев Д.В. Мониторинг и планирование развития систем информационной поддержки корпоративного управления и стратегического менеджмента // Бизнес-информатика. 2012. № 3 (21). С. 63-69.
5. 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.
6. Акопов А.С. К вопросу проектирования интеллектуальных систем управления сложными организационными структурами. Ч.2. Программная реализация системы управления инвестиционной деятельностью вертикально-интегрированной нефтяной компании // Проблемы управления. 2011. № 1. С. 47-54.
7. Акопов А.С. К вопросу проектирования интеллектуальных систем управления сложными организационными структурами. Ч.1. Математическое обеспечение системы управления инвестиционной деятельностью вертикально-интегрированной нефтяной компании // Проблемы управления. 2010. № 6. С. 12-18.
8. Akopov A.S., Hevencev M.A. A Multi-agent genetic algorithm for multi-objective optimization // Proceedings of 2013 IEEE International Conference on Systems, Man and Cybernetics. Manchester, UK, October 13-16, 2013. IEEE, 2013. P. 1391-1396.
9. 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). Seoul, Korea, May 27-30, 2001. P. 536-543.
10. Хивинцев М.А., Акопов А.С. Распределенная эволюционная сеть для решения многокритериальных оптимизационных задач в системах имитационного моделирования // Бизнес-информатика. 2013. № 3(25). С. 35-41.
11. Gordon V.S., Whitley D., Bohn A. Dataflow parallelism in genetic algorithms // Parallel Problem Solving from Nature, Amsterdam: Elsevier Science. 1992. No.2. P.533-542.
12. Pareto Front Viewer. <http://www.ccas.ru/mmes/mmeda/soft/third.htm> (date of access 01.01.2014).

APPLICATION OF MULTI-AGENT GENETIC ALGORITHM FOR SEARCH OF OPTIMUM STRATEGIC AND OPERATIONAL DECISIONS

Maxim KHIVINTCEV,

Post-Graduate Student, Department of Business Analytics, Faculty of Business Informatics,
National Research University Higher School of Economics

Andranik AKOPOV,

Professor, Department of Business Analytics, Faculty of Business Informatics,
National Research University Higher School of Economics

Address: 20, Myasnitskaya str., Moscow, 101000, Russian Federation

E-mail: mkhivintsev@hse.ru, aakopov@hse.ru

The article presents a new approach to applying a multi-agent genetic algorithm (MAGAMO) for search of optimum strategic and operational solutions in large-scale simulation models.

The purpose of the paper is to develop a simulation model of a referential Internet shop on the basis of the system dynamics methods and to apply the multi-agent genetic algorithm (MAGAMO) for solution a multi-criteria optimizing problem of the strategic and operational control parameters related to the class of large-scale problems. An imitation modeling system Powersim Studio is used for implementing the mathematical model of referential internet shop.

The research is focused on the large-scale multi-criteria optimizing problems run in simulation systems.

For the solution of such problems, a multi-agent genetic algorithm (MAGAMO) is offered. The feature of this algorithm is the distribution of a set of the system operating parameters

between agents on the basis of the preliminary cluster analysis. Each agent represents independent genetic algorithm with its own evolution of the decisions, corresponding to the preset control parameters. Information exchange between the agents functioning in parallel processes is carried out through divided memory of system (a multidimensional database). Here, the central process is responsible for selecting solutions of the highest rank of Pareto. Using a specialized software of Pareto Front Viewer visualization, Pareto's front is provided.

The developed simulation model is integrated with algorithm of MAGAMO, system of visualization of Pareto front and a multidimensional database.

The results of numerical experiments, which have been carried out on real data of the internet shop, have demonstrated high efficiency of the developed multi-agent genetic algorithm for search of optimum solutions in systems of imitating modeling of big dimension.

Key words: multicriteria optimization, system dynamics, parallel genetic algorithms, large-scale problem, Pareto front.

References

1. Forrester J.W. (1961) *Industrial Dynamics*. MIT Press.
2. Sterman J. (2000) *Business Dynamics*. Irwin McGraw-Hill.
3. Gorbunov A.R. (2005) *Upravlenie finansovymi potokami. Proekt «sorkaholdinga»* [Management of financial streams. «Holding assembly» project]. Moscow: Globus. (in Russian)
4. Isaev D.V. (2012) Monitoring i planirovanie razvitiya sistem informacionnoj podderzhki korporativnogo upravleniya i strategicheskogo menedzhmenta [Monitoring and planning of development of information support systems for corporate governance and strategic management]. *Business informatics*, no. 3 (21), pp. 63-69. (in Russian)
5. Akopov A.S. (2012) Designing of Integrated System-dynamics Models for an Oil Company. *International Journal of Computer Applications in Technology*, vol. 45, no. 4, pp. 220-230.
6. Akopov A.S. (2011) K voprosu proektirovaniya intellektual'nyh sistem upravleniya slozhnymi organizacionnymi strukturami. Ch.2. Programmnaya realizacija sistemy upravleniya investicionnoj dejatel'nost'ju vertikal'no-integrirovannoj neftjanoy kompanii [On the issue of developing of intelligent control systems of complex organizational structures. Part 2. Software support for control system of the vertically integrated oil company investment activities]. *Problemy Upravleniya*, no. 1, pp. 47-54. (in Russian)
7. Akopov A.S. (2010) K voprosu proektirovaniya intellektual'nyh sistem upravleniya slozhnymi organizacionnymi strukturami. Ch.1. Matematicheskoe obespechenie sistem ypravleniya investicionnoj dejatel'nost'ju vertikal'no-integrirovannoj neftjanoy kompanii [On the issue of developing of intelligent control systems of complex organizational structures. Part 1. Mathematical support for control system of the vertically Integrated oil company investment activities]. *Problemy Upravleniya*, no. 6, pp. 12-18. (in Russian)

8. Akopov A.S., Hevencev M.A. (2013) A Multi-agent Genetic Algorithm for Multi-objective Optimization. Proceedings of 2013 IEEE International Conference on Systems, Man and Cybernetics. Manchester, UK, October 13-16, 2013. IEEE, pp. 1391-1396.
9. Bleuer S., Brack M., Thiele L., Zitzler E. (2001) Multiobjective genetic programming: Reducing bloat by using SPEA 2. Proceedings of the 2001 Congress on Evolutionary Computation (CES-2001). Seoul, Korea, May 27-30, 2001, pp. 536-543.
10. Hivintsev M.A., Akopov A.S. (2013) Raspredeleonnaja jevoljucionnaja set' dlja reshenija mnogokriterial'nyh optimizacionnyh zadach v sistemah imitacionnogo modelirovanija [Distributed evolutionary network for the solving of multi-criteria optimization problems in simulation systems]. *Business Informatics*, no. 3(25), pp. 35-41. (in Russian)
11. Gordon V.S., Whitley D., Bohn A. (1992) Dataflow parallelism in genetic algorithms. *Parallel Problem Solving from Nature, Amsterdam: Elsevier Science*, no.2, pp. 533-542.
12. Pareto Front Viewer. <http://www.ccas.ru/mmes/mmeda/soft/third.htm> (date of access 01.01.2014).

ВНЕДРЕНИЕ РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ В МНОГОФИЛИАЛЬНОЙ ОРГАНИЗАЦИИ: ОПЫТ ФГУП «ПОЧТА РОССИИ»

А.Ф. Моргунов,

*кандидат технических наук, доцент кафедры корпоративных
информационных систем факультета бизнес-информатики,*

Национальный исследовательский университет «Высшая школа экономики»

Адрес: 101000, г. Москва, ул. Мясницкая, д. 20

E-mail: amorgunov@hse.ru

Внедрение информационной системы в крупной компании, имеющей широкую филиальную сеть, как правило, сопряжено с рядом трудностей организационного и технического характера. Проекты внедрения растягиваются во времени и не всегда приводят к желаемым результатам.

В статье рассматриваются проблемы внедрения информационной системы автоматизации технологических процессов и документооборота подписки на периодические печатные издания в подразделениях российского почтового оператора ФГУП «Почта России». Оценивается состояние автоматизации бизнес-процесса до начала внедрения, исследуются организационные и технические проблемы, возникающие на этапе ввода в эксплуатацию новой информационной системы. К числу таких проблем относятся: обеспечение непрерывности технологического процесса, невозможность начала работы с «пустой базы данных», проблемы работы с персоналом, унификация бизнес-процесса и большое количество объектов внедрения (более 42 000). В заключение анализируются преимущества, которые может получить почтовый оператор от внедрения единой информационной системы.

Материалы статьи могут быть интересны специалистам, занимающимся внедрением информационных систем.

Ключевые слова: информационная система, внедрение, унификация, автоматизация, подписка на периодические печатные издания, почта.

Введение

Внедрение информационной системы (ИС) в крупной компании сам по себе процесс не простой, порой весьма болезненный для организации. Когда речь идет о внедрении единого

информационного решения в крупной организации с большим количеством филиалов по стране, то это только усложняет задачу, растягивает ее по времени и накладывает другие ограничения на внедрение. В данной статье будем рассматривать организацию, в филиалах которой уже был частично автоматизиро-

ван бизнес-процесс, при этом в каждом филиале в свое время было разработано собственное программное решение, учитывающее некоторые местные особенности. Цель работы – выявить закономерности, проявляющиеся при внедрении новой информационной системы взамен существующей, а также исследовать проблемы, которые могут возникнуть на этапе замены информационной системы, для того, чтобы по возможности решить проблемы до начала проведения работ или, по меньшей мере, снизить риски неудачи в процессе проведения работ. В статье будет рассмотрено влияние выявленных закономерностей на примере проекта внедрения единой информационной системы, которая должна заменить локальные разработки и унифицировать один из бизнес-процессов в ФГУП «Почта России».

1. Состояние автоматизации бизнес-процесса

Российский почтовый оператор ФГУП «Почта России» – уникальная по своим масштабам организация, имеющая около 42 000 представительств по всей стране [1]. Структура почтового оператора представлена на рис. 1. В каждом областном (краевом) центре имеется управление федеральной почтовой связи (УФПС) – филиал ФГУП «Почта России» (всего 86 филиалов из них 82 УФПС). В состав каждого филиала входит от 3 до 21 почтамтов [1]. Услуги населению оказываются в отделениях почтовой связи (ОПС), подчиняющихся почтамтам. Кроме этого в структуру ФГУП «Почта России» входит 2900 газетных узлов (точек магистральной доставки периодики по стране).

Одна из услуг, оказываемых российским почтовым оператором, – прием подписки на периодические печатные издания. Несмотря на внешнюю простоту (пришел на почту, заполнил бланк, заплатил деньги, получишь издание по указанному адресу), технологически это достаточно сложный набор операций, в

которых задействовано большое количество персонала: сотрудники почты, издательств, логистических служб [2-4].

Автоматизация технологических операций и документооборота по подписке на периодические печатные издания в каждом УФПС в течение последних 25 лет проводилась самостоятельно. В результате в настоящий момент практически в каждом УФПС функционируют собственные программные решения, разработанные на различных платформах. В ходе проведенного исследования выяснилось, что практически в половине УФПС работают программы, разработанные под DOS.

Еще хуже положение дел с автоматизацией приема подписки в отделениях почтовой связи. В большом количестве ОПС прием подписки производится с использованием бумажных каталогов периодических печатных изданий, после чего информация на бумажных носителях передается на вышестоящий уровень для последующей обработки. Есть почтамты, где обработка информации о подписке производится вручную.

Следствием такой «лоскутной» автоматизации стало наличие целого ряда проблем. В головной организации (аппарате управления ФГУП «Почта России») отсутствует оперативная и, самое главное, достоверная информация о ходе проведения и результатах подписных кампаний. Причиной сложившейся ситуации стало следующее обстоятельство: в аппарате управления нет никакой информационной системы, позволяющей собирать и обрабатывать информацию о подписке с нижестоящего уровня, информация из подчиненных подразделений поступает в виде консолидированных отчетов. Результатом низкой информированности стала слабая оперативная управляемость бизнес-процессом (оперативное точечное вмешательство в ход подписной кампании). Аппарат управления не может опера-

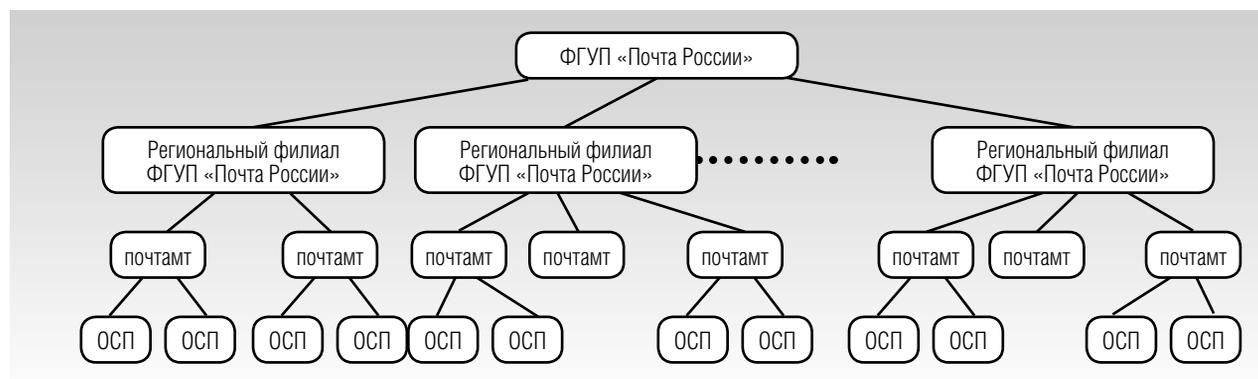


Рис. 1. Структура ФГУП «Почта России» (ОПС – отделение почтовой связи)

тивно влиять на сроки проведения, проводить более гибкую ценовую политику, организовывать акции по привлечению новых клиентов и удержанию старых. Быстрое внедрение новых технологических решений во всех филиалах не представляется возможным, что влечет за собой сложность перевода услуги на новый технологический уровень (распространение электронных версий изданий, оказание услуги с использованием иных, отличных от традиционного, способов, таких как Интернет, платежные терминалы и т.п.). Наличие большого количества локальных разработок приводит к большим затратам на сопровождение программного обеспечения.

Выход из сложившейся ситуации — замена локальных информационных систем на единое решение.

2. Некоторые закономерности, проявляющиеся при замене одной информационной системы на другую

При замене одной информационной системы организации на другую проявляется ряд закономерностей. Рассмотрим их более подробно.

Первое, на что следует обратить внимание, — это наличие у персонала опыта работы с информационными системами (по меньшей мере, хотя бы с одной из систем). Следовательно, сотрудников не требуется специально обучать принципам работы с информационными системами и больше внимания можно уделить особенностям работы с внедряемой ИС. Вместе с тем есть и обратная сторона: персонал привык работать в старой системе и, работая в новой ИС, необходимо привыкать к другим интерфейсам, иным ограничениям по работе с данными. Не всегда новшества положительно воспринимаются персоналом: сотрудники обычно сравнивают две системы, и поскольку опыта работы в новой системе еще мало, то результаты сравнения, как правило, не в пользу внедряемого решения. Поэтому при обучении необходимо акцентировать больше внимания на преимущества внедряемого решения, что позволит быстрее побороть психологическое непринятие нового решения.

Второе, что стоит заметить, — наличие реализации бизнес-процессов компании в старой информационной системе. При внедрении нового решения необходимо учесть опыт автоматизации бизнес-процессов компании. Особенно важен предыдущий положительный опыт при доработке внедряемой ИС. Не стоит исключать также вариант реинжиниринга бизнес-процессов компании, если имеется негативный опыт эксплуатации старой ИС.

При замене информационной системы разумно использовать имеющуюся справочную информацию. Перенос справочников из старой ИС в новую — отдельная задача. Прежде всего, нужно получить ответ на вопрос о том, что выгоднее: перенести справочную информацию вручную (здесь возможны ошибки ввода данных, длительность этапа ввода, положительный момент — будут перенесены только корректные и актуальные данные) или разрабатывать специальное программное обеспечение (это ПО будет запускаться однократно, разработка такого ПО стоит значительно дороже тиражных решений). В первую очередь следует оценить объем переносимой информации и трудоемкость разработки программного обеспечения, обеспечивающего миграцию данных. При небольших объемах справочной информации может быть логичным вариант отказа от автоматизации этого процесса. После длительной эксплуатации старой ИС в справочной информации могут быть данные, потерявшие актуальность (например, в справочнике контрагентов могут присутствовать организации, с которыми долгое время нет сотрудничества и такое сотрудничество в перспективе не предвидится, тогда в каждом конкретном случае надо решать отдельно: переносить данные таких контрагентов в новую информационную систему или нет). При переносе данных в новую ИС, конечно, не стоит переносить «мусор».

Результат замены одной ИС на другую во многом зависит от удачного выбора времени начала работы в новой информационной системе. Как правило, запуск новой ИС приурочивают к началу нового календарного года. Отчеты за предыдущий год формируют в старой ИС, в новую систему переносят только актуальные данные на конец года. К этому моменту в новой ИС уже должны быть заполнены справочники, проведены все настройки. Остается только импортировать данные на конец предыдущего года. Запуск ИС в середине календарного года потребует переноса всех данных из старой ИС, так как формирование отчетов из новой ИС невозможно без наличия ретроспективных данных.

Посмотрим теперь, как указанные закономерности проявляются в конкретном проекте замены информационной системы в ФГУП «Почта России».

3. Замена информационной системы: проблемы и пути их решения на примере ФГУП «Почта России»

В ФГУП «Почта России» завершена разработка Единой автоматизированной информационной системы «Прием подписки, экспедирование и достав-

ка периодических печатных изданий», которая позволяет автоматизировать документооборот на всех уровнях объектов почтовой связи: ОПС, почтамт, сортировочный узел, УФПС, аппарат управления ФГУП «Почта России». Разработанное решение предоставляет возможность создания единого информационного пространства по подписке на всей территории страны.

В настоящее время стоит задача внедрения этой сложной распределенной многоуровневой информационной системы взамен существующих локальных систем. Предварительный анализ и проработка методологии внедрения информационной системы в филиалах ФГУП «Почта России» показал наличие некоторых проблем и необходимость их решения на этапе подготовки к внедрению. Рассмотрим подробнее проблемы внедрения и попытаемся предложить пути их решения.

Прием подписки и последующая доставка периодических изданий циклически повторяются из года в год. Основная подписная кампания на полугодие проводится два раза в год: с сентября по декабрь и с апреля по июнь. В оставшиеся месяцы проводится так называемая текущая подписная кампания [2]. Следует заметить, что периоды проведения основной и текущей подписных кампаний могут пересекаться (например, с сентября по декабрь 2013 года проводится основная подписная кампания на 1 полугодие 2014 года, а с августа по ноябрь 2013 года проходит текущая подписная кампания на 2 полугодие 2013 года). Данные о тиражах и местах доставки изданий используются в течение всего подписного периода, длящегося полгода, а для некоторых изданий в течение года [3; 4]. Таким образом, исключена возможность начала работы в новой информационной системе «с чистого листа», т.е. с нового квартала, с нового полугодия или нового года. Первая проблема, которая возникает в процессе внедрения – перенос всех данных из старых информационных систем в новую. Поскольку в каждом из 82 УФПС работает своя информационная система, необходимо разработать 82 программы для переноса данных в УФПС и не менее 82 аналогичных программ для почтамтов (в силу технологических особенностей в почтамтах стоит программное обеспечение, отличное от ПО УФПС). Почему для почтамтов необходимо разработать больше программ, чем для УФПС? Как показало исследование, в некоторых УФПС работает несколько вариантов (два и более) программного обеспечения автоматизации технологических процессов и документооборота по подписке в почтамтах.

Единственное решение указанной проблемы – разработка для каждого филиала своего комплекта программ для переноса информации из старых баз данных в новую. Эти программы необходимо разработать и оттестировать до начала внедрения. По предварительным оценкам, это самый трудоемкий этап из всех подготовительных операций перед началом работ по внедрению. В связи с этим не играет никакой роли момент перехода на работу в новой системе.

Вторая проблема: внедрение должно производиться без остановки технологических процессов. Газеты выходят ежедневно и должны быть доставлены в день выхода. Кроме того, подписка в отделениях связи принимается ежедневно, и данные должны быть своевременно введены и обработаны. Опоздание с обработкой данных может привести к ситуации, когда не будет заказан тираж у издательства (многие издания печатают тиражом, соответствующим количеству подписчиков, такие издания, как правило, в розничную продажу не поступают и распространяются только по подписке). Предлагается устанавливать и настраивать новую информационную систему без остановки обработки информации в существующих системах. Конвертацию данных следует производить на месте внедрения после полного завершения установки и настройки новой системы, что позволит краткосрочно (на 1 - 2 часа) остановить только ввод данных. Запуск системы должен проводиться одновременно на трех уровнях сбора и обработки данных (ОПС – почтамт – УФПС). Запуск системы на уровне аппарата управления (АУП) можно производить после первичного накопления данных на уровне УФПС.

Третья проблема: обучение персонала. Больше половины персонала в настоящее время работает в операционной системе DOS, и переход на работу с информационной системой под OS Windows обычно вызывает определенные трудности. Требуется больше внимания уделять особенностям работы в новых интерфейсах и правилам ввода информации в Windows-формы. Уже разработаны программы и методики обучения в зависимости от степени подготовки персонала, таким образом можно повысить качество обучения и сократить его сроки. Кроме обучения персонала с выездом на площадки внедрения планируется использовать дистанционные средства обучения и консультирования. Для дистанционного обучения опробовано совместное использование таких программных средств как Skype (голосовое общение с обучаемыми) и TeamViewer (демонстрация порядка и правил работы в новой информационной системе, контроль за действиями обучаемых). Одна

из особенностей подготовки персонала: невозможность одновременного обучения всех сотрудников подразделений подписки филиала, так как часть сотрудников должна продолжать работу по приему и обработке информации. С целью повышения эффективности обучения, обеспечения возможности обучения сотрудников в любое удобное для них время, а также снижения количества обращений персонала в службу поддержки записаны видеоматериалы с демонстрацией порядка работы оператора при выполнении отдельных наиболее часто выполняемых технологических операций.

Четвертая проблема: несмотря на регламентацию бизнес-процесса (выпущено большое количество нормативно-технической документации [2-5]), в каждом УФПС существуют отличия, связанные с местными особенностями, возможностями действующего программного обеспечения, распределением должностных обязанностей. Наибольшую опасность на этапе внедрения представляет выполнение доработок информационной системы по всем предложениям и пожеланиям персонала каждого региона внедрения. Одна из целей внедрения единого решения – унификация бизнес-процесса во всех подразделениях почтового оператора. Следует отказаться от доработки информационной системы под конкретные региональные особенности, необходимо принимать консолидированные решения, следствием которых может стать корректировка бизнес-процесса одновременно во всех подразделениях почтового оператора. В результате такой политики внедрения клиенты по всей стране должны получить одинаковое качество оказания услуг, а почтовый оператор – унифицированный бизнес-процесс.

Пятая проблема: оппортунистическое поведение персонала. Эта проблема носит по большей мере психологический характер. Персонал привык работать в старых информационных системах, ему хорошо знаком интерфейс, известны сильные и слабые стороны работающих систем. В этой ситуации внедрение нового решения, каким бы передовым оно ни было и какие бы новые перспективы и возможности оно не предоставляло, будет с недоверием восприниматься персоналом бизнес-подразделений. Работники будут сравнивать новую систему со старой, искать недостатки новой информационной системы. Особую критику обычно вызывает интерфейс нового программного продукта (старый интерфейс всегда кажется проще и удобнее), и критика интерфейса будет сильнее, если до начала внедрения эксплуатировалось решение, разработанное под опе-

рационную систему DOS. Со стороны персонала бизнес-подразделений возможен поиск предлогов отказаться от перехода на новое решение. Вопросы, связанные с привыканием к новому интерфейсу программного продукта, отпадают по мере увеличения интенсивности использования нового программного продукта. Преодоление сопротивления персонала на начальном этапе внедрения обычно решается административными методами.

Перед началом работ был разработан регламент внедрения, в котором подробно расписаны все действия команды внедренцев, а также условия начала работ и обязанности заказчика. Регламентом предусматриваются работы по установке и настройке программного обеспечения, обучению персонала, унификации бизнес-процесса, миграции данных.

Внедрение планируется начать в 2014 году и поэтапно провести в каждом УФПС в течение одного-двух лет. На *рис. 2* [7] представлена схема информационных потоков во внедряемой системе. На схеме показаны четыре информационных уровня. В левой части схемы обозначен обмен информацией в процессе оформления подписки, в правой части – доставка периодики клиентам. На схеме хорошо видны все объекты внедрения. При планировании внедрения учитывалось и большое количество объектов внедрения: 40 000 отделений почтовой связи, около 1000 почтамтов, 82 УФПС. Объекты внедрения расположены по всей территории России. В условиях ограниченного бюджета, сжатых сроков и небольшой команды внедренцев необходимо искать новые методы работы. В качестве одной из мер, позволяющей ускорить процесс внедрения, предлагается специалистам из Москвы устанавливать программное обеспечение в ОПС и почтамтах, расположенных в областном центре и населенных пунктах, расположенных вблизи областного центра. При этом необходимо обучить ИТ-специалистов заказчика, которые смогли бы самостоятельно установить и настроить программное обеспечение в оставшихся ОПС и почтамтах.

4. Преимущества, получаемые после завершения внедрения единой информационной системы

Отдельно следует выделить преимущества, которые получит почтовый оператор в результате внедрения единого решения. Основное технологическое преимущество – это создание единого информационного пространства по подписке во ФГУП «Почта России», что предоставит возможность обмена информаци-

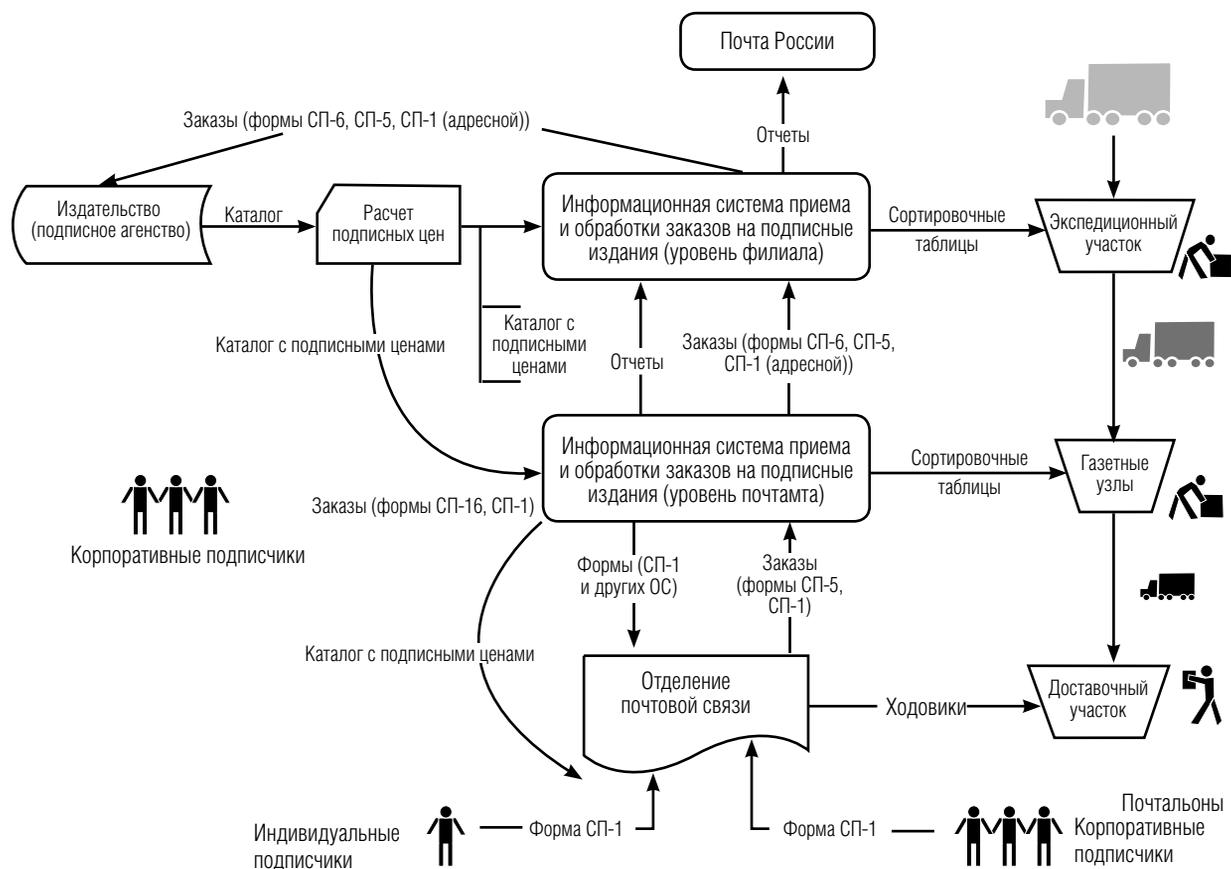


Рис. 2. Информационное пространство по подписке

ей между всеми объектами почтовой связи внутри системы. Единое информационное решение проще поддерживать специалистам центрального офиса компании, что в конечном счете приведет к снижению затрат на доработку и сопровождение программного обеспечения, позволит значительно разгрузить ИТ-специалистов в филиалах. Единое решение – это обеспечение одинакового качества оказания услуги во всех регионах страны, поддержание единой технологии подписки во всех филиалах с возможностью ее оперативной корректировки. Совершенствование технологии позволит, наконец, добиться одной из целей, которую уже давно ставит почтовый оператор – сокращение сроков доставки газет и журналов подписчикам («газета к утреннему чаю»). Централизованный расчет подписных цен, с функцией передачи их в подписные агентства для печати каталогов периодических печатных изданий с конечными подписными ценами, позволит сократить тудозатраты в бизнес-подразделениях филиалов, как минимум, на 164 чел./мес. и, как следствие, обеспечить поддержание единой тарифной политики.

Единое информационное пространство – это путь «стиранию границ» оказания услуги, а именно воз-

можность оформления подписки в каждом регионе с доставкой в адрес любого другого региона страны (в настоящее время нельзя оформить подписку в одном почтамте в адрес другого почтамта, кроме Москвы и Санкт-Петербурга). Подобное нововведение должно привести к увеличению количества подписчиков, особенно это касается филиалов организаций, основные офисы которых находятся в Москве и других крупных городах. При переезде в другой регион клиент сможет оформить переадресацию подписки.

Новые технологии позволят использовать современные каналы для оказания услуги (например, Интернет, платежные терминалы). Интернет-технологии – шаг к внедрению широкой подписки на электронные версии изданий.

Следует ожидать значительного улучшения качества управления на основе организации информационного обмена с подписными агентствами сводными данными о планируемых тиражах, прогнозирование тиражей изданий на основе ретроспективного анализа. Наличие консолидированной информации предоставит возможность оперативно влиять на ход подписной кампании на местах из центрального офиса

компании. Для повышения количества подписчиков можно рассматривать, например, такие мероприятия как проведение дополнительных рекламных акций в регионе, проведение гибкой ценовой политики в регионе и продление подписной кампании в регионе.

В настоящее время сотрудники центрального офиса получают информацию о ходе и результатах проведения подписных кампаний в виде отчетов с мест, проверка достоверности которых не представляется возможной. Оперативное получение аналитических отчетов непосредственно в центральном офисе позволит получать достоверную информацию в различных разрезах: о ходе подписки по филиалам, о приеме подписки по почтамтам и ОПС, о тиражах конкретных изданий, о взаиморасчетах с издателями, о доходах от услуги.

На основании имеющихся данных можно принимать решения о поощрении сотрудников ОПС за прием подписки, отслеживать нагрузку на сортировщиков и почтальонов по доставке изданий. Впервые в единой системе реализована функция подготовки пакета документации для доставки периодических изданий (в настоящее время организация доставки на уровне газетного узла осуществляется без использования средств вычислительной техники). Появляется возможность отслеживать, какие издания поступили и были доставлены, какие были недопоставки (или перепоставки) и брак.

Заключение

В статье выявлены закономерности, проявляющиеся при внедрении новой информационной системы взамен существующей. На примере проекта замены локальных информационных систем на единое решение во ФГУП «Почта России» было проведено исследование проблем и «узких мест», которые могут возникнуть в процессе внедрения распределенной информационной системы в многофилиальной организации взамен существующих локальных решений, предложены возможные пути их решения.

В качестве рекомендаций следует отметить следующее: из-за особенностей бизнес-процесса начало запуска новой информационной системы можно не привязывать к конкретной дате, при этом необходима миграция всех данных из старых систем в новую на всех уровнях; для успешного перехода к работе в новой системе нужно больше внимания уделить подготовке персонала, чтобы облегчить его работу в новых интерфейсах и снизить негативную реакцию на новое программное обеспечение.

В заключение хотелось бы отметить, что переход российского почтового оператора на работу с единой информационной системой позволит вывести услугу приема подписки на новый качественный уровень. ■

Литература

1. Официальный сайт ФГУП «Почта России». [Электронный ресурс]: <http://www.russianpost.ru/gr/companу/ru/home> (дата обращения 01.12.2013).
2. Инструкция по приему подписки и обработке подписной документации на периодические печатные издания / Утверждена приказом ФСПС России от 22.10.1996, № 87. М., 1996. 78 с.
3. Инструкция о порядке экспедирования периодических печатных изданий / Утверждена приказом ФСПС России от 22.10.1996, № 87. М., 1996. 62 с.
4. Инструкция о порядке обработки периодических печатных изданий в газетных узлах и отделениях почтовой связи / Утверждена приказом ФСПС России от 22.10.1996, № 87. М., 1996. 52 с.
5. Инструкция о перевозке периодических печатных изданий / Утверждена приказом ФСПС России от 22.10.1996, № 87. М., 1996. 32 с.
6. Правила распространения периодических печатных изданий по подписке / Утверждены Постановлением Правительства РФ от 01.11.2001, № 759.
7. Моргунов А.Ф. Подписка на периодические печатные издания: продление сроков приема заказов и ускорение доставки // Логистика. 2011. № 3. С. 36-38.
8. Грекул В.И., Денищенко Г.Н., Коровкина Н.Л. Проектирование информационных систем. Учебное пособие. М.: ИНТУИТ, 2005. 299 с.
9. Калянов Г.Н. Теория и практика реорганизации бизнес-процессов. М.: СИНТЕГ, 2000. 212 с.
10. Заичкин Н.И. Управленческие решения в информационном пространстве промышленных организаций. Автореф. дис. на соиск. учен. степ. д-ра экон. наук. М.: ГУУ, 2003. 57 с.

IMPLEMENTATION OF A DISTRIBUTED INFORMATION SYSTEM IN A MULTI-BRANCH ORGANIZATION: FSUE «RUSSIAN POST» CASE STUDY

Alexander MORGUNOV,

Associate Professor, Department of Corporate Information Systems, Faculty of Business Informatics, National Research University Higher School of Economics

Address: 20, Myasnitskaya str., Moscow, 101000, Russian Federation

E-mail: amorgunov@hse.ru

Implementation of information systems in a large company with a wide network of branches is usually associated with a number of difficulties of organizational and technical nature. Implementation projects are time consuming not always lead to desired results.

The paper discusses the problems related to implementation of the information system of automation of technological processes and workflow of subscription for periodical printed publications in branches of Federal State Unitary Enterprise (FSUE) «Russian Post». General performance of the business process automation before the implementation is estimated, organizational and technical

problems that arise during implementation of a new information system are examined. Such problems are associated with continuity of the technological process, impossibility of performing operation with «empty database», problems of personnel development, unification of business processes, as well as large quantity of implementation objects (more than 42 000). As a result of the study the advantages of implementation of the unified information system, which may be gained by post operators are analyzed.

The article's matters can be of interest to experts studying the implementation of information system.

Key words: information system, implementation, unification, automation, subscriptions for periodical printed publications, post.

References

1. Official website of FSUE «Russian Post». Available at: <http://www.russianpost.ru/rp/company/ru/home> (accessed 01 December 2013).
2. Federal Post Communications Service of Russia (1996) Instrukcija po priemu podpiski i obrabotke podpisnoj dokumentacii na periodicheskie pechatnye izdanija [The Statement on Reception of Subscription and Subscription Processing Documentation to Periodicals], Moscow. (in Russian).
3. Federal Post Communications Service of Russia (1996) Instrukcija o porjadke jekspedirovanija periodicheskikh pechatnyh izdanij [Instruction on the Order of Forwarding of Periodicals], Moscow. (in Russian).
4. Federal Post Communications Service of Russia (1996) *Instrukcija o porjadke obrabotki periodicheskikh pechatnyh izdanij v gazetnyh uzlah i otdelenijah pochtovoj svjazi* [Instruction on the procedure of processing of periodic printed publications in the newspaper nodes and post offices], Moscow. (in Russian).
5. Federal Post Communications Service of Russia (1996) Instrukcija o perevozke periodicheskikh pechatnyh izdanij [The Instruction on Transportation of Periodicals], Moscow. (in Russian).
6. The Government of the Russian Federation (2001) Pravila rasprostraneniya periodicheskikh pechatnyh izdanij po podpiske [The Rules of Distribution of Periodical Printed Editions by Subscription]. Resolution of the Government of the Russian Federation, 01.11.2001, No. 759. (in Russian).
7. Morgunov A. (2011) Podpiska na periodicheskie pechatnye izdanija: prodlenie srokov priema zakazov i uskorenje dostavki [Subscription to periodicals: extension of deadlines for taking orders and acceleration of delivery]. *Logistic*, no 3 (56). pp. 36-38. (in Russian).
8. Grekul V., Denishhenko G., Korovkina N. (2005) Proektirovanie informacionnyh sistem. Uchebnoe posobie [Designing of information systems. School-book]. Moscow: INTUIT. (in Russian).
9. Kaljanov G. (2000) Teorija i praktika reorganizacii biznes-processov [Theory and practice of business process re-engineering]. Moscow: SINTEG. (in Russian)
10. Zaichkin N. (2003) Upravlencheskie reshenija v informacionnom prostranstve promyslennyh organizacij [Management decisions in the information space of industrial organizations]. PhD Thesis. Moscow. (in Russian)

К ВОПРОСУ ОБ ОЦЕНКЕ УРОВНЯ РАЗВИТИЯ СИСТЕМ УПРАВЛЕНИЯ ЭФФЕКТИВНОСТЬЮ

Д.В. Исаев,

кандидат экономических наук, доцент кафедры бизнес-аналитики, факультет бизнес-информатики, Национальный исследовательский университет «Высшая школа экономики»

Адрес: 101000, г. Москва, ул. Мясницкая, д. 20

E-mail: disaev@hse.ru

Статья посвящена вопросам оценки уровня развития систем управления эффективностью (performance management systems, PMSs), которые рассматриваются в одной из существующих трактовок, как системы информационной поддержки корпоративного управления и стратегического менеджмента. Такие системы предназначены для решения задач сбора, хранения, аналитической обработки и представления информации, являющейся ключевой для обеспечения информационной прозрачности организаций и поддержки принятия стратегических управленческих решений внешними и внутренними заинтересованными лицами. Цель исследования – предложить подход к оценке уровня развития таких систем. Для этого в статье рассмотрены существующие подходы к определению уровня развития систем управления и информационных систем, сформулированы общие принципы оценки, разработаны методические рекомендации в части оценки уровня развития систем управления эффективностью.

Предлагаемый подход к оценке уровня развития систем управления эффективностью основывается на иерархической концептуальной модели, включающей такие элементы, как функциональные блоки, функциональные модули и аналитические функции. В этом случае применим «восходящий» принцип, когда оценки вышестоящих элементов иерархической структуры (вплоть до системы в целом) основываются на оценках субординированных элементов нижестоящих уровней. При этом каждый элемент оценивается с точки зрения методов и процессов обработки управленческой информации, информационных систем, персонала, а также интеграции со смежными элементами, качества данных, результативности и управления. Также обосновывается целесообразность оценки уровня развития системы в динамике, в сопоставлении с определенными целевыми уровнями.

Ключевые слова: управление эффективностью, система управления эффективностью, уровень развития, стратегический менеджмент, информационная поддержка управления, траектория развития.

Введение

Актуальность задачи оценки уровня развития систем управления эффективностью (performance management systems, PMSs) объясняется тем, что в современных условиях роль

таких систем в управлении предприятиями и организациями становится все более значительной. При этом, несмотря на определенные различия в существующих определениях и концептуальных моделях [1 - 5], основное предназначение систем управления эффективностью большинством авторов трактуется

ся примерно одинаково: это системы, деятельность которых направлена на преодоление разрыва между стратегическим и текущим уровнями управления, а также на обеспечение гибкости и адаптивности стратегического развития организации.

Одна из возможных трактовок понятия системы управления эффективностью предполагает ее рассмотрение как системы информационной поддержки корпоративного управления и стратегического менеджмента [6]. Данная трактовка системы управления эффективностью является более узкой, чем трактовки некоторых других авторов (например, в модели А.Феррейры и Д.Отли, помимо информационных потоков, систем и сетей, содержатся такие элементы, как видение и миссия, ключевые факторы успеха, организационная структура, система мотивации и некоторые другие [1]). Тем не менее, в ходе дальнейшего изложения будем рассматривать системы управления эффективностью именно с точки зрения информационного аспекта.

1. Роль мониторинга в развитии систем управления эффективностью

Управление развитием систем управления эффективностью предусматривает три уровня иерархии, на каждом из которых присутствуют функции проектирования, планирования и контроля [7; 8]. Разным уровням соответствуют разные объекты управления: на верхнем уровне в качестве объекта управления выступает система в целом, на среднем – входящие в ее состав подсистемы, на нижнем – проекты внедрения или совершенствования отдельных элементов.

На каждом из трех уровней управления значительную роль играет мониторинг – регулярный процесс сбора, структурирования и обобщения информации о характеристиках, событиях и явлениях, относящихся к объекту мониторинга – системе в целом или какой-то ее части. Именно мониторинг позволяет сформировать мнение о текущем состоянии и динамике развития системы. Более того, одни и те же показатели могут использоваться не только для мониторинга (отражения состояния системы в текущий момент времени или в прошлом), но и для планирования (описания состояния системы в будущем). Это, в свою очередь, позволяет контролировать развитие системы путем анализа отклонений фактических значений ее параметров от запланированных.

Показатели развития также играют существен-

ную роль на этапе проектирования системы. В этом случае они выступают в качестве средства целевой ориентации развития: проектирование системы ведется таким образом, чтобы придать ей определенные свойства и обеспечить желаемые (целевые) характеристики.

При разработке рекомендаций в области оценки уровня развития системы следует принимать во внимание некоторые основополагающие принципы, к числу которых относятся принципы декомпозиции и измеримости [9]. Принцип декомпозиции исходит из того, что уровень развития определяется не только для системы в целом, но и для отдельных ее составляющих, а также для отдельных аспектов, определяющих уровень развития. Принцип измеримости предусматривает описание объектов управления в виде количественных показателей, которые могут быть как ретроспективными (мониторинг), так и перспективными (планирование). При этом в случаях, когда определенные характеристики имеют качественный характер, необходима квантификация – выражение качественных характеристик в количественном виде.

Говоря о применимости разных типов показателей для оценки уровня развития системы (и, соответственно, для мониторинга и планирования), следует отметить ведущую роль качественных показателей, которые нельзя измерить, но можно оценить и соотнести с определенной шкалой. В результате происходит выражение качественных показателей в количественном виде.

Оценка уровня развития также может применяться при анализе эффективности инвестиций, направляемых на развитие системы – внедрение новых методов, регламентов и информационных систем, реорганизацию систем управления, повышение квалификации персонала и т.д. Для такого анализа необходимо сопоставлять финансовые характеристики затратной части проектов с изменениями показателей уровня развития системы, многие из которых по своей сущности являются качественными.

2. Практика применения моделей, основанных на уровнях зрелости

Поскольку существенную роль в оценке уровня развития некоторых типов организационно-технических систем (организаций, систем управления, информационных систем) играют качественные показатели, в области оценки уровня развития таких

систем широкое распространение получил подход, основанный на упорядоченных по возрастанию пороговых уровнях зрелости (maturity levels), отражающих качественные различия в развитии оцениваемых объектов.

Одной из первых таких моделей является модель зрелости процессов разработки программного обеспечения (Capability Maturity Model for Software), разработанная для оценки уровня развития процессов разработки программного обеспечения [10]. Впоследствии принципы оценки процессов на основе пороговых уровней зрелости нашли развитие в разных версиях документации «Задачи контроля для информационных и смежных технологий» (Control Objectives for Information and Related Technology, COBIT) [11; 12].

Модели, основанные на уровнях зрелости, применяются не только по отношению к процессам, но и для оценки уровня развития информационных систем. Одной из таких разработок является модель оценки уровня зрелости систем бизнес-интеллекта, разработанная Институтом хранилищ данных (The Data Warehouse Institute, TDWI) [13]. Еще одна методика, разработанная TDWI, связана с оценкой решений, предназначенных для анализа больших данных (big data) [14].

В качестве примеров других разработок, основанных на уровнях зрелости, можно привести методики оценки уровня развития в таких областях, как управление информационными технологиями организации [15], архитектура предприятия [16 - 18], соответствие бизнеса и информационных технологий [19], сервисно-ориентированная архитектура [20].

Одна из методик оценки уровня развития разработана специально для оценки систем управления эффективностью и предусматривает определение индекса управления эффективностью (performance management index, PMI) [21; 22]. Методика предусматривает пять пороговых уровней зрелости. На первом («начальном») уровне развития характеризуется как «информационный силос», этому уровню свойственны локальный характер применяемых решений и средств формирования управленческой отчетности, низкое качество данных, отсутствие корпоративных стандартов управления эффективностью и централизованного управления системой. На втором («повторяемом») уровне организация приходит к пониманию значимости управления эффективностью для достижения своих целей, при этом сами цели также осознаны и определены в яв-

ном виде. На этом уровне применяются финансовые информационные системы, а также витрины данных и интерактивные системы управленческой отчетности. На третьем («определенном») уровне обеспечивается мониторинг реализации корпоративной стратегии и принятия управленческих решений на основе фактической информации. Для этого применяются системы корпоративного планирования и прогнозирования, управление по ключевым показателям эффективности, а также хранилища данных, панели индикаторов и средства централизованной поддержки основных данных. На четвертом («управляемом») уровне активно применяются аналитические методы, модели и системы. Наконец, на пятом («оптимизированном») уровне система управления эффективностью превращается в основной инструмент стратегического управления организацией.

Несомненным достоинством моделей, основанных на пороговых уровнях, является возможность рассмотрения не только количественных, но и качественных характеристик оцениваемых объектов. Кроме того, следует отметить системный подход и аналитический характер моделей: уровень развития оцениваемого объекта (системы) в целом складывается из уровней развития по каждому из аспектов.

В то же время существующие модели имеют ряд существенных ограничений.

Прежде всего, следует отметить то, что в качестве объекта оценки выступает только система в целом, возможность оценки отдельных ее элементов не рассматривается. Кроме того, не рассматриваются вопросы привязки этих элементов к организационной структуре или структуре управления организацией. Еще одна проблемная область связана с формальностью некоторых критериев, которая, как правило, сопровождается недостаточно четким описанием процедур и алгоритмов формирования оценок по этим критериям. Что касается консолидации (обобщения) оценок, то здесь применяется метод взвешенной суммы критериев. Помимо ряда достоинств (понятность, вычислительная простота) этому методу присущ целый ряд неустраняемых недостатков принципиального характера, к числу которых относятся отсутствие у интегрального показателя содержательного смысла, постоянство коэффициентов важности критериев при любых их значениях, невыполнение условия взаимонезависимости критериев, необоснованное рассмотрение общей шкалы критериев как количественной, необоснованное принятие допущения о равномерности общей шка-

статистических методов. Для машинной обработки данных информация должна быть четко формализована и как можно лучше структурирована. Поэтому наиболее подходящим для решения задачи методом является анкетный опрос с подготовленными формализованными вопросами.

Для повышения достоверности экспертных оценок чаще всего применяются метод решающих матриц [5] и метод Дельфи [6]. Метод решающих матриц предполагает разбиение проблемы на несколько подзадач и пошаговое получение экспертных оценок. Экспертам проще ответить на детализированные частные вопросы, что приводит к улучшению качества анкетирования. В процессе сбора последовательно формируется все более полная картина, позволяющая всесторонне оценить ситуацию и выработать обоснованное управленческое решение. При данной методике для создания базы знаний достаточной полноты необходимо многократное анкетирование большого числа экспертов, что является технически сложной и затратной процедурой.

Наиболее распространенным методом сбора и обработки экспертных сведений является метод Дельфи. При использовании этого метода проводится многоэтапное анкетирование с обратной связью с целью приведения ответов экспертов к единому мнению большинства. При этом мнения экспертов, которые значительно отклоняются от главного направления, фактически отбрасываются. Компетентность экспертов в существующих системах также оценивается с точки зрения близости их мнений к основному вектору. Однако во многих случаях именно особые мнения могут оказаться верными, и классический подход приводит к существенному искажению конечного результата.

Исходя из вышесказанного, предлагается использовать модифицированный алгоритм сбора и обработки данных, основанный на классическом подходе и предусматривающий первичное анкетирование большого числа экспертов и учет всех экспертных мнений, но в дальнейшем с проведением опросов выделенной группы экспертов.

Для правильной оценки разброса экспертных мнений необходимо учитывать компетенцию экспертов. К методам оценки компетенции экспертов относят следующие группы: априорные, апостериорные и тестовые. Априорные методы позволяют получить оценку до проведения сессий, в то время как апостериорная оценка фактически является

оценкой по результатам сбора экспертных сведений. К апостериорным методам, например, относится метод определения компетенции на основе степени отклонения мнения эксперта от консолидированного мнения большинства.

Оценка компетенции эксперта по степени отклонения его мнения от мнения большинства может привести к тому, что мнение высококомпетентного эксперта будет проигнорировано, так как оно не совпадает с основной совокупностью мнений менее компетентных экспертов.

Тестовый метод наиболее надежен, так как использует заранее известное суждение, принятое как эталон, что позволяет оценить степень приближенности знаний эксперта к истинным. Этот метод используется на первом этапе опроса и может выдаваться как экспертное задание. Однако, при проведении форсайт-исследований такой метод сложен в реализации, так как правильность ответов экспертов можно проверить только по прошествии длительного периода времени.

Априорные оценки компетенции могут быть получены с применением одного из описанных ниже методов. В основе метода самооценки лежит самостоятельное определение экспертом уровня собственных знаний. Данный метод основан на предположении, что более уверенный в своих знаниях эксперт, как правило, имеет более высокий уровень самооценки. Судейский метод предусматривает взаимную оценку экспертами уровня компетентности (при условии знакомства с результатами научной деятельности) с заданием критериев компетентности или формированием группы признанных компетентных экспертов, которые оценивают других. Этот метод имеет существенный недостаток — влияние личностных отношений между экспертами на результат оценки. Наиболее объективным считается метод оценки компетенции по вторичным источникам (публикации, индекс цитирования, справочники профессиональных сообществ, рейтинги и топ-листы профессиональных журналов, итоги конкурсов и т.п.).

Также к методам оценки по вторичным источникам можно отнести метод «скрытых коллективов» [7]. В процессе построения графа отношений научных деятелей на основе анализа совместной публикационной активности и картины цитирования публикаций выявляются устойчивые группы экспертов. Метод позволяет ранжировать специалистов заданной предметной области внутри групп, а

лы критериев и некоторые другие [23]. Еще одним фактором ограниченности существующих методик является отсутствие целевой ориентации развития и игнорирование того факта, что разные организации (в зависимости от бизнеса, рынка или этапа своего развития) имеют разные потребности в развитии своих систем управления и информационных систем. Наконец, ограниченность существующих методик проявляется еще и в том, что все оценки носят статический характер, тогда как несомненный интерес представляет динамика развития, т.е. изменение уровня развития системы с течением времени.

3. Оценка уровня развития системы управления эффективностью на основе ее концептуальной модели

Детальный анализ системы управления эффективностью, позволяющий оценить уровень развития как системы в целом, так и ее составляющих, может быть произведен на основе ее концептуальной модели. Концептуальное моделирование системы управления эффективностью (напомним, что мы рассматриваем такие системы в одной из существующих трактовок, как системы информационной поддержки корпоративного управления и стратегического менеджмента) предусматривает построение информационно-логической модели, основными элементами которой являются информационные потоки, внешние информационные объекты, функциональные блоки, функциональные модули и аналитические функции [24]. Под информационным потоком понимается обособленная часть управленческой информации, подлежащая передаче между, как минимум, двумя информационными объектами – источником информации и ее получателем. Источники и получатели информационных потоков, находящиеся за пределами системы, представляют собой внешние информационные объекты. Информационные объекты, находящиеся внутри системы и осуществляющие преобразование управленческой информации, укрупненно представляются в виде функциональных блоков. Базовая информационно-логическая модель, являющаяся общей для предприятий и организаций разных отраслей, включает четыре функциональных блока: блок стратегического анализа и стратегического выбора, блок целевого управления, блок корпоративного планирования и бюджетирования и блок корпоративной отчетности [25]. Дальнейшая детализация функциональных блоков предусматривает их декомпозицию на

функциональные модули, которые характеризуют определенный класс задач обработки управленческой информации, находящихся в рамках той или иной предметной области. Преобразование входящих информационных потоков в исходящие осуществляется посредством аналитических функций, которые могут иметь как проектный, так и процессный характер.

В случае наличия концептуальной информационно-логической модели системы оценка уровня развития может быть организована по восходящему («снизу вверх») принципу: сначала оцениваются конечные («листовые») элементы системы, т.е. аналитические функции, затем, на основе оценок функций, – функциональные модули, затем – функциональные блоки и, наконец, система в целом. Поскольку многие характеристики системы носят качественный характер и не могут быть надежно измерены в количественном выражении, то для оценки уровня развития могут быть приняты пороговые уровни зрелости, как и в модели «performance management index» (PMI) [21; 22].

Оценка аналитических функций (листовых элементов) может осуществляться на основе перспектив (аналог компонентов в модели PMI), аспектов (аналог субкомпонентов в модели PMI) и критериев.

Перспективы носят укрупненный характер, каждая из них представляет собой взгляд на систему с определенной точки зрения и, таким образом, служит для лучшего понимания системы в целом. Основные перспективы характеризуют внутреннее содержание оцениваемых аналитических функций: к ним относятся применяемые аналитические методы и модели, процессы и регламенты обработки управленческой информации, информационные системы, при помощи которых происходит обработка информации, а также соответствующий управленческий персонал. Дополнительные перспективы характеризуют интеграцию (каналы взаимодействия той или иной функции с другими элементами модели), качество входящих данных, результативность (качественные характеристики информации, являющейся исходящей для рассматриваемых функций), а также управление и контроль деятельности оцениваемых аналитических функций.

Каждая перспектива рассматривается в разрезе аспектов. Аспекты представляют собой характеристики системы, в совокупности определяющие уровень ее развития в рамках данной перспективы.

Каждый из аспектов оценивается относительно пороговых уровней зрелости по одному или нескольким критериям. Для этого для каждого критерия должна быть определена его «привязка» к уровням зрелости: в каком случае аналитическая функция может быть отнесена к тому или иному уровню зрелости. На этой основе осуществляется квантификация – оценка рассматриваемой аналитической функции относительно выбранной шкалы по соответствующему критерию, даже если этот критерий носит качественный характер.

После того, как рассматриваемая аналитическая функция оказывается оцененной по всем критериям, становится возможным сформировать интегральную оценку функции: сначала – по каждому из аспектов, затем – по каждой из перспектив и, наконец, оценку функции в целом.

Обобщение оценок уровня развития аналитической функции по критериям, аспектам и перспективам может осуществляться разными способами. Прежде всего, может применяться правило: вышестоящая оценка равна минимальной из нижестоящих оценок (например, оценка по аспекту равна минимальной из оценок по всем критериям, входящим в состав данного аспекта). Кроме того, если есть возможность определить относительную важность элементов оценки (критериев в составе аспектов, аспектов в составе перспектив и самих перспектив), то может быть применен «традиционный» метод взвешенной суммы критериев: оценка вышестоящего уровня формируется путем суммирования оценок нижестоящего уровня, умноженных на коэффициенты их относительной важности. Кроме того, метод взвешенной суммы критериев может быть модифицирован путем замены значений некоторых оценок в соответствии со следующим правилом: сначала определяется минимальная (среди всех критериев) оценка, после чего все оценки, превышающие минимальную на две и более позиции, понижаются до уровня «минимальная оценка плюс один».

Оценка вышестоящих по отношению к аналитическим функциям элементов концептуальной модели (функциональных модулей, функциональных блоков, системы в целом) строится аналогичным образом: здесь тоже могут применяться обобщение на основе минимальной оценки нижестоящего элемента, взвешенной суммы критериев или взвешенной суммы критериев с усеченными значениями. Важно отметить, что оценку аналитических функций и вышестоящих элементов (функциональных модулей, функциональных блоков, системы в це-

лом) можно осуществлять как по полному набору аспектов и перспектив, так и отдельно по конкретному аспекту, конкретной перспективе или некоторому их сочетанию.

Как для системы в целом, так и для отдельных ее элементов целесообразно применять целевые уровни развития. Они, как и фактические значения уровня развития, формируются на основе уровней зрелости. Однако если уровни зрелости определяются объективными критериями и являются общими для всех организаций (независимо от типов, размеров и отраслевой принадлежности), то целевые уровни развития определяются каждой организацией индивидуально и отражают видение конкретной организации в отношении «приемлемого» или «желаемого» уровней развития своей системы управления эффективностью. При этом вполне допускается применение как одного, так и нескольких целевых уровней. Например, организация может установить «уровень необходимости» (уровень развития, достижение которого является обязательным, а недостижение может привести к серьезным проблемам в системе управления) и «уровень достаточности» (желаемый уровень развития, по достижении которого дальнейшие инвестиции в развитие уже не будут существенно влиять на эффективность системы управления).

Наконец, и фактические, и целевые значения показателей развития должны рассматриваться в динамике. Что касается фактических значений, то их изменение с течением времени, отражающее совершенствование системы и улучшение ее характеристик, выглядит вполне естественным. Это позволяет перейти от отдельных показателей к траектории развития системы, представляющей собой изменение значений фактических показателей уровня развития с течением времени. Более того, траектория развития может быть не только фактической, но и плановой (т.е. относящейся не к прошлому, а к будущему). Плановая траектория определяется программой развития и результативностью запланированных проектов, поскольку именно они определяют положительные изменения характеристик системы и, соответственно, динамику общего уровня развития. Более того, поскольку сроки реализации проектов и их влияние на систему часто имеют вероятностный характер, программе развития может соответствовать семейство перспективных траекторий, каждая из которых может иметь место с определенной вероятностью.

Помимо траектории развития системы в целом, аналогичные траектории могут строиться для отдельных ее элементов, а также для отдельных перспектив и аспектов развития. Кроме того, следует отметить, что требования, предъявляемые организацией к своей системе управления, также могут изменяться с течением времени. Поэтому целевые уровни развития, наряду с фактическими и плановыми траекториями, также должны рассматриваться в динамике.

Заключение

Вопросы оценки уровня развития систем управления эффективностью имеют большое значение для мониторинга развития таких систем, планирования их развития, а также для проектирования систем и их составляющих. При оценке систем управления эффективностью существенную роль играют качественные характеристики, которые можно оценить экспертным путем и затем соотнести с определенной шкалой. Поэтому наиболее перспективным в данном случае является подход, основанный на упорядоченных по возрастанию пороговых уровнях зрелости, отражающих качественные различия в развитии оцениваемых объектов. Различные методики, основанные на уровнях зрелости, сегодня находят широкое применение для оценки процессов, систем управления и информационных систем. В то же время существующие модели имеют ряд существенных ограничений, к числу которых относятся отказ от оценки отдельных составляющих системы, отсутствие привязки к организационной структуре, формальность применяемых критериев, недостаточная четкость процедур и алгоритмов

оценки, отсутствие практики применения целевых значений уровня развития, а также статичность.

В качестве подхода к оценке уровня развития системы управления эффективностью предлагается подход, основанный на ее концептуальной модели, основными элементами которой являются информационные потоки, внешние информационные объекты, функциональные блоки, функциональные модули и аналитические функции. В этом случае оценка уровня развития может быть организована по восходящему принципу: сначала оцениваются аналитические функции, затем – функциональные модули, функциональные блоки и система в целом. При этом сама оценка осуществляется на основе перспектив, аспектов и критериев. Это позволяет оценивать аналитические функции и вышестоящие элементы как по полному набору аспектов и перспектив, так и отдельно по конкретному аспекту, конкретной перспективе или некоторому их сочетанию.

Помимо фактических значений уровня развития, применяются целевые уровни развития, которые также формируются на основе уровней зрелости. Целевые уровни определяются каждой организацией индивидуально и отражают видение конкретной организации в отношении «приемлемого» или «желаемого» уровней развития своей системы управления эффективностью.

Также представляется целесообразным, чтобы значения показателей уровня развития (включая целевые) рассматривались в динамике. При этом может рассматриваться как траектория развития системы в целом, так и траектории развития отдельных элементов. ■

Литература

1. Ferreira A., Otley D. The design and use of performance management systems: An extended framework for analysis // Management Accounting Research. 2009. No.20. P. 263-282.
2. Business Performance Management Industry Framework Document. Final version 5.0. BPM Standards Group, 2005. 27 pp.
3. Стратегический разрыв: технологии воплощения корпоративной стратегии в жизнь / М.Ковени и [др.]: Пер. с англ. М.: Альпина Бизнес Букс, 2004. 232 с.
4. Кокинз Г. Управление результативностью: Как преодолеть разрыв между объявленной стратегией и реальными процессами: Пер. с англ. М.: Альпина Бизнес Букс, 2008. 318 с.
5. Эккерсон У.У. Панели индикаторов как инструмент управления: ключевые показатели эффективности, мониторинг деятельности, оценка результатов: Пер. с англ. М.: Альпина Бизнес Букс, 2007. 396 с.
6. Исаев Д.В. Корпоративное управление и стратегический менеджмент: информационный аспект. М.: Изд. дом ГУ-ВШЭ, 2010. 219 с.
7. Исаев Д.В. Процесс управления развитием систем информационной поддержки корпоративного управления и стратегического менеджмента // Аудит и финансовый анализ. 2012. №4. С. 276-281.
8. Исаев Д.В. Развитие систем информационной поддержки корпоративного управления и стратегиче-

- ского менеджмента // Бизнес-информатика. 2011. №2 (16). С. 56-62.
9. Исаев Д.В. Мониторинг и планирование развития систем информационной поддержки корпоративного управления и стратегического менеджмента // Бизнес-информатика. 2012. №3 (21). С. 63-69.
 10. Capability Maturity Model for Software, Version 1.1. Technical Report. CMU/SEI-93-TR-024; ESC-TR-93-177. February 1993 // Software Engineering Institute, Carnegie Mellon University, 1993. 64 pp.
 11. COBIT 4.1. IT Governance Institute. 2007. 196 pp.
 12. COBIT 5: A Business Framework for the Governance and Management of Enterprise IT. ISACA. 2012.
 13. Eckerson W.W. Beyond the Basics: Accelerating BI Maturity. Renton, WA: TDWI, 2007. 15 pp.
 14. Halper F., Krishnan K. TDWI Bid Data Maturity Model Guide: Interpreting Your Assessment Score. Renton, WA: TDWI, 2013. 18 pp.
 15. Weill P., Ross J. IT Governance: How Top Performers Manage IT Decision Rights for Superior Results. Boston: Harvard School Press, 2004. 267 pp.
 16. Ross J., Weill P., Robertson D. Enterprise Architecture as Strategy: Creating a Foundation for Business Execution. Boston: Harvard School Press, 2006. 234 pp.
 17. NASCIO Enterprise Architecture Maturity Model. Version 1.3. December 2003. Lexington, KY: National Association of State Chief Information Officers (NASCIO), 2003. 16 pp.
 18. Extended Enterprise Architecture Maturity Model Support Guide. Version 2.0 / Editorial writer: J.Schekkerman. Amersfoort: Institute for Enterprise Architecture Developments (IFEAD), 2006. 4 pp.
 19. Luftman J., Kempaiah R. An Update on Business-IT Alignment: A Line Has Been Drawn. MIS Quarterly Executive, 2007. Vol. 6. No. 3 (September 2007). P. 165-177.
 20. Perko J. IT Governance and Enterprise Architecture as Prerequisites for Assimilation of Service-Oriented Architecture: An Empirical Study of Large Finnish Companies / Thesis for the Degree of Doctor of Technology. Tampere: Tampere University of Technology, 2008. 248 pp.
 21. Aho M. A Capability Maturity Model for Corporate Performance Management – An Empirical Study in Large Finnish Manufacturing Companies // Proceedings of EBRF 2009 – Research Forum to Understand Business in Knowledge Society. Jyväskylä, Finland. September 23-25, 2009.
 22. Aho M. What is your PMI? A Model for Assessing the Maturity of Performance Management in Organizations // Proceedings of 'Performance Management: From Strategy to Delivery' (PMA 2012) conference. University of Cambridge, UK. July 11-13, 2012.
 23. Подиновский В.В., Потапов М.А. Метод взвешенной суммы критериев в анализе многокритериальных решений: Pro et contra // Бизнес-информатика. 2013. № 3 (25). С. 41-48.
 24. Исаев Д.В. Проектирование систем информационной поддержки корпоративного управления и стратегического менеджмента // Аудит и финансовый анализ. 2013. №6. С. 329-334.
 25. Isaev D. Performance Management Information Support System: A Conceptual Model // European Journal of Economics, Finance and Administrative Sciences. 2012. Issue 52. P. 6-20.

ON EVALUATION OF PERFORMANCE MANAGEMENT SYSTEMS MATURITY

Dmitry ISAEV,

Associate Professor, Department of Business Analytics, Faculty of Business Informatics, National Research University Higher School of Economics.

Address: 20, Myasnitskaya str., Moscow, 101000, Russian Federation

E-mail: disaev@hse.ru

The paper focuses on estimating maturity of performance management systems, which are considered according to one of the possible treatments, as systems for information support of corporate governance and strategic management. Such systems are focused into the tasks of gathering, storage, analytical processing and presentation of information, which is critical for organizations' information transparency and strategic decision making performed by external and internal stakeholders. The purpose of this paper is to advance a approach to evaluating maturity of such systems. For this purpose, we have considered existing approaches to maturity evaluation for management and information systems, formulated general principles of maturity evaluation, and developed methodological recommendations in the field of performance management systems maturity evaluation.

The proposed approach to evaluating performance management system maturity relies on its hierarchical conceptual model that includes such elements as functional blocks, functional modules and analytical functions. In this case a «bottom-up» principle is applicable: evaluation of higher level elements maturity (up to the system as a whole) is performed relying on estimates of subordinated lower level elements. At that, every element is estimated from viewpoints of data processing methods and processes, information systems, personnel, as well as integration with complementary elements, data quality, effectiveness and governance. Advisability of maturity evaluation in dynamics, as well as comparison with certain target levels is also justified.

Key words: performance management, performance management system, maturity level, strategic management, information support, development trajectory.

References

1. Ferreira A., Otley D. (2009) The design and use of performance management systems: An extended framework for analysis. *Management Accounting Research*, no. 20, p. 263-282.
2. BPM Standards Group (2005) Business Performance Management Industry Framework Document. Final version 5.0. BPM Standards Group.
3. Coveney M., Ganster D., Hartlen B., King D. (2004) *Strategicheskij razryv: tehnologii voploshhenija korporativnoj strategii v zhizn'* [The Strategy Gap: Leveraging Technology to Execute Winning Strategies]. Moscow: Alpina Business Books. (in Russian)
4. Cokins G. (2008) *Upravlenie rezul'tativnost'ju: Kak preodolet' razryv mezhdru ob'javlennoj strategiej i real'nymi processami* [Performance Management: Finding the Missing Pieces (to Close the Intelligence Gap)]. Moscow: Alpina Business Books. (in Russian)
5. Eckerson W.W. (2007) *Paneli indikatorov kak instrument upravlenija: kljuchevye pokazateli jeffektivnosti, monitoring dejatel'nosti, ocenka rezul'tatov* [Performance Dashboards: Measuring, Monitoring and Managing Your Business]. Moscow: Alpina Business Books. (in Russian)
6. Isaev D.V. (2010) *Korporativnoe upravlenie i strategicheskij menedzhment: informacionnyj aspekt* [Corporate Governance and Strategic Management: Information Aspect]. Moscow: HSE. (in Russian)
7. Isaev D.V. (2012) *Process upravlenija razvitiem sistem informacionnoj podderzhki korporativnogo upravlenija i strategicheskogo menedzhmenta* [A Management Process of Corporate Governance and Strategic Management Information Support Systems Development]. *Audit and Financial Analysis*, no. 4, p. 276-281. (in Russian)
8. Isaev D.V. (2011) *Razvitie sistem informacionnoj podderzhki korporativnogo upravlenija i strategicheskogo menedzhmenta* [Development of Information Support Systems for Corporate Governance and Strategic Management]. *Business Informatics*, no. №2 (16), p. 56-62. (in Russian)
9. Isaev D.V. (2012) *Monitoring i planirovanie razvitija sistem informacionnoj podderzhki korporativnogo upravlenija i strategicheskogo menedzhmenta* [Monitoring and Planning of Development of Information Support Systems for Corporate Governance and Strategic Management]. *Business Informatics*, no. 3 (21), p. 63-69. (in Russian)
10. Carnegie Mellon University (1993) *Capability Maturity Model for Software, Version 1.1. Technical Report. CMU/SEI-93-TR-024; ESC-TR-93-177. February 1993*. Software Engineering Institute, Carnegie Mellon University.
11. IT Governance Institute (2007) COBIT 4.1. IT Governance Institute.
12. ISACA (2012) *COBIT 5: A Business Framework for the Governance and Management of Enterprise IT*. ISACA.

13. Eckerson W.W. (2007) *Beyond the Basics: Accelerating BI Maturity*, Renton, WA: TDWI.
14. Halper F., Krishnan K. (2013) *TDWI Bid Data Maturity Model Guide: Interpreting Your Assessment Score*, Renton, WA: TDWI.
15. Weill P., Ross J. (2004) *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results*, Boston: Harvard School Press.
16. Ross J., Weill P., Robertson D. (2006) *Enterprise Architecture as Strategy: Creating a Foundation for Business Execution*, Boston: Harvard School Press.
17. National Association of State Chief Information Officers (2003) *NASCIO Enterprise Architecture Maturity Model. Version 1.3. December 2003*, Lexington, KY: NASCIO.
18. Institute for Enterprise Architecture Developments (2006) *Extended Enterprise Architecture Maturity Model Support Guide. Version 2.0* / Editorial writer: J.Schekkerman, Amersfoort: IFEAD.
19. Luftman J., Kempaiah R. (2007) An Update on Business-IT Alignment: A Line Has Been Drawn. *MIS Quarterly Executive*, vol. 6, no. 3 (September 2007), p. 165-177.
20. Perko J. (2008) *IT Governance and Enterprise Architecture as Prerequisites for Assimilation of Service-Oriented Architecture: An Empirical Study of Large Finnish Companies*. Thesis for the Degree of Doctor of Technology. Tampere: Tampere University of Technology.
21. Aho M. (2009) A Capability Maturity Model for Corporate Performance Management – An Empirical Study in Large Finnish Manufacturing Companies. Proceedings of *EBRF 2009 – Research Forum to Understand Business in Knowledge Society*. Jyväskylä, Finland. September 23-25, 2009.
22. Aho M. (2012) What is your PMI? A Model for Assessing the Maturity of Performance Management in Organizations. Proceedings of *'Performance Management: From Strategy to Delivery' (PMA 2012) conference*. University of Cambridge, UK. July 11-13, 2012.
23. Podinovski V.V., Potapov M.A. (2013) Metod vzveshennoj summy kriteriev v analize mnogokriterial'nyh reshenij: Pro et contra [Weighted Sum Method in The Analysis of Multicriterial Decisions: Pro et Contra]. *Business Informatics*, no. 3 (25), p. 41-48. (in Russian)
24. Isaev D.V. (2013) Proektirovanie sistem informacionnoj podderzhki korporativnogo upravlenija i strategicheskogo menedzhmenta [Design of Corporate Governance and Strategic Management Information Support Systems]. *Audit and Financial Analysis*, no. 6, p. 329-334. (in Russian)
25. Isaev D. (2012) Performance Management Information Support System: A Conceptual Model. *European Journal of Economics, Finance and Administrative Sciences*, issue 52, p. 6-20.

АВТОМАТИЗАЦИЯ ПРОЕКТИРОВАНИЯ ТЕХНОЛОГИЧЕСКИХ ДОРОЖНЫХ КАРТ. СБОР ИНФОРМАЦИИ И КОНСОЛИДАЦИЯ ЭКСПЕРТНЫХ МНЕНИЙ

К.В. Нагаев,

*старший научный сотрудник центра информационно-аналитических систем,
Институт статистических исследований и экономики знаний,
Национальный исследовательский университет «Высшая школа экономики»*

Е.М. Курбатова,

*младший научный сотрудник центра информационно-аналитических систем,
Институт статистических исследований и экономики знаний,
Национальный исследовательский университет «Высшая школа экономики»*

Адрес: 101000, г. Москва, ул. Мясницкая, д. 20

E-mail: knagaev@hse.ru, ekurbatova@hse.ru

В статье обсуждаются вопросы сбора, обработки и оценки экспертных сведений для заданной предметной области. Предлагаются методы обработки экспертных сведений с применением автоматизированных вычислительных средств. Полученные результаты используются в проектах форсайтных исследований, долгосрочного прогнозирования и разработки технологических дорожных карт. Основным преимуществом предложенных методов является учет степени компетенции экспертов, а также адаптация массивов экспертных сведений к аспектам предметной области, соответствующих сфере интересов заказчика технологической дорожной карты.

Рассматриваемый метод позволяет получать срез мнений экспертов для определенной версии опроса, либо актуальные (последние высказанные) мнения экспертов для заданного момента времени. Дополнительным фактором гибкости в модели представления экспертных знаний служат метаданные, характеризующие компетенции экспертов в различных аспектах: наука, технологии, бизнес и государственное управление. Метаданные позволяют обеспечить расчет различных аналитических индикаторов форсайтных исследований, например, слабых сигналов или форс-мажорных факторов. С целью поиска слабых сигналов метод выделяет оценки высококомпетентного эксперта, которые существенно не совпадают с мнением большинства. Форс-мажорные факторы определяются на основе оценки вероятности реализации элемента дорожной карты с высоким коэффициентом влияния на предметную область.

Другой областью применения методов обработки экспертных сведений с использованием метаданных является построение различных прогнозов развития предметной области с учетом значимости атрибутов в различных контекстах (экономическом, политическом, социальном, экологическом и технологическом). Совместный анализ этих сведений с данными о компетентности экспертов позволяет автоматически генерировать уточненные целевые анкеты для различных групп экспертов.

Предложенные методы реализованы в программном комплексе «Интерактивная дорожная карта с обратной связью» с апробированием на двух пилотных проектах: «Каталитический крекинг» и «Биотехнологии в медицине».

Ключевые слова: экспертный опрос, анкетирование, консолидация данных, формализация знаний, технологическая дорожная карта, форсайт.

Введение

На современном этапе для принятия решений и долгосрочного прогнозирования разрабатываются различные новые подходы анализа массивов информации, получаемых при опросе экспертных групп. Значительное внимание уделяется вопросам обработки экспертных сведений при проведении форсайт-исследований (от английского Foresight – предвидение) [1]. Одним из наиболее перспективных инструментов планирования и достижения качественно новых результатов форсайт-исследований является технология дорожного картирования (составления дорожных карт). В результате анализа данных в форсайт-проектах создаются дорожные карты, представляющие собой генеральный план развития исследуемой области на заданном временном горизонте [2]. Дорожные карты дают возможность получить наглядное представление о перспективах развития предметной области с целью принятия важных решений, как в области технологий, так и в стратегическом управлении бизнесом и государственной политике, и являются одним из наиболее удобных инструментов долгосрочного стратегического планирования [3].

При реализации программного комплекса «Интерактивная дорожная карта с обратной связью» были проведены исследования процесса подготовки данных и расчета технологических дорожных карт и разработан оригинальный метод сбора, хранения и обработки информации. В настоящей работе обсуждаются организационные и технические аспекты сбора и консолидации данных при построении технологических дорожных карт (ТДК), одного из наиболее информационно насыщенных видов дорожных карт.

1. Постановка задачи

Создание базы знаний по предметной области для составления технологической дорожной карты является сложным процессом, так как требует обработки и анализа большого объема разнородной информации, ее формализации в соответствии с моделью предметной области дорожной карты и преобразования в структуры данных, соответствующих аспектам. Один из подходов при формализации предметной области предусматривает задание метаописания в виде атрибутированных объектов, отношений между объектами (при необходимости, также атрибутов отношений). Однотипные объекты объединяются в классы, что определяет требуемую структуру

ТДК. Наполнение этой структуры необходимыми типами объектов и отношений между ними предполагает решение следующих подзадач: формализация предметной области, формирование экспертных групп (поиск экспертов и оценка их компетенции), формирование анкет, обеспечивающих определение структуры классов, анкетирование и обработка результатов анкетирования. Результатом выполнения этих задач является заполненная модель ТДК, обеспечивающая проведение аналитических исследований. Целью данной статьи является описание метода формализации предметной области, сбора и консолидации экспертных мнений.

При применении большинства классических методов консолидации экспертных сведений существует проблема потери важной информации вследствие того, что мнения меньшинства экспертов отсекаются без учета их компетентности. Описываемый метод предполагает сохранение всей полученной от экспертов информации с добавлением метаданных, описывающих качественные характеристики сведений.

Для получения достоверных данных классические методы анкетирования, как правило, используют итеративный подход, где на каждой итерации идет повторный опрос группы экспертов постоянного состава с уменьшением количества вариантов ответов. Для повышения достоверности данных используются многочисленные группы экспертов, что, стоит отметить, ведет к значительным материальным и временным затратам при большом количестве итераций анкетирования. В целях снижения влияния этой проблемы был разработан адаптивный метод составления уточненной анкеты для опроса меньшего числа экспертов.

2. Современные методы сбора и обработки информации

При прогнозировании и принятии решений применяется широкий спектр методов сбора исходных данных, в их числе синектика, мозговой штурм, проведение фокус-групп, анкетный опрос, формализованное интервью и многие другие [4]. Особенностью долгосрочного планирования является значительная неопределенность будущего, что при построении дорожных карт требует для повышения объективности обработки большого массива начальных данных, полученных от представительной группы экспертов. Как правило, эти данные подготавливаются и консолидируются с применением специализированных

также в соответствии с межгрупповым взаимодействием, и формировать списки экспертов с распределением компетентности.

3. Использование мультиатрибутивного подхода при формализации данных

Для получения исходных данных при построении ТДК требуется четкая формализация задачи, при этом наилучшие результаты дает представление данных в виде объектов, характеризующихся набором атрибутов различной природы. Данный метод позволяет сформулировать вопросы анкеты таким образом, чтобы получить ответы, пригодные для машинной обработки. В целях повышения управляемости модели ТДК объекты, принадлежащие одному слою (например, НИР, технологии, продукты, рынки) и обладающие одним и тем же набором атрибутов, объединяются в класс [8]. Такая классификация позволяет структурировать свойства объектов.

Основной проблемой методов экспертного анализа является декомпозиция сложных абстрактных вопросов на последовательность простых задач, доступных для понимания и адекватного восприятия экспертом [9]. Мультиатрибутивный подход, широко применяемый, например, в маркетинге при измерении полезности того или иного товара [10], позволяет декомпозировать такие задачи на совокупность элементов и их атрибутов.

Однако, при формализации характеристик объектов ТДК в виде атрибутов разработчиком ТДК возможно появление проблем, связанных с ошибками декомпозиции. Такие ошибки основаны, прежде всего, на неполном понимании разработчиком ТДК всех возможных параметров исследуемых объектов на начальных этапах работы. Выявление проблем, связанных с ошибками декомпозиции, обеспечивается в предлагаемом методе при помощи оценки распределения экспертных мнений. В качестве показателя корректности формализации конкретного параметра объекта ТДК предлагается рассматривать унимодальное распределение мнений экспертов. Получение унимодального распределения показывает, что большинство экспертов понимают значение параметра одинаково, и дают сходные оценки объекта по данному показателю. Би- или мульти-модальное распределение указывает на ошибку декомпозиции. В качестве одной из наиболее вероятных причин такой ошибки нами предлагается рассматривать смещение в одном

атрибуте нескольких свойств объекта. В таком случае разработчику ТДК рекомендуется декомпозировать этот атрибут на последовательность более детальных характеристик объекта.

Различные свойства объекта могут быть описаны с помощью количественных либо качественных характеристик. Для облегчения работы экспертов количественные характеристики объектов описываются с помощью метрических шкал, где значения атрибутов вводятся в виде диапазона чисел, а не определенного числа. Значения, указываемые с использованием диапазонов шкал, имеющих однородную структуру единиц измерения, могут сравниваться с определенными погрешностями, которые задаются на этапе определения шкалы.

Качественные характеристики определяются с помощью номинальных и ранговых шкал, и эксперт при заполнении анкеты должен выбрать один из предложенных ему вариантов значения атрибута. Примером номинальной шкалы может служить шкала цветов. Нельзя сказать, что синий цвет лучше зеленого, а красный находится между ними по значимости. В отличие от номинальных, в ранговых шкалах определен порядок следования вариантов (например, низкий – средний – высокий), хотя вне конкретного контекста использования не всегда можно однозначно определить, какой из вариантов является наилучшим или наихудшим.

В предлагаемом методе атрибуты также характеризуются значимостью по аспектам (политический, экономический, социальный, технологический, экологический). Эта значимость определяется с помощью иерархического подхода [11] и сохраняется в базе знаний в виде метаданных.

Все данные, полученные в ходе одного этапа анкетирования экспертов, считаются версией опроса. Метод предполагает хранение всех сведений, полученных от экспертов, с обязательным указанием версии опроса. Один и тот же эксперт на разных итерациях анкетирования может дать различные варианты ответа на вопрос, поскольку его мнение будет меняться по мере накопления знаний, повышения квалификации и появления новых сведений по предметной области. Например, при составлении дорожной карты эксперты предполагали, что существует вероятность появления инновационной технологии со значительным потенциалом применения. По прошествии полугода группа ученых совершает прорыв, и данная технология становится реальностью. В этом случае в новой версии

экспертных мнений эта технология будет фигурировать уже как реально существующая, что окажет огромное влияние на прогноз развития предметной области. Кроме того, сведения о вновь созданной технологии будут конкретизированы, что приведет к изменению параметров элементов дорожной карты, связанных с ее появлением.

В различных версиях опросов состав принимающих участие экспертов может оказаться различным. Рассматриваемый метод позволяет получать срез мнений экспертов на момент определенной версии опроса (мнения экспертов, принявших участие в опросе), либо актуальные (последние высказанные) мнения экспертов на любой заданный момент времени. Мнения экспертов будут консолидированы и представлены в виде элементов ТДК с конкретными значениями параметров и связей между ними.

4. Консолидация данных, полученных от экспертов

Построение дорожных карт представляет собой адаптивный процесс, предусматривающий учет интересов заказчика ТДК в тех или иных областях планирования. Объекты дорожной карты могут рассматриваться с разных точек зрения. Большое влияние на результат разработки дорожных карт оказывает соотношение интересов заказчика ТДК в необходимых сферах планирования. Данные интересы представлены в предлагаемом методе в виде метаданных, характеризующих прогнозируемые процессы развития НИР и разработки технологий в различных аспектах: экономическом, политическом, социальном, экологическом и технологическом. Эти аспекты участвуют в определении значимости различных атрибутов элементов ТДК для

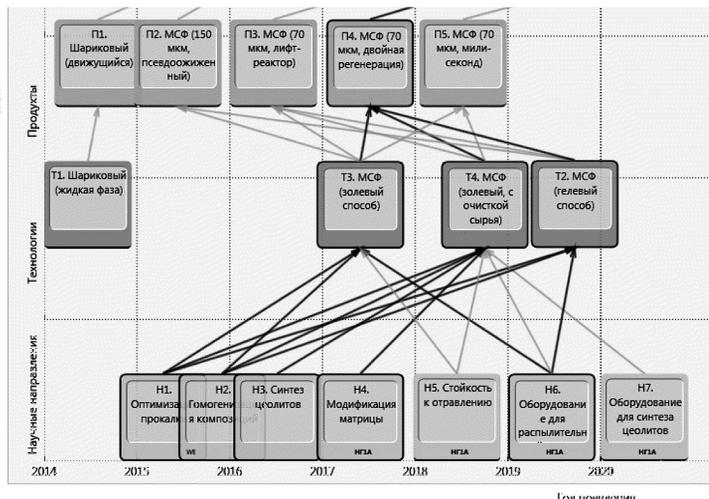
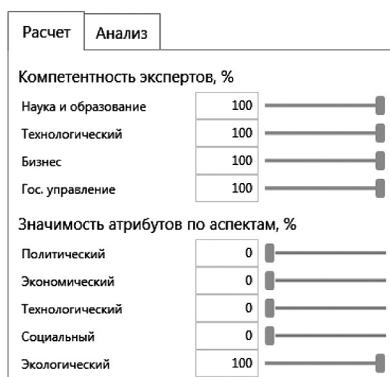
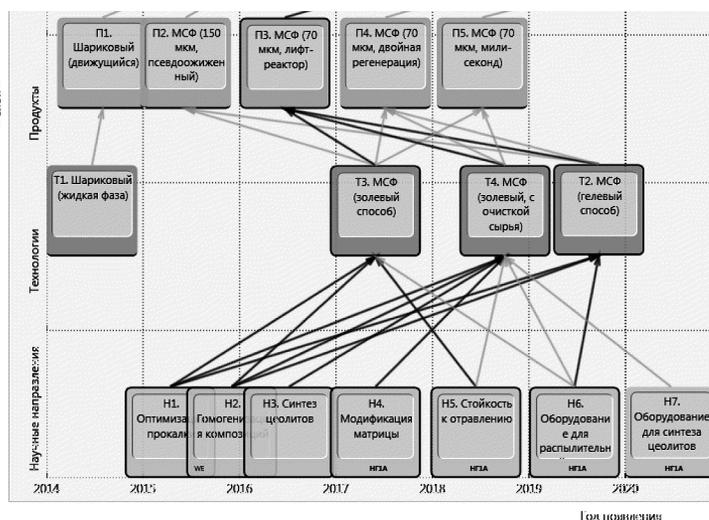
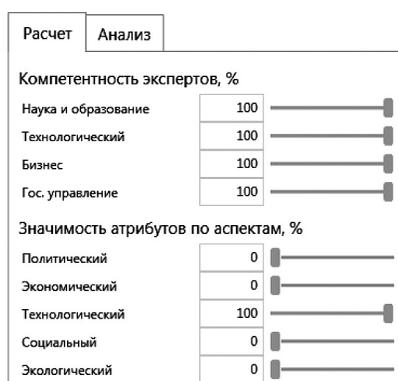


Рис. 1. Рассмотрение ТДК с точки зрения технологичности и экологии

вычисления их применимости в предметной области, как будет показано ниже.

Хранение в базе знаний всех мнений, полученных от экспертов, позволяет получать различные прогнозы развития предметной области в зависимости от условий, накладываемых с помощью метаданных. Например, разработчику ТДК может понадобиться оценить перспективы развития отрасли с точки зрения экологии. В этом случае понадобится повысить значимость экологического аспекта по сравнению с остальными в настройках расчета дорожной карты. На рис. 1 можно наблюдать, что при рассмотрении дорожной карты с учетом экологического аспекта элемент НИР «Н5. Стойкость к отравлению» исключается из основной траектории (показана синими стрелками) как обладающий низкой экологичностью и заменяется на «Н6. Оборудование для распылительной сушки».

Технологические дорожные карты в своем базовом представлении обычно имеют четыре основных слоя: рынки, продукты, технологии и НИР [12]. В программном комплексе «Интерактивная дорожная карта с обратной связью» предусмотрен также особый тип слоя «Глобальный контекст развития» (ГКР). Элементы этого слоя могут оказывать влияние на элементы других слоев, увеличивая либо уменьшая значение применимости атрибута по различным аспектам, если затронутый влиянием ГКР элемент удовлетворяет фильтру, описанному в объекте ГКР. Множитель для функции применимости элемента, учитывающий различные аспекты, определяется по формуле:

$$a = \frac{\sum_{k=1}^q (A_k \cdot P_k \cdot G_k)}{\sum_{k=1}^q A_k \cdot \sum_{k=1}^q P_k \cdot \sum_{k=1}^q G_k} \cdot q, \quad (1)$$

q — количество учитываемых аспектов (в данный момент их пять — технологический, экономический, политический, социальный и экологический);

A_k — значимость атрибута по аспекту k (от 0 до 1);

P_k — желаемый коэффициент учета аспекта k , задается при расчете ДК (от 0 до 1);

G_k — коэффициент влияния объекта ГКР по аспекту k (от 0 до ∞).

Множитель влияет на полученный при расчете частный показатель применимости атрибута, затем на суммарный показатель применимости элемента,

и, в конечном итоге, на построение основной траектории и на общую картину ТДК.

Метаданные также включают в себя желаемые характеристики компетенции экспертов, чьи мнения являются определяющими для принятия решений по заданной тематике. Эксперты характеризуются, прежде всего, степенью компетенции в предметной области ТДК, а также компетентностью по следующим направлениям: наука и образование, технологии, бизнес, государственное управление. В предлагаемом методе степень компетенции экспертов оценивается с помощью метода скрытых коллективов, а также индекса цитируемости [7].

При консолидации экспертных сведений вычисляются средневзвешенные значения характеристик элементов ТДК:

$$\bar{x} = \frac{\sum_{i=1}^n w_i \cdot x_i}{\sum_{i=1}^n w_i} \quad (2)$$

Вес эксперта w_i в данной формуле определяется с учетом его компетентности в предметной области и по различным аспектам при помощи инъекции введенных в систему метаданных:

$$w_i = C_i \cdot \frac{\sum_{k=1}^m C_{ik} \cdot K_k}{m} \quad (3)$$

C_i — степень компетенции эксперта в предметной области ТДК;

C_{ik} — степень компетенции эксперта по аспекту k ;

K_k — желаемый коэффициент учета компетенции по аспекту k , задается при расчете ДК;

m — количество учитываемых аспектов.

В программном комплексе «Интерактивная дорожная карта с обратной связью» компетентность эксперта в предметной области задается в виде безразмерной величины, исходя из базовой компетенции, равной 1 (рядовой специалист). Для эксперта высокого уровня может быть установлен коэффициент компетенции, существенно превышающий единицу. Практика показывает, что ограничение таких величин как компетенция экспертов приводит к искусственному уплотнению верхних значений интервала, либо необходимости повторной нормировки в случае необходимости добавления значений, превышающих верхнюю границу.

Кроме степени компетенции в предметной области ПК учитывает также общую компетент-

Выявление слабых сигналов

Х атрибут	Год появления	элемент	Год появления	X	WE	~ Слои ~
Y атрибут	~ Слои ~	B1. Евро-5	2014	2014	<input type="checkbox"/>	Глобальный контекст развития
Версия опроса	Уточненная анкета эксперта	B2. Отсталая структура производства	2014	2014	<input type="checkbox"/>	Глобальный контекст развития
Компетентность экспертов, %		N1. Оптимизация прокатки	2015	2015	<input type="checkbox"/>	Научные направления
Наука и образование	100	N2. Гомогенизация композиций	2015	2015	<input checked="" type="checkbox"/>	Научные направления
Технологический	100	N3. Синтез цеолитов	2016	2016	<input type="checkbox"/>	2 016 — 1 experts вления
Бизнес	100	N4. Модификация матрицы	2017	2017	<input type="checkbox"/>	Научные направления
Гос. управление	100	N5. Стойкость к отравлению	2018	2018	<input type="checkbox"/>	Научные направления
Выявление слабых сигналов		N6. Оборудование для распылители	2019	2019	<input type="checkbox"/>	Научные направления
Компетенция, %	91	N7. Оборудование для синтеза цеол	2020	2020	<input type="checkbox"/>	Научные направления
Чувствительность X	1	П1. Шариковый (движущийся)	2014	2014	<input type="checkbox"/>	Продукты
Чувствительность Y	1	П2. МСФ (150 мкм, псевдооживленнь	2015	2015	<input type="checkbox"/>	Продукты
		П3. МСФ (70 мкм, лифт-реактор)	2016	2016	<input type="checkbox"/>	Продукты
		П4. МСФ (70 мкм, двойная регенера	2017	2017	<input type="checkbox"/>	Продукты

Рис. 2. Выявление слабых сигналов

ность эксперта по четырем аспектам: наука, технологии, бизнес и государственное управление. Комплекс позволяет в интерактивном режиме повышать и понижать значимость того или иного аспекта компетенции при расчете, в результате разработчик ТДК получает консолидированные значения параметров элементов с учетом мнений экспертов, компетентных в той или иной области деятельности.

Метод также позволяет обратить внимание разработчика ТДК на особые мнения высококомпетентных экспертов путем выявления слабых сигналов. Слабый сигнал определяется как значительное отклонение мнения высококомпетентного эксперта от мнений большинства. Для выявления слабых сигналов разработчик ДК задает степень компетенции искомым экспертам в предметной области (в процентах от наибольшей компетенции), и величины порогов в единицах измерения вычисляемых атрибутов. Пример слабого сигнала можно наблюдать на рис. 2. В данном случае мнение одного эксперта, степень компетенции которого выше 91% от максимального уровня компетенции экспертов, участвовавших в опросе, отличается от мнения большинства. Большинство экспертов прогнозируют появление результатов НИР «Н2. Гомогенизация композиций» в 2015 году, однако, высококомпетентный эксперт считает, что результаты будут только в 2016. Таким образом, по атрибуту «Год появления» система указывает на наличие слабого сигнала (колонок WE).

Также предусмотрена возможность оценки форс-мажорных факторов, то есть появления в будущем высокоэффективных технологий и раз-

работок, вероятность создания которых мала. Для оценки форс-мажорных факторов необходимо указать год появления, максимальную величину вероятности появления, а также минимальную применимость искомым элементам. Пример оценки форс-мажорных факторов показан на Рис.3. Здесь оценивается вероятность появления до 2016 года результатов НИР, у которых значение применимости выше 10. Система указывает (колонок WI) на форс-мажорный фактор, поскольку вероятность получения результатов НИР «Н4. Модификация матрицы» до 2016 года ниже 0.5.

В отличие от применяемых в настоящее время методов прогнозирования, предлагается уделять слабым сигналам и форс-мажорным факторам особое внимание, а именно: проводить повторный опрос экспертов с целью дополнительного исследования таких объектов.

5. Адаптивный метод составления уточненной анкеты

Метаданные, хранящиеся в базе знаний, содержат информацию о степени компетенции эксперта в предметной области ТДК и о его компетенции в различных областях деятельности. Также в метаданных содержится информация о значимости каждого атрибута по различным аспектам. Анализ этих данных позволяет автоматически генерировать уточненные анкеты для различных групп экспертов. Уточненные анкеты составляются с учетом компетенции целевой группы экспертов, что позволяет сделать последующие этапы анкетирования менее трудоемкими и затратными (анкеты становятся ко-

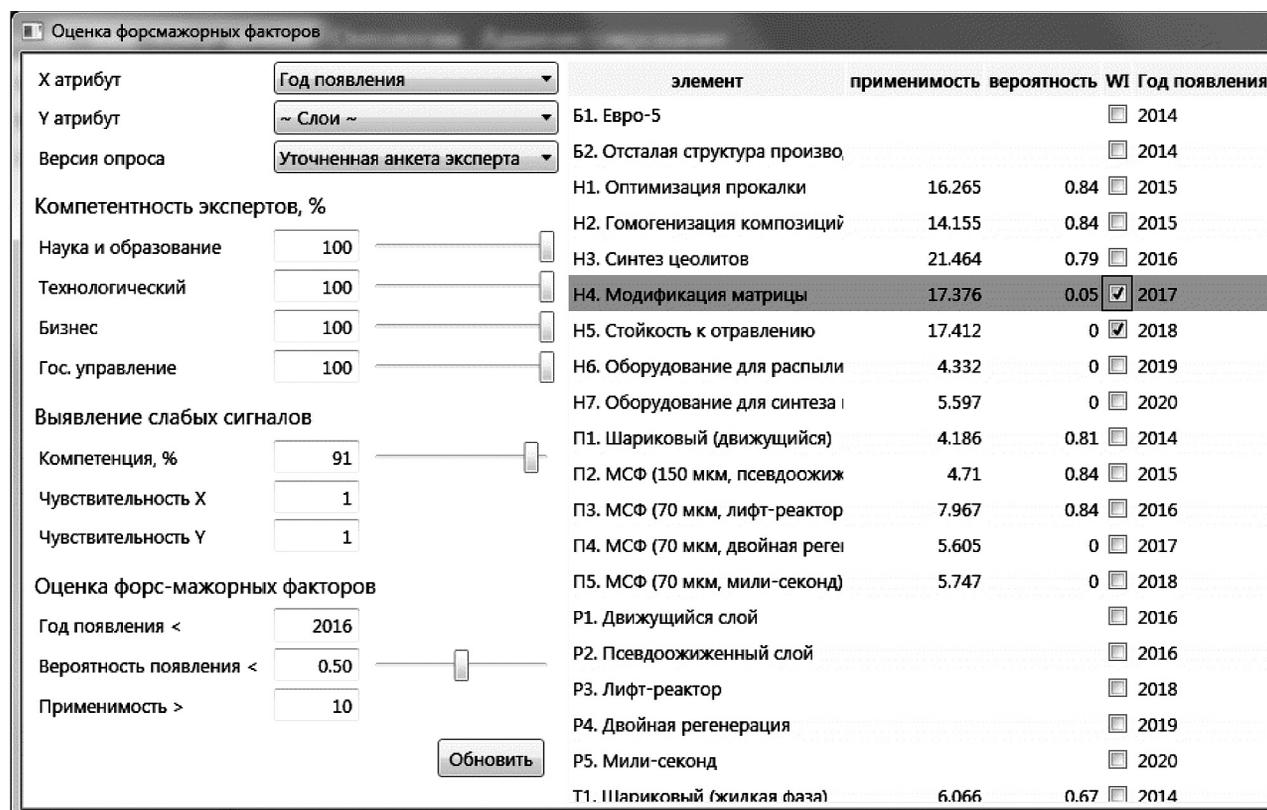


Рис. 3. Оценка форс-мажорных факторов

роче и эффективнее). Экспертам предоставляются вопросы, ориентированные на их область интересов. Это существенно облегчает работу, а также улучшает качество получаемых сведений.

Например, для экспертов в области науки и технологий (у которых компетентность в бизнесе и государственном управлении менее 10%) в анкету включаются только атрибуты объектов ТДК, связанные с технологическим аспектом, и не включаются экономические, политические или социальные. Кроме того, при расчете элементов дорожной карты их отбор из большого множества, получаемого при первичном опросе экспертов, осуществляется с учетом настроек проекта. В результате такого расчета получается набор элементов и связей между ними, отфильтрованный с учетом соотношения аспектов, указанных в постановке задачи на создание дорожной карты. Этот фильтр также оказывает влияние на формирование уточненной анкеты.

Метод адаптивного составления анкеты позволяет на основе результатов, полученных из уточненных анкет, обновлять имеющуюся базу знаний, и, исходя из новых данных, получать другую модификацию уточненной анкеты для прежних

условий выборки. Дополнительным преимуществом разработанного метода является возможность использования данных от любой итерации опроса, благодаря версионной организации базы знаний. Каждая версия данных соответствует определенному опросу экспертов. Такая структура данных позволяет отслеживать динамику изменения представлений экспертов о предметной области во времени. Например, на базовой графической схеме ТДК при задании различных версий опросов экспертов можно увидеть, как перемещаются элементы дорожной карты по временной оси, изменяются связи между ними, и перестраивается основная технологическая траектория, что отражает различные стороны использованного множества экспертных заключений.

Заключение

В статье предложены методы сбора и хранения данных для создания базы знаний при составлении технологической дорожной карты для заданной предметной области. Описан метод консолидации данных с учетом значимости атрибутов объектов по различным аспектам деятельности, а также степени компетентности опрашиваемых

экспертов. Метод позволяет анализировать получаемые от экспертов сведения с формированием отчетов, представляющих собой прогноз или план развития предметной области с учетом интересующих разработчика ТДК аспектов (экономический, социальный и т.д.) Также в работе представлен адаптивный метод составления анкет, позволяющий эффективно производить анкетирование экспертов в несколько этапов.

Данные методы реализованы в программном комплексе «Интерактивная дорожная карта с обратной связью», предназначенном для построения дорожных карт и принятия решений на основе экспертных сведений. Комплекс позволяет создавать базы знаний, содержащие мнения экспертов и метаданные, производить консолидацию данных и расчет технологических траекторий и частных

показателей элементов ТДК с учетом различных аспектов заданной тематики (например, с точки зрения экономики или государственного управления) [8]. Данный комплекс нацелен на управления и департаменты стратегического развития, инновационную деятельность министерств, ведомств, государственных корпораций, коммерческих корпораций, концернов, холдингов и крупных компаний. Комплекс является универсальным и может применяться при долгосрочном прогнозировании развития любой предметной области.

Описанная технология была апробирована на двух пилотных проектах: «Каталитический крекинг» и «Биотехнологии в медицине». Полученный опыт показал практическую применимость этой методологии для создания сложных моделей в широком спектре областей применений. ■

Литература

1. Соколов А.В. Форсайт: взгляд в будущее // Форсайт. 2007. № 1. С. 8-15.
2. Инвестиционная деятельность Роснано – дорожные карты // Роснано [Электронный ресурс]: <http://www.rusnano.com/investment/roadmap/oil> (дата обращения: 08.08.2013)
3. Farrukh C., Phaal R., Probert D. Technology Roadmapping: Linking Technology Resources into Business Planning // International Journal of Technology Management. 2003. No. 26. P. 2-19.
4. Орлов А.И. Организационно-экономическое моделирование: В 3 ч. Ч.2. Экспертные оценки. М.: Изд. МГТУ им. Н. Э. Баумана, 2011.
5. Поспелов Г.С., Ириков В.А. Программно-целевое планирование и управление. М.: Советское радио, 1976.
6. Rowe G., Wright G. The Delphi Technique as a Forecasting Tool: Issues and Analysis // International Journal of Forecasting. 1999. Vol. 15. Issue 4.
7. Хорошевский В.Ф. Извлечение информации из текстов на конференциях серии «ДИАЛОГ: взгляд со стороны по лестничной клетке» // Труды международной конференции «Диалог 2010». М.: Наука, 2010.
8. Ена О.В., Нагаев К.В. Автоматизация процессов разработки технологических дорожных карт. Расчет интегральных показателей применимости // Бизнес-информатика. 2013. №3 (25). С. 56-62.
9. Ременников В.Б. Управленческие решения: Учебное пособие для вузов. М.: ЮНИТИ-ДАНА, 2005.
10. Wittink D., Walsh J. Conjoint Analysis: Its Reliability, Validity and Usefulness. Paper presented at Sawtooth Conference Proceedings. 1988.
11. Saati T.L. Decision Making with Dependence and Feedback: The Analytic Network Process. Rws Publications, 2001.
12. Группа «Конструирование будущего»: Создание дорожных карт. Санкт-Петербургская региональная общественная организация работников науки и культуры: «Энциклопедия» // [Электронный ресурс]: <http://future-designing.org/biblioteka/logyroadmappingaguideforgovernmentemployees.html> (дата обращения: 08.08.2013).

TECHNOLOGY ROADMAPPING. METHOD FOR GATHERING AND CONSOLIDATION EXPERT OPINIONS

Konstantin NAGAEV,

Senior Researcher, Center for Information Analysis Applications, Institute for Statistical Studies and Economics of Knowledge, National Research University Higher School of Economics

Elena KURBATOVA,

Junior Researcher, Center for Information Analysis Applications, Institute for Statistical Studies and Economics of Knowledge, National Research University Higher School of Economics

Address: 20, Myasnitskaya str., Moscow, 101000, Russian Federation

E-mail: knagaev@hse.ru, ekurbatova@hse.ru

The paper discusses the problems related to data collection, processing and expert opinions evaluating in terms of a specified knowledge domain. Methods for expert data processing based on computer systems are presented. The results are used in foresight projects, long-term forecasting and in technology road mapping. The main advantage of proposed methods is the consideration of expert competence degrees. Additionally, expert datasets are adapted to aspects of the knowledge domain in accordance with a customer's interest area.

The proposed method provides obtaining subsets of expert data for a required version of expert polls, or expert data values effective on a specified date. An additional driver of expert data model flexibility is metadata for expert competence identification in different aspects: science, technology, business and governance. The metadata is designed to run evaluation of various analytic indicators for foresight investigations, for example, for weak signals

and wild cards. The method gives accent to the opinion of a high-competence expert which may diverge significantly from the major opinion. This very property of the method is extremely important for weak signal retrieval. Wild cards are founded with an estimation of realization probability of a roadmap element with high influence rate in an interest area.

Another area of the method applications is the development of different forecasts of knowledge domain's evolution taking into account the importance of properties in special contexts (economic, political, ecological and technological). Joint analysis of these data and experts' competence data allows generating improved specific questionnaires for distinct expert groups automatically.

The methods described here have been implemented in the software system «Interactive roadmap with feedback link» and tested in two pilot projects: «Catalytic cracking» and «Biotechnologies medical».

Key words: expert survey, data gathering, data consolidation, knowledge formalization, roadmapping, foresight.

References

1. Sokolov A.V. (2007) Forsajt: vzgljad v budushhee [Foresight: Look into the Future]. *Foresight*, no 1, pp. 8-15. (in Russian)
2. *Investitsionnaya deyatel'nost' Rosnano – dorozhnye karty* [Rosnano Investment Activity – Roadmaps] Available at: <http://www.rusnano.com/investment/roadmap/oil> (accessed 8 August 2013). (in Russian)
3. Farrukh C., Phaal R., Probert D. (2003) Technology Roadmapping: Linking Technology Resources into Business Planning. *International Journal of Technology Management*, no. 26, pp. 2-19.
4. Orlov A.I. (2011) *Organizatsionno-ekonomicheskoe modelirovanie. Ekspertnye otsenki* [Business modelling. Expert evaluation]. Moscow: Bauman Moscow State Technical University. (in Russian)
5. Pospelov G.S., Irikov V.A. (1976) *Programmno-celevoe planirovanie i upravlenie* [Programmed-objective planning and control]. Moscow: Sovetskoye Radio. (in Russian)
6. Rowe G., Wright G. (1999) The Delphi Technique as a Forecasting Tool: Issues and Analysis. *International Journal of Forecasting*, vol. 15, issue 4.
7. Horoshevkiy V.F. (2010) *Izlyechenie informacii iz tekstov na konferencijah serii «DIALOG: vzgljad soseda po lestnichnoj kletke»* [Information Extraction from texts at conferences «DIALOG: Point of View of Housemate»]. Moscow: Nauka. (in Russian)
8. Ena O.V., Nagaev K.V. (2013) Avtomatizacija processov razrabotki tehnologicheskikh dorozhnykh kart. Raschet integral'nykh pokazatelej primenimosti [Automation of the Technology Roadmap Development. Calculation of the Integral Indicators of Applicability]. *Business Informatics*, no. 3(25), pp. 56–62. (in Russian)
9. Remennikov V.B. (2005) *Upravlencheskie reshenija: Uchebnoe posobie dlja vuzov* [Management Decisions: Study Guide for High School]. Moscow: UNITY-DANA. (in Russian)
10. Wittink D., Walsh J. (1988) *Conjoint Analysis: Its Reliability, Validity and Usefulness*. Paper presented at Sawtooth Conference Proceedings.
11. Saati T.L. (2001) *Decision Making with Dependence and Feedback: The Analytic Network Process*, Rws Publications.
12. *Gruppa «Konstruivovanie budushego»: Sozdanie dorozhnykh kart* [The Future Designing Group: Roadmapping]. Available at: <http://future-designing.org/biblioteka/yroadmappingaguideforgovernmentemployees.html> (accessed 8 August 2013).

МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ТОРГОВОГО ПРЕДПРИЯТИЯ

С.М. Ямпольский,

кандидат технических наук, доцент кафедры бизнес-аналитики,
факультет бизнес-информатики, Национальный исследовательский
университет «Высшая школа экономики»

Адрес: 101000, г. Москва, ул. Мясницкая, д. 20

А.С. Шаламов,

доктор технических наук, научный сотрудник отдела
статистических проблем информатики и управления,
Институт проблем информатики Российской академии наук

Адрес: 119333, г. Москва, ул. Вавилова, д. 44

E-mail: syampolsky@hse.ru, a-shal5@yandex.ru

В статье рассматриваются вопросы повышения эффективности управления оборотными средствами торгового предприятия на основе автоматизированного планирования. Применяемый подход дает возможность использования классических методов оптимизации динамических систем с целью определения основных параметров экономической политики торгового предприятия, обеспечивающей выбор наилучшей стратегии развития. С помощью предложенного подхода могут быть созданы алгоритмы, позволяющие оценивать как управляемые, так и неуправляемые риски, а также находить обоснованные варианты решений по их минимизации или снижению.

При изложении материала среда функционирования торгового предприятия описывается в традиционных терминах торгово-финансового рынка, что позволяет использовать ее везде, где используется информационная система «1С Предприятие», в виде дополнительного программного комплекса прогнозирования и оптимизации.

Разработанная математическая модель позволяет решать задачи, облегчающие руководству торгового предприятия выбор и обоснование различных решений, включая прогнозирование функций времени, определяющих математические ожидания процессов изменения активов и пассивов торгового предприятия, определение уставного капитала, обеспечивающего решение стратегических задач, определение начальной товарно-денежной политики, расчет эффективности вариантов вырабатываемых планов.

Результаты прогнозирования динамики состояния оборотных средств представляются в интегрированном графическом виде, дающем возможность увидеть целостную картину предстоящего состояния оборотных средств торговой организации, а также основных участников торгово-экономической деятельности на заданном интервале времени.

Ключевые слова: торговое предприятие, поставщики, оптовые и розничные покупатели, оборотные средства, автоматизированное планирование, управление заказами, дискретный марковский процесс с непрерывным временем.

Введение

Принципы математического моделирования торгового предприятия (ТП), изложенные в работе [1], базируются на марковских свойствах случайных процессов, порожденных пуассоновскими потоками. Эти потоки по своим свойствам наиболее близки к тем, посредством которых происходит обмен материальными и финансовыми ресурсами между ТП и экономическими агентами (ЭА).

В теории марковских процессов существует понятие «обобщенных» пуассоновских потоков [3]. В этом случае имеет место аксиоматика, допускающая нестационарность и неординарность потоков, при обязательном сохранении свойства отсутствия последствия, что и позволяет называть порожденные ими процессы марковскими.

Плотности вероятностей интервалов времени в пуассоновском потоке имеют максимальную энтропию среди других плотностей вероятностей при одинаковой дисперсии интервалов. Поэтому применение гипотезы о пуассоновости потока дает завышенные значения дисперсий порождаемых ими процессов, что перекрывает возможный недоучет каких-либо факторов (все другие возможные оценки находятся внутри получаемой в расчетах области) и тем самым способствует снижению неуправляемых рисков.

Практическая ценность предложенного подхода заключается в том, что с его помощью могут быть созданы алгоритмы ЭВМ, обеспечивающие, в частности, технико-экономическое обоснование контрактов, заключаемых между участниками торгово-финансовых процессов, обуславливающих взаимные обязательства, а также заданный уровень доходности ТП в процессе функционирования. Особенностью этих алгоритмов является то, что при правильном учете основных приводящих факторов они становятся уникальным инструментальным средством, позволяющим оценивать всевозможные риски, а также находить обоснованные варианты решений по их минимизации или снижению.

Все это предоставляет новые возможности по управлению ТП, в том числе по сложным экономическим критериям, обеспечивающим выполнение *главного требования* – минимум затрат ресурсов при гарантированном обеспечении заданных (максимальных) показателей доходности.

В работе [1] с целью изложения основных принципов предлагаемого подхода читателю была представлена упрощенная модель процессов между ТП и его ЭА. В ней были опущены экономические

отношения, например, между ТП и возможными заказчиками, поставщиками, собственниками, а также инвесторами, предоставляющими заемные финансовые средства и др. Модель представлялась в виде графа состояний оборотных средств (ОС) без конкретной математической постановки задачи.

В настоящей работе предлагается рассмотреть несколько более детальную модель за счет включения в нее экономических агентов в лице поставщиков товарной продукции. Ставится задача создания алгоритма, с помощью которого можно было бы рассчитать основные параметры деятельности ТП и его партнеров по рынку, обеспечивающие максимальную выручку от продажи товаров на достаточно коротком промежутке времени. Это позволит получить предварительные заключения об эффективности использования ОС в целях получения прибыли, достаточной для жизнедеятельности предприятия.

С этой целью формулируется математическая постановка задачи моделирования динамики ОС и предлагаются основные соотношения в виде дифференциальных уравнений для математических ожиданий и ковариаций процессов ТП. Решения этих уравнений при определенных параметрах характеризуют прогнозируемое изменение активов и пассивов ТП при его функционировании в течение заданного промежутка времени. В конечном случае, оцениваются показатели доходности ТП с определением их нижних и верхних границ.

Следует отметить, что на продолжительных промежутках времени деятельности ТП приведенный перечень ЭА следует еще более расширить. Кроме того, придется учитывать, помимо оборотных, еще и движение основных средств. Это необходимо, например, для выработки долговременной политики поведения на товарно-денежном рынке.

Применяемый подход дает возможность использования в дальнейшем классических методов оптимизации динамических систем с целью определения основных параметров экономической политики ТП, обеспечивающей выбор наилучшей стратегии развития.

Особенностью работы является то, что при изложении материала среда функционирования ТП описывается в традиционных терминах торгово-финансового рынка, что позволяет расширить понятие «торговое предприятие» и тем самым получить универсальную модель процессов. Это означает, что ее можно использовать с определенными доработками везде, где используется информационная система «1С Предприятие», в виде дополни-

тельного программного комплекса прогнозирования и оптимизации.

В конечном счете, предлагаемые подходы могут существенно облегчить решение главных задач, стоящих перед руководством ТП, например:

♦ минимизация риска оказаться в ситуациях, связанных с неплатежеспособностью («программа-минимум»);

♦ принятие и реализация политики обеспечения оптимальных показателей торгово-экономической деятельности предприятия в течение заданного периода времени;

♦ постоянное наращивание собственного капитала ТП в соответствии со стратегическими целями по развитию предприятия («программа-максимум») и др.

1. Граф состояний оборотных средств

Торговое предприятие (ТП) является одним из экономических агентов (ЭА) рынка финансов, товаров, работ и услуг и в процессе своей деятельности вступает с остальными в экономические отношения, определяемые характером движения оборотных и основных средств.

На кратковременных периодах деятельности ТП, где наиболее характерным является изменение состояний ОС, экономическими агентами являются:

- торговое предприятие (ТП) в лице головного офиса торговой сети (состояние 1);
- торговые партнеры (ТОП) – магазины, наделенные определенной самостоятельностью – реализующие товары через сеть продаж (состояние 3);
- поставщики товаров (ПТ) (состояние 0₄);
- покупатели товаров (П) (состояние 0₅).

Для учета потерь ОС, связанных с порчей товаров, списанием неликвидов, ошибками при планировании, недобросовестностью персонала и пр., обусловленными исключительно деятельностью торговых партнеров предприятия, вводится состояние 2.

Отметим, что вершины графа состояний оборотных средств соответствуют приведенному списку, т.к. ОС, в том или ином виде, при своем движении проходят через финансовые счета этих ЭА. Обозначения указанных вершин соответствуют количеству ОС, находящихся в этих состояниях: $y_1(t)$ соответствует состоянию 1, $y_2(t)$ – состоянию 2, $y_3(t)$ – состоянию 3.

На рис. 1 представлен граф рассматриваемой системы.

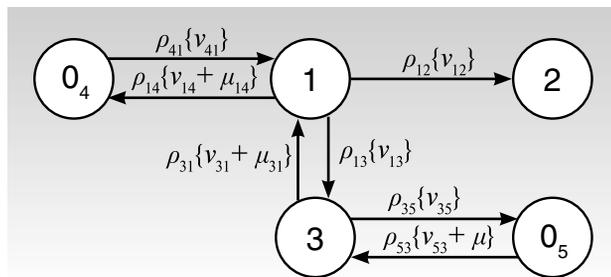


Рис. 1. Граф оборотных средств торгового предприятия

Через 0₄ обозначается поставщик товаров, а через 0₅ обозначается покупатель товаров, возможности которых по поставке и оплате считаются неограниченными. Это необходимо для того, чтобы при моделировании не учитывать начальные капиталы ЭА.

На графе указаны следующие потоки оборотных средств:

1 → 2 – поток потерь товарных запасов (в денежном выражении) и финансовых средств суммарным объемом v_{12} , с известной величиной ρ_{12} интенсивности.

1 → 3 – продажа торговым предприятием оптовым покупателям товарной продукции в объеме, обусловленном *пакетным* заказом v_{13} с интенсивностью ρ_{13} .

1 → 0₄ – перевод денежного транша v_{14} поставщикам товаров (ПТ) для закупки *пакетного* заказа товарной продукции с целью ее дальнейшей реализации, с учетом оговоренной маржи μ_{14} . В одних случаях величины таких траншей могут приниматься в виде параметров управления. В других – считаться статистическими величинами, оцениваемыми на основе данных наблюдений в течение определенного интервала времени. То же можно сказать о средней интенсивности таких переводов и марже поставщиков.

3 → 1 – возврат торговому предприятию денежных средств в объеме ранее сделанных поставок *пакетных* заказов v_{31} , с учетом маржи ТП μ_{31} , с интенсивностью ρ_{31} .

0₄ → 1 – поставка товаров торговому предприятию в объеме *пакетного* заказа v_{41} (в денежном выражении) с интенсивностью ρ_{41} .

3 → 0₅ – поставка товаров розничным покупателям в соответствии с *пакетными* заказами v_{35} (в денежном выражении) с известной интенсивностью ρ_{35} для реализации через торговую сеть.

0₅ → 3 – оплата розничными покупателями полученных товаров в объеме v_{53} по цене с максимальной торговой надбавкой μ , включающей в себя

сумму маржей μ_{14} , μ_{31} , а так же маржу оптового покупателя μ_3 , с интенсивностью ρ_{53} .

Необходимо подчеркнуть, что объемы пакетов товарных запасов и финансовых средств v_{kh} ($k, h = 1, \dots, 7$) могут рассматриваться как параметры управления в определенных диапазонах изменения или приниматься случайными с известными математическими ожиданиями и дисперсиями, в зависимости от особенности решаемых задач. То же касается величины торговой надбавки μ , а также маржи торгового предприятия μ_{31} , поставщиков μ_{14} и оптовых покупателей (на графе не показана) μ_3 , которая определяется как разность $\mu_3 = \mu - (\mu_{31} + \mu_{14})$. Эти величины, вместе с ранее упомянутыми, могут являться количественным выражением условий контрактов. В зависимости от их сочетания заключаемые контракты могут быть либо выгодными, либо невыгодными.

Таким образом, общая маржа, полученная от реализации товара покупателям, определяется как максимальная торговая надбавка μ , установленная соответствующими нормативами.

Итак, очевидно, что для управления деятельностью ТП необходимо *прогнозировать* и *оптимизировать* как текущее, так и итоговое состояние счетов ЭА. Это должно базироваться на оценках предстоящих поступлений и отчислений финансовых средств, обеспечивающих как выполнение собственных обязательств, так и обязательств торговых партнеров и поставщиков, и при этом получение выгоды ТП (прибыли) в условиях определенных ограничений.

Прогнозируемая динамика движения ОС описывается с помощью переменных – случайных функций времени, введенных выше.

Эти переменные участвуют в формировании бухгалтерского баланса активов и пассивов.

В решаемой задаче в качестве актива выступают оборотные активы, а в качестве пассива – краткосрочные обязательства, исполняемые участниками в виде потоков поступлений денежных средств и пакетных заказов. Баланс активов и пассивов определяется текущим состоянием счетов ТП.

Отметим, что в качестве некоторых упрощений в данной задаче не будут рассматриваться поступления и отчисления в виде различных штрафов и пени, связанные с нарушениями договорных обязательств.

Дифференциальные уравнения для вектора математических ожиданий $m(t)$ и ковариационной матрицы $\theta(t)$, которые определяют статистические характеристики прогнозируемых переменных, имеют вид [2,3]:

$$\dot{m}(t) = M[S^T \rho] \quad (1)$$

$$\dot{\theta}(t) = M[S^T \rho Y^T + \dot{Y} \rho^T S + S^T \text{diag}(\rho) S], \quad (2)$$

где: M – оператор математического ожидания;

S – матрица переходов (структурная матрица системы);

$\dot{Y}(t) = Y(t) - m(t)$ – центрированный вектор, составленный из переменных $\dot{y}_1(t), \dots, \dot{y}_3(t)$;

$\text{diag}(\rho)$ – диагональная матрица, содержащая в диагонали вектор ρ , составленный из интенсивностей ρ_{ij} , в соответствии со строками матрицы S .

Существует скалярная форма уравнений (1 – 2): для математических ожиданий составляющих вектора $Y(t)$:

$$\dot{m}_\eta(t) = M \left[\sum_{k,h=0}^n S_{kh} (\vartheta_{kh}) \rho_{kh} (Y, t) \right]$$

для составляющих ковариационной матрицы:

$$\begin{aligned} \dot{\theta}_{\eta\xi}(t) = M \left\{ \sum_{k,h=0}^n \rho_{kh} (Y) \left[S_{kh} (\vartheta_{kh}) S_{kh\xi} (\vartheta_{kh}) + \right. \right. \\ \left. \left. + S_{kh\eta} (\vartheta_{kh}) \dot{y}_\xi + S_{kh\xi} (\vartheta_{kh}) \dot{y}_\eta \right] \right\} \quad (3) \end{aligned}$$

где \dot{y}_η – центрированная составляющая вектора $Y(t)$.

Здесь $\eta, \xi = 1, \dots, n$.

Если какие-либо параметры системы являются случайными величинами, то их известные вероятностные характеристики могут быть без труда введены в модель.

При моделировании системы необходимо вводить условия и ограничения, отражающие специфику решаемой задачи. Так, необходимо учитывать возможную связь между величинами v_{31} и v_{13} , отражающую договорные обязательства ТОП по возврату полученных ранее средств. Подобная связь существуют и между величинами v_{14} , v_{41} , v_{35} , v_{53} . В конечном счете, должны быть установлены соотношения между указанными величинами и с помощью математической модели – влияние этих связей на качество функционирования ТП, обеспечивающее его необходимую доходность. В частности, такими соотношениями могут быть линейные зависимости типа $v_{31} = \alpha v_{13}$, где $\alpha_1 \leq 1$ – варьируемая величина. Аналогичным образом могут быть установлены соотношения между другими величинами. Один из возможных вариантов – равенство $v_{14} = v_{41}$ и $v_{35} = v_{53}$

С учетом разметки графа и (1 – 3) запишем матрицы переходов S и интенсивностей ρ , на основе которых осуществляется математическое модели-

$$S = \begin{pmatrix} -v_{12} & v_{12} & 0 & 0 & 0 \\ -v_{13} & 0 & v_{13} & 0 & 0 \\ -(v_{14} + \mu_{14}) & 0 & 0 & 0_4 & 0 \\ v_{31} + \mu_{31} & 0 & -(v_{31} + \mu_{31}) & 0 & 0 \\ 0 & 0 & -v_{35} & 0 & 0_5 \\ v_{41} & 0 & 0 & 0_4 & 0 \\ 0 & 0 & v_{53} + \mu & 0 & 0_5 \end{pmatrix} \quad (4) \quad \rho = \begin{pmatrix} \rho_{12} \\ \rho_{13} \\ \rho_{14} \\ \rho_{31} \\ \rho_{35} \\ \rho_{41} \\ \rho_{53} \end{pmatrix} \quad (5)$$

рование непрерывной (пуассоновской) части системы:

Диагональная матрица $diag(\rho)$ представляет собой квадратную матрицу, диагональ которой содержит все составляющие вектора $\vec{\rho}$.

С использованием (1) – (5) получим дифференциальные уравнения для математических ожиданий и ковариаций переменных $y_1(t), y_2(t), y_3(t)$:

$$\begin{aligned} \dot{m}_1(t) &= \rho_{41} m_{\theta_{41}} - \rho_{12} m_{\theta_{12}} - \rho_{14} (m_{\theta_{14}} + \mu_{14}) - \rho_{13} m_{\theta_{13}} + \\ &\quad + \rho_{31} (m_{\theta_{31}} + \mu_{31}) \\ \dot{m}_2(t) &= \rho_{12} m_{\theta_{12}} \\ \dot{m}_3(t) &= -\rho_{31} (m_{\theta_{31}} + \mu_{31}) + \rho_{13} m_{\theta_{13}} - \rho_{35} m_{\theta_{35}} + \rho_{53} (m_{\theta_{53}} + \mu_{53}) \\ \dot{\theta}_{11}(t) &= \rho_{12} (m_{\theta_{12}}^2 + \theta_{\theta_{12}}) + \rho_{13} (m_{\theta_{13}}^2 + \theta_{\theta_{13}}) + \rho_{41} (m_{\theta_{41}}^2 + \theta_{\theta_{41}}) + \\ &\quad + \rho_{14} [(m_{\theta_{14}} + \mu_{14})^2 + \theta_{\theta_{14}}] + \rho_{31} [(m_{\theta_{31}} + \mu_{31})^2 + \theta_{\theta_{31}}] \\ \dot{\theta}_{22}(t) &= \rho_{12} (m_{\theta_{12}}^2 + \theta_{\theta_{12}}) \\ \dot{\theta}_{33}(t) &= \rho_{13} (m_{\theta_{13}}^2 + \theta_{\theta_{13}}) + \rho_{35} (m_{\theta_{35}}^2 + \theta_{\theta_{35}}) + \\ &\quad + \rho_{31} [(m_{\theta_{31}} + \mu_{31})^2 + \theta_{\theta_{31}}] + \rho_{53} [(m_{\theta_{53}} + \mu_{53})^2 + \theta_{\theta_{53}}] \\ \dot{\theta}_{12}(t) &= -\rho_{12} (m_{\theta_{12}}^2 + \theta_{\theta_{12}}) \\ \dot{\theta}_{13}(t) &= -\rho_{13} (m_{\theta_{13}}^2 + \theta_{\theta_{13}}) - \rho_{31} [(m_{\theta_{31}} + \mu_{31})^2 + \theta_{\theta_{31}}] \end{aligned} \quad (6)$$

Системы уравнений (6) решается при известных (заданных) начальных условиях численным мето-

дом Рунге-Кутта 4-го порядка. Результат решения представлен в виде графика (рис. 3).

Данная модель может в дальнейшем последовательно наращиваться за счет расширения количества участников рынка капиталов, товаров и услуг, а также учета дополнительной информации о параметрах функционирования экономических агентов. Это позволит более корректно обеспечивать поддержку решения различных задач. Наиболее полной модель станет тогда, когда станут известными и будут введены в модель достаточно достоверные оценки капиталов, которыми располагают упомянутые агенты, стратегия их использования, а также ценовые параметры рынка товаров, труда и капиталов.

В целях дальнейшего развития модели необходимо будет ввести ограничения на параметры, вытекающие из требований нормативной документации, а также учитывать различные привходящие факторы.

Разработанная математическая модель позволит решать следующие задачи, облегчающие руководству ТП выбор и обоснование решений:

- ◆ получать (прогнозировать) функции времени $m_i(t) \pm \varepsilon \sigma_i(t)$, определяющие математические ожидания процессов изменения активов и пассивов ТП в процессе его функционирования, а также доверительные интервалы, вычисляемые через среднеквадратические отклонения σ_i с квантилем ε , определяемым через заданную доверительную вероятность прогнозов;
- ◆ при заданном наборе параметров (исходных данных) подбирать начальное условие $y_1(0)$ – *уставный капитал*, обеспечивающий решение стратегических задач;
- ◆ при любом уставном фонде определять *начальную товарно-денежную политику*, выражаемую совокупностью различных процентных ставок, часть

Рис. 2. Ввод исходных данных для моделирования

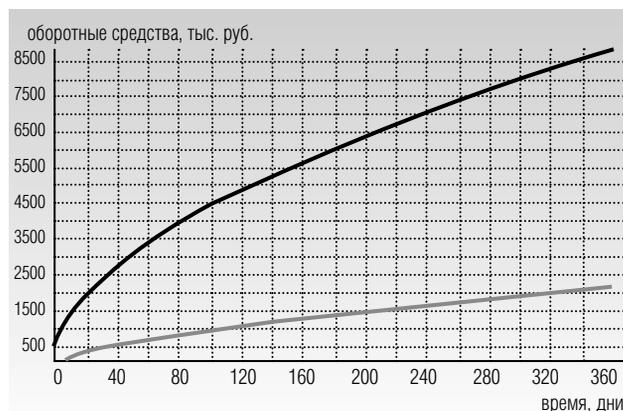


Рис. 3. Динамика счетов торгового предприятия

из которых устанавливается регулятором, другие части — рыночными механизмами и договорными обязательствами участников сделок, что позволит решать поставленные задачи на начальном периоде функционирования;

- ♦ обеспечивать поддержку планирования работы ТП в виде расчета эффективности вариантов выработываемых планов и др.

2. Результаты прогноза

Результаты прогнозирования динамики состояния оборотных средств представляются в интегрированном графическом виде, дающем возможность увидеть целостную картину предстоящего состояния оборотных средств торговой организации, а также основных участников торгово-экономической деятельности на заданном интервале времени.

Окно ввода исходных данных для проведения моделирования представлено на рис. 2. Динамика счетов торгового предприятия в виде нижней и верхней границы представлена на рис. 3.

Интенсивность использования оборотных средств

формируется в соответствии с коммерческими планами организации с учетом политики ценообразования, размера скидок и форм оплаты товара. Количество оборотных средств в группе задается в виде диапазона равномерно распределенного в пределах от минимального до максимального значения.

Верхняя и нижняя границы доверительных интервалов на графиках означают максимальное и минимальное прогнозируемые значения оборотных средств торгового предприятия, вычисленных с заданной доверительной вероятностью.

Заключение

Разработанная модель и полученные результаты прогнозирования пока не вполне соответствуют какой-либо реальной картине, а лишь отражают принципиальные возможности логистического подхода к выработке решений по управлению торговыми сетями.

Дальнейшее развитие данного подхода заключается в том, что на основании полученной модели, прошедшей определенную сертификацию, могут быть разработаны алгоритмы оптимизации, которые позволят:

- ♦ определять параметры *оптимальной товарно-денежной политики* для любого критерия оптимальности, в соответствии с «программой минимум», «программой максимум» или промежуточными вариантами;
- ♦ обеспечивать *оптимальное планирование* работы ТП на различных временных этапах с учетом выбранной стратегии развития;
- ♦ обеспечивать *оптимальную корректировку планов*, начиная с произвольного текущего момента времени и др. ■

Литература

1. Ямпольский С.М., Шаламов А.С. Логистический подход к автоматизации управления оборотными средствами торгового предприятия // Логистика и управление цепями поставок. 2013. №4 (57). С. 65-70.
2. Шаламов А.С. Интегрированная логистическая поддержка. М.: Университетская книга, 2008. 254 с.
3. Синицын И.Н., Шаламов А.С. Лекции по теории систем интегрированной логистической поддержки. М.: Торус Пресс, 2011. 546 с.

MODELING OF TRADE ENTERPRISE PROCESSES

Sergey YAMPOLSKY,

Associate Professor, Department of Business Analytics, Faculty of Business Informatics, National Research University Higher School of Economics.

Address: 20, Myasnitskaya str., Moscow, 101000, Russian Federation

Anatoly SHALAMOV,

Researcher, Department of Statistical Problems of Informatics and Management, Institute of Informatics Problems of the Russian Academy of Sciences

Address: 44, Vavilova str., Moscow, 119333, Russian Federation

E-mail: syampolsky@hse.ru, a-shal5@yandex.ru

The article covers issues of effectiveness of working capital management of a trading enterprise on the basis of the automated planning. This approach gives the possibility to use the classical methods of optimization of dynamic systems to determine the main parameters of the economic policy of the commercial enterprise for providing the best strategy for development. Using the proposed approach, algorithms we can be created for evaluating both managed and unmanaged risk and also for finding reasonable solutions for preventing them.

When presenting the material, operational environment of the commercial enterprise is describing in traditional terms of trade and financial market, allowing its use everywhere where the information system «IC» is run, in the form of incremental software complex for forecasting and optimization.

The developed mathematical model allows to solve the tasks of easing the selection and justification of decisions for the managers of a shopping enterprise, including these: prediction of time functions defining mathematical expectations of assets and liabilities, definition of share capital providing strategic objectives, identification of primary commodity-monetary policy, evaluation of efficiency of variants of produced plans.

Results of forecasting the dynamics of the situation of circulating assets are presented in an integrated graphical form, which provides an opportunity to see the full picture of the forthcoming state of current assets of a trade organization, and the main participants of the trade and economic activity in a given time interval.

Key words: trade enterprise, suppliers, wholesale and retail buyers, current assets, automated planning, order management, discrete Markov process with continuous time.

References

1. Yampolsky S., Shalamov A. (2013) Logisticheskij podhod k avtomatizacii upravlenija oborotnymi sredstvami torgovogo predpriyatija [The Logistic Approach to the Automation of a Trade enterprise's Working Capital Management]. *Logistics and Supply Chain Management*, no. 4 (57), pp. 65–70.
2. Shalamov A. (2008) Integrirovannaja logisticheskaja podderzhka [Integrated Logistic Support]. Moscow: University Book. (in Russian).
3. Sinicin I., Shalamov A. (2011) *Lekcii po teorii sistem integrirovannoj logisticheskaj podderzhki* [Lectures on the Theory of Systems of Integrated Logistics Support]. Moscow: Torus Press. (in Russian).

ПРОГНОЗ ПОВЕДЕНИЯ КЛИЕНТОВ СУПЕРМАРКЕТОВ С ПОМОЩЬЮ ВЕСОВЫХ СХЕМ ОЦЕНОК ВЕРОЯТНОСТЕЙ И ПЛОТНОСТЕЙ¹

А.Г. Дьяконов,

доктор физико-математических наук, профессор кафедры математических методов прогнозирования, факультет вычислительной математики и кибернетики, Московский государственный университет им. М.В. Ломоносова, старший научный сотрудник, Вычислительный центр им. А.А. Дородницына Российской академии наук

*Адрес: 119991, Москва, ГСП-1, Ленинские горы, МГУ им. М.В. Ломоносова, 1, стр. 52
E-mail: djakonov@mail.ru*

Рассмотрены две задачи, связанные с поведением клиентов сети супермаркетов: прогнозирование даты следующего визита каждого клиента и суммы его покупок. Первая задача сведена к задаче оценки вероятностей визитов, вторая – к задаче восстановления плотностей распределений сумм покупок каждого пользователя. Для решения указанных задач предложено использовать взвешенные схемы: каждой точке выборки ставится в соответствие вещественное неотрицательное число (вес). Веса позволяют учитывать дополнительную информацию, например устаревание данных (точки соответствующие старым данным имеют меньшие веса). В работе рассмотрено несколько весовых схем (способов приписывания весов точкам выборки), произведена их настройка (оптимизация качества оценки вероятности или плотности по параметрам весовой схемы). Показано, что использование весовых схем не приводит к переобучению, т.е. настройка весов на обучении не понижает качество на независимой контрольной выборке. Показана возможность использования ансамблирования для повышения качества решения рассмотренных задач, т.е. построения нескольких алгоритмов и составления их линейной комбинации. Все эксперименты произведены на реальных данных крупного Международного конкурса по разработке алгоритмов анализа данных. Специфика данных (отсутствие праздников на финальном временном отрезке статистики) позволила при решении указанных задач сосредоточиться исключительно на статистических методах решения. Кроме того, рассмотрены вопросы построения алгоритмов, которые одновременно решают обе задачи: прогнозирования даты следующего визита и суммы покупок. Показано, что не всегда их можно решать независимо. Предложен метод оптимизации функционала, который оценивает решение обеих задач.

Ключевые слова: прогнозирование, оценивание вероятности, восстановление плотности, непараметрические методы, прикладные задачи.

¹ Работа поддержана грантами РФФИ № 12-07-00187, №14-07-00965.

Введение

Пусть есть конечное множество клиентов сети супермаркетов. Для построения и исследования алгоритмов оно предполагается фиксированным. Для каждого клиента известны все даты его визитов и суммы покупок в каждую из этих дат. Статистика дана за последние d недель. Обычно подобная статистика собирается по покупкам с использованием скидочных карт (что позволяет идентифицировать клиентов), поэтому рассматриваемое множество не просто множество людей, посещающих магазины сети, а людей, которые достаточно часто посещают эти магазины. Необходимо предсказать дату следующего визита и сумму покупок. Что понимается под точностью прогноза, будет оговорено далее.

Сначала введём необходимый формализм. Статистику посещений конкретного клиента можно записать в виде матрицы с неотрицательными элементами

$$S = \|s_{ij}\|_{d \times 7}, \quad (1)$$

где s_{ij} — сумма покупок клиента на j -й день i недель назад. Например, если сегодня конец воскресенья, то среда этой недели соответствует элементу s_{13} , а пятница прошлой — s_{25} . Считаем, что визит был тогда и только тогда, когда сумма покупок положительна. Хотя теоретически допустима ситуация «клиент пришёл, но ничего не купил», практически она не отображается в статистике (которая, как уже говорилось, формируется по чекам покупок с дисконтными картами). В дальнейшем будем рассматривать каждого клиента по отдельности (поэтому обозначение (1) не содержит номер клиента).

Отметим, что данные, на которых тестировались методы, охватывали период чуть больше года. Последние месяцы этого периода не включали праздники (Новый год, 8 марта и т.п.) и особые дни (например, 1 сентября в РФ). Поэтому считаем статистику однородной, без влияния сезонных факторов. С одной стороны, это существенно упрощает задачу. С другой, делает возможным тестирование различных методов оценки вероятностей и плотностей распределений.

Все графики и результаты экспериментов представлены для данных компании dunnhumby, которые использовались для проведения Международного конкурса [1]. Автор работы стал победителем этого соревнования среди 279 участников. Здесь представлены результаты последующих исследо-

ваний, кроме того, все выводы согласуются с экспериментами, которые были сделаны также и на данных российских Интернет-магазинов.

В работе некоторые модификации алгоритмов дают улучшения на доли процентов. Отметим, что это действительно улучшения (качество увеличивается и на независимом контроле и на других данных). Кроме того, в современных бизнес-задачах даже незначительные улучшения могут дать существенный доход, применение же нескольких модификаций увеличивает качество на проценты. В режиме соревнования [1] именно эти модификации позволили построить лучший алгоритм.

1. Решение задачи прогноза даты следующего визита

1.1. Оценивание вероятностей, пересчёт, весовые схемы

Для каждого клиента необходимо предсказать день его следующего визита: $j \in \{1, 2, \dots\}$ (1 соответствует завтрашнему дню, 2 — послезавтрашнему и т.д.). Прогноз считается верным, если мы угадываем этот день, т.е. клиент придёт в j -й день и не придёт в 1, 2, ..., $(j-1)$ -й день. Далее при оценке методов будем указывать процент верных прогнозов (по всем клиентам): из данных удаляется информация о последней неделе, метод прогнозирует день визита каждого клиента на этой удалённой неделе, процент таких верных прогнозов и считается качеством метода.

Имеет смысл перейти от матрицы сумм покупок $S = \|s_{ij}\|_{d \times 7}$ к бинарной матрице визитов $V = \|v_{ij}\|_{d \times 7}$:

$$v_{ij} = 1 \Leftrightarrow s_{ij} > 0.$$

Предложим вероятностную модель поведения клиента: в j -й день недели он с вероятностью p_j посещает магазин. Напомним, что первый день недели — это день, которому соответствует $j=1$ при формировании матрицы (1) (прогнозирование происходит в конце 7-го дня, сразу перед началом новой недели). В такой вероятностной модели вполне естественно вычислить вероятности первых визитов в j -е дни, что делается по следующей «формуле пересчёта»:

$$\tilde{p}_j^1 = p_j \prod_{r=1}^{j-1} (1 - p_r), \quad (2)$$

(считаем, что $\tilde{p}_1^1 = p_1$) и выбрать из них максимальное значение:

$$j = \operatorname{argmax}_{r \in \{1, 2, \dots, 7\}} \tilde{p}_r^1.$$

В этой вероятностной модели неявно делаются следующие предположения:

1. Поведения разных клиентов сети супермаркетов независимы (поэтому ответ для конкретного клиента зависит только от его статистики посещений).

2. Каждый клиент обязательно посетит магазин в течение следующей недели.

Отметим, что хотя первое предположение вряд ли верно, пока не удалось предложить методы, использующие зависимости в поведении разных клиентов, которые превосходили бы по качеству описанные в данной работе.

Второе предположение легко обойти, «увеличив неделю», например, до 14 дней. В этом случае считаем, что клиент обязательно посетит магазин в течение ближайших 14 дней. Как показывают эксперименты, качество алгоритма не меняется при таком «увеличении недели». Важно понимать, что очень много зависит от свойств данных, на которых предполагается использовать метод. Далее приводятся результаты экспериментов на реальных данных [1], для которых алгоритм «клиент придёт завтра» уже даёт более 25% верных прогнозов (см. рис. 1). Отметим также, что в соответствии с этими данными за неделю магазины посещают около 85% клиентов (держателей дисконтных карт).

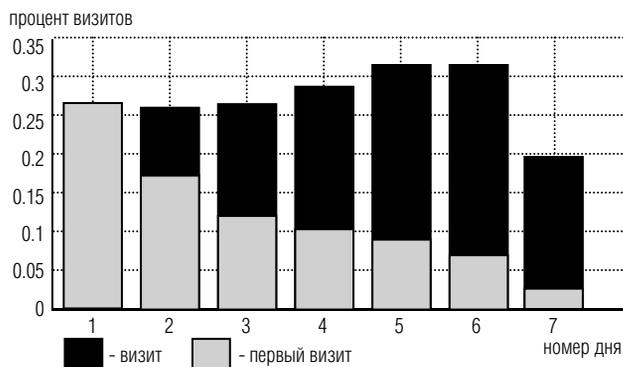


Рис. 1. Процент визитов и первых визитов на неделе в зависимости от дня недели

Простейший (и самый распространенный на практике) способ оценивания вероятности – через частоту, в данном случае, вероятности p_j можно оценить через частоты визитов за последние d недель:

$$p_j = \frac{1}{d} \sum_{i=1}^d v_{ij}. \quad (3)$$

Однако необходимо учитывать, что информация последних недель ценнее информации d недель назад, поэтому можно рассмотреть взвешенную схему оценки вероятности:

$$p_j = \sum_{i=1}^d w_i v_{ij},$$

$$w_1 \geq w_2 \geq \dots \geq w_d \geq 0, \sum_{i=1}^d w_i = 1.$$

Подобные весовые схемы применяются в анализе данных, например во взвешенном методе ближайших соседей [2]. Часто достаточно эффективный способ задания весов следующий:

$$w_i^N = \left(\frac{d-i+1}{d} \right)^\delta, i \in \{1, 2, \dots, d\}, \quad (4)$$

$$w_i = \frac{w_i^N}{\sum_{i=1}^d w_i^N}, i \in \{1, 2, \dots, d\}. \quad (5)$$

Здесь выбор конкретной весовой схемы сводится к выбору параметра $\delta \in [0, +\infty)$. При этом, значение $\delta = 0$ соответствует равным весам, т.е. вычислению вероятностей по формуле (3).

В дальнейшем при задании весов будем как в (4) использовать верхний индекс N , подразумевая, что далее необходима нормировка по сумме (5), которую явно указывать не будем.

Как видно из рис. 2 (сплошная линия) применение схемы весов позволяет немного повысить качество прогноза: простой метод по формулам (3), (2) даёт точность 35.9%, взвешенный – около 36.36%. Перестановка и прямой метод (качество которых также изображено на рис. 2) будут введены чуть ниже (в конце этого раздела и в разделе 1.2).

Как и во взвешенном методе ближайших соседей [2], можно выбирать ещё параметр «число ненулевых весов» k :

$$w_i^N = \begin{cases} \left(\frac{d-i+1}{d} \right)^\delta, & i \in \{1, 2, \dots, k\}, \\ 0, & i \in \{k+1, \dots, d\}. \end{cases}$$

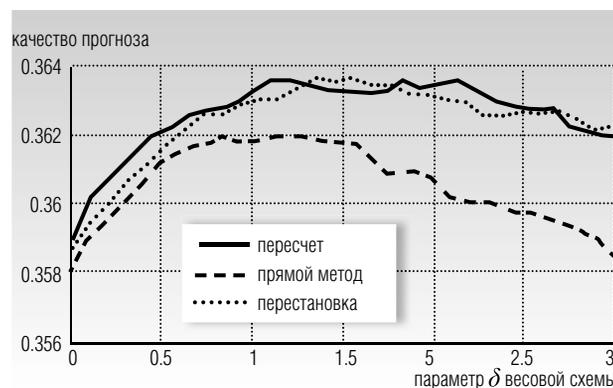


Рис. 2. Зависимость качества прогноза от степени δ

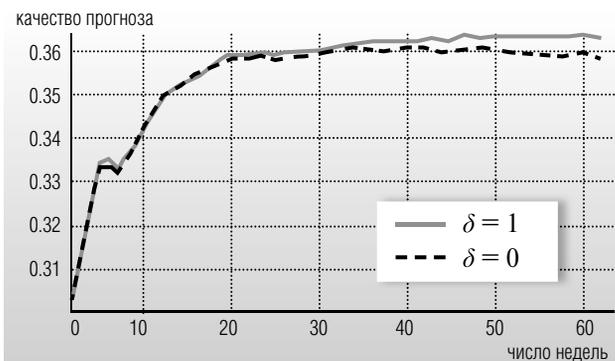


Рис. 3. Зависимость качества прогноза от числа учитываемых недель

Но, как видно на рис. 3, здесь работает принцип «чем больше данных, тем лучше». Кстати, простой алгоритм «будет как на прошлой неделе» ($k=1$) показывает качество 29.3% (лучше константного $j=1$, но существенно хуже рассматриваемых).

Также на практике иногда удаляют куски данных «без информации». В данном случае – нулевые строки из матрицы $V = \|v_{ij}\|_{d \times 7}$, чтобы не учитывались недели, в которые не было визитов. Для удобства (это упрощает программирование во многих пакетах, например Matlab), они остаются в матрице, просто «съезжают вниз», т.е. если

$$I = \{i_1, \dots, i_r\} = \{i \mid v_{i1} + \dots + v_{i7} > 0\}, 1 \leq i_1 < \dots < i_r \leq d,$$

то матрица $V = \|v_{ij}\|_{d \times 7}$ заменяется матрицей

$$\begin{pmatrix} v_{i_1,1} & \dots & v_{i_1,7} \\ \dots & \dots & \dots \\ v_{i_r,1} & \dots & v_{i_r,7} \\ 0 & \dots & 0 \\ \dots & \dots & \dots \end{pmatrix}$$

Как видно из рис. 2 (чёрные точки), такая перестановка недель не сильно увеличивает качество (36.37%), но делает функцию качества от параметра $\delta \in [0, +\infty)$ «почти унимодальной».

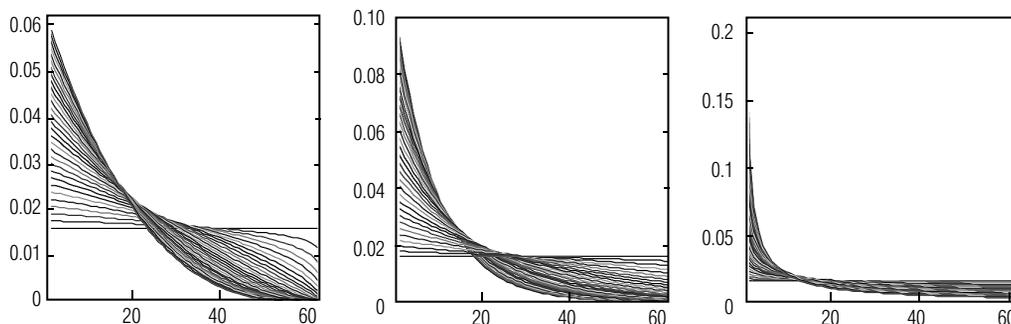


Рис. 4. Три разные весовые схемы: вес недели в зависимости от её номера

Вообще, весовые схемы можно выбирать по-разному. Например,

$$w_i^N = \lambda^i, i \in \{1, 2, \dots, d\}, \lambda \in (0, 1], \quad (6)$$

или

$$w_i^N = \frac{1}{i^\gamma}, i \in \{1, 2, \dots, d\}, \gamma \in [0, +\infty). \quad (7)$$

На рис. 4 показаны распределения весов (после нормировки по сумме (5)) при первой схеме (4) для различных $\delta \in [0, 3]$ (слева), при второй (6) для различных $\lambda \in [0.9, 1]$ (в центре) и при третьей (7) для различных $\gamma \in [0, 1]$ (справа). В каждом случае отрезок возможных значений параметра делился на равные части 30-ю точками и строились графики для 30-ти различных значений параметра. Зависимости качества от параметров при различных нормировках похожи, ср., например, рис. 2 и рис. 5.

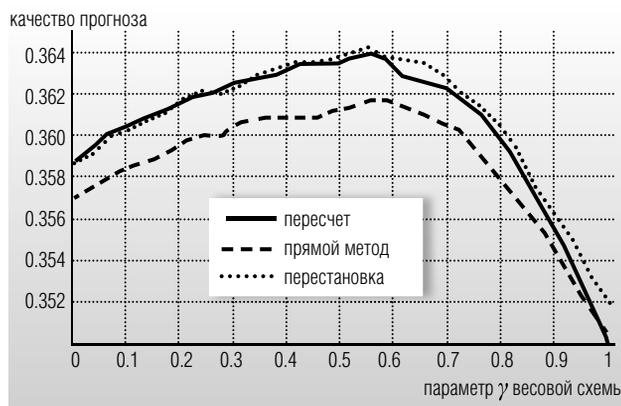


Рис. 5. Зависимость качества прогноза от $\gamma \in [0, 1]$

1.2. Прямое оценивание вероятностей

Оценить вероятности первых визитов можно и непосредственно (без пересчёта по оценкам вероятностей визитов):

$$\tilde{p}_j^2 = \frac{1}{d} |\{i \in \{1, 2, \dots, d\} : v_{i1} = \dots = v_{i,j-1} = 0, v_{ij} = 1\}|, \quad (8)$$

т.е. для каждого дня недели посчитать долю недель, в которые в этот день был первый визит. Такой метод кажется даже «более естественным». Если в матрице V в каждой строке оставить лишь первый единичный элемент (если он есть), а остальные занулить, т.е. перейти к матрице первых визитов $V' = \|v'_{ij}\|_{d \times 7}$, то для оценки вероятности можно использовать следующую взвешенную схему:

$$\tilde{p}_j^2 = \sum_{i=1}^d w_i v'_{ij} \quad (9)$$

Равные веса соответствуют формуле (8). Отметим, что это окончательная формула (пересчёта вероятностей (2) не требуется).

Из рис. 2, 5 (прерывистая линия) видно, что качество прямого метода чуть ниже, чем описанного ранее метода с пересчётом вероятностей.

1.3. Ансамблирование

Часто качество прогнозирования повышается при «ансамблировании» нескольких методов, т.е. построении алгоритма на базе разных методов, тогда недостатки одного метода, компенсируются достоинствами других. На принципе ансамблирования построены как многие современные эффективные алгоритмы машинного обучения (бэггинг, бустинг и т.д.), так и некоторые теории анализа и построения алгоритмов (например, алгебраический подход, комитетный подход и т.д.) [3].

Выше предложены две оценки вероятности первого визита (2) и (8). Первый способ ансамблирования – «стандартный ансамбль» – взять их выпуклую комбинацию:

$$\tilde{p}_j = \alpha \tilde{p}_j^1 + (1 - \alpha) \tilde{p}_j^2, \quad (10)$$

где $\alpha \in [0, 1]$ – параметр для настройки.

Предложим другой способ ансамблирования: будем брать не выпуклую комбинацию \tilde{p}_j^1 и \tilde{p}_j^2 , а выпуклую комбинацию p_j и \tilde{p}_j^2 :

$$\begin{aligned} \alpha p_j + (1 - \alpha) \tilde{p}_j^2 &= \alpha \sum_{i=1}^d w_i v_{ij} + (1 - \alpha) \sum_{i=1}^d w_i v'_{ij} = \\ &= \sum_{i=1}^d w_i (\alpha v_{ij} + (1 - \alpha) v'_{ij}) \end{aligned} \quad (11)$$

(после этого требуется пересчёт вида (2) для определения вероятностей первых визитов). Такой метод не имеет строгого теоретического обоснования (хотя в [4] представлены некоторые аргументы в его пользу), но очень неплохо показал себя на практике. Его можно интерпретировать как более сложное задание весов, см. (11). На рис. 6 видно, что предложенный

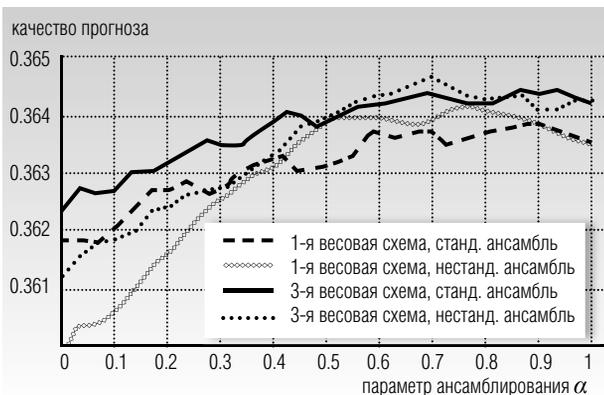


Рис. 6. Качество ансамблирования от параметра $\alpha \in [0, 1]$

«нестандартный ансамбль» лучше «стандартного» и позволяет повысить качество с 36.35% до 36.41% в случае первой весовой схемы (4) и с 36.42% до 36.46% – в случае использования третьей (7).

1.4. Проблема переобучения

Переобучением в анализе данных называется эффект, при котором на новых данных алгоритм работает существенно хуже, чем на исходных (на основе которых он был построен) [5]. Как правило, переобучение связывают с излишней сложностью модели алгоритмов. Для проверки отсутствия переобучения используют отложенный контроль: часть данных не используют при настройке параметров, затем на этой части (отложенной выборке) проверяют качество алгоритма.

Все приводимые графики получены для фиксированной выборки. Естественно, встаёт вопрос о надёжности алгоритма, т.е. как он будет работать на новых данных. Оказывается, что предложенные взвешенные схемы с оптимальными параметрами практически не склонны к переобучению. На рис. 7 представлена столбцовая диаграмма с результатами экспериментов на обучающей выборке и отложенном контроле (в качестве контроля использовалась последняя неделя, за которую есть статистика визитов клиентов, она не использовалась для настройки параметров). Показано качество следующих алгоритмов:

- 1) константный («клиент придёт на следующий день»),
- 2) визит клиента как на прошлой неделе,
- 3) вероятности (3) оценены по последним 5 неделям,
- 4) вероятности оценены по всем неделям,
- 5) оптимальные значения весов (7),
- 6) оптимальное нестандартное ансамблирование (11).

Несмотря на кажущееся усложнение алгоритма (ввод весов, ансамблирование), не только повы-

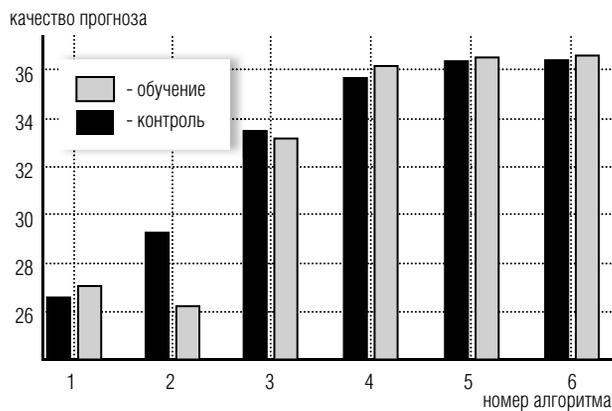


Рис. 7. Качество на обучении и отложенном контроле для шести алгоритмов

шается качество, но и уменьшается разница между обучением и контролем. Поэтому далее качество на отложенном контроле не будет приводиться (оно почти такое же, как на обучении).

2. Решение задачи прогноза суммы покупок

В день визита клиента в супермаркет необходимо предсказать его траты с точностью до $\varepsilon = 10$ у.е., т.е. для t -го клиента ответ a^t алгоритма прогнозирования считается верным, если $|s^t - a^t| \leq \varepsilon$, где s^t – истинные траты клиента. Качество на всей выборке оценивается как процент верных ответов:

$$\frac{1}{n} \sum_{t=1}^n \begin{cases} 1, & |s^t - a^t| \leq \varepsilon, \\ 0, & |s^t - a^t| > \varepsilon, \end{cases}$$

где n – число клиентов (для которых сделан прогноз). По нашей договорённости дальше индекс клиента t будем опускать, решая задачу прогнозирования независимо для каждого клиента.

Эту задачу можно решать как обычную задачу восстановления плотности. Пусть клиент делал покупки на суммы s_1, \dots, s_m . Будем считать, что это реализации некоторой случайной величины. Если оценить плотность её распределения, то разумно в качестве прогноза выдать значение s , в котором плотность максимальна. Есть три подхода к оцениванию плотности [6]:

- 1) параметрический (когда известно распределение с точностью до параметров),
- 2) смеси распределений (когда известно, что плотность представляется в виде выпуклой комбинации плотностей, известных с точностью до параметров),
- 3) непараметрический (когда не известен вид распределения).

Для каждого клиента траты достаточно сильно отличаются (см. рис. 8), кроме того, часто от дня недели зависят суммы покупок и разброс этих сумм (рис. 9). Для решения задачи был выбран достаточно универсальный непараметрический подход, поскольку:

- 1) вид распределения априорно не известен,
- 2) проще настраивать плотность, а функционал качества автоматически определяет ядерную функцию (см. далее),
- 3) подход допускает обобщения с использованием весовых схем (и, соответственно, учёт времени покупок).

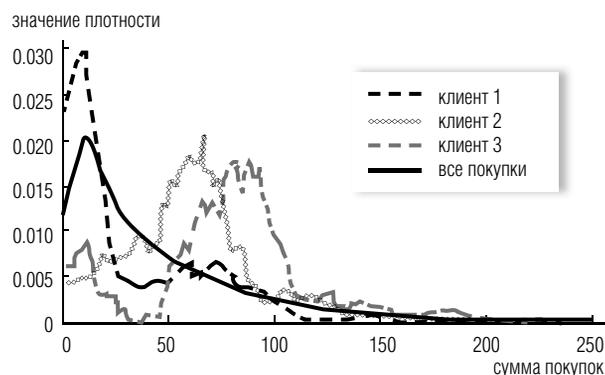


Рис. 8. Плотности распределения покупок

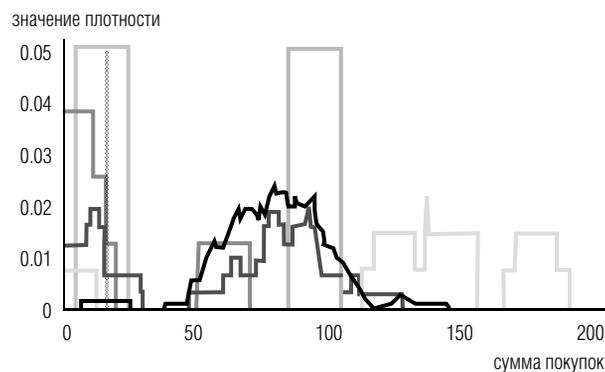


Рис. 9. Плотности покупок одного пользователя в разные дни недели

Для оценки плотности непараметрическим методом Розенблатта–Парзена [7, 8] пользуемся формулой

$$f(x) = \frac{1}{m} \sum_{i=1}^m K(s_i - x), \quad (12)$$

где функция $K(x)$ – ядерная (ядро), т.е. симметричная неотрицательная вещественная функция, монотонно не возрастающая на положительной полуоси, интегрируемая, причём

$$\int_{-\infty}^{+\infty} K(x) dx = 1.$$

Эти требования обеспечат, чтобы выражение (12) имело смысл плотности.

Функционал качества диктует выбор ядра:

$$K(s-x) = \begin{cases} 1/2\varepsilon, & |s-x| \leq \varepsilon, \\ 0, & |s-x| > \varepsilon. \end{cases}$$

Это становится ясно из различных модельных примеров. Например, если $s_1 = \dots = s_m = s$, т.е. клиент всё время делает покупки на одну и ту же сумму, то алгоритм может в качестве прогноза выдавать любое значение из отрезка $[s - \varepsilon, s + \varepsilon]$. Если же, например, клиент совершает покупки только на суммы из двухэлементного множества $\{s - \varepsilon, s + \varepsilon\}$, то алгоритм должен прогнозировать значение s (только так он «угадывает» в обоих случаях). При описанном выборе ядра выражение (12) в первом случае достигает максимального значения на отрезке $[s - \varepsilon, s + \varepsilon]$, а во втором – в точке s .

Отметим, что на рис. 8 – 9 изображены плотности, полученные по формуле (12). На рис. 8 изображена плотность для всех покупок (всех пользователей) и для трёх пользователей, совершивших 182, 70 и 80 покупок соответственно. Для последнего на рис. 9 показана более подробная информация по дням недели.

Формула (12) легко обобщается на случай использования весов:

$$f(x) = \sum_{i=1}^m w_i K(s_i - x), \quad (13)$$

поэтому оценка плотности однозначно определяется весами w_i , $i \in \{1, 2, \dots, m\}$. Веса имеет смысл использовать, поскольку ценность информации о покупке на сумму s_i зависит от того, в какой день недели была сделана покупка и как давно. Далее тот факт, что i -й покупке на сумму s_i соответствует вес w_i будем обозначать так:

$$s_i \leftrightarrow w_i$$

(это не вызовет путаницы). При этом если сумма всех весов w_i отлична от единицы, то считаем, что перед использованием формулы (13) необходимо сделать нормировку: поделить каждый вес на эту сумму.

Рассмотрим следующий выбор весов. Пусть s_1, \dots, s_m – все покупки пользователя, упорядоченные по дате от самой последней до самой первой, s'_1, \dots, s'_m – покупки, сделанные в день недели, на который делается прогноз, упорядоченные аналогично. Плотность по формуле (13) будем восстанавливать для расширенного набора $s'_1, \dots, s'_m, s_1, \dots, s_m$. Некоторые покупки здесь учитываются дважды. Это соответствует тому, что их вес будет равен сумме двух весов

(для первого учёта и для второго). Пусть

$$s'_i \leftrightarrow \beta \frac{(m' - i + 1)^{\rho'}}{\sum_{j=1}^{m'} j^{\rho'}}, \quad i \in \{1, 2, \dots, m'\} \quad (14)$$

$$s_i \leftrightarrow (1 - \beta) \frac{(m - i + 1)^{\rho}}{\sum_{j=1}^m j^{\rho}}, \quad i \in \{1, 2, \dots, m\}. \quad (15)$$

Параметр β определяет вклад каждого из двух поднаборов, степени ρ, ρ' – весовые схемы для каждого из поднаборов. Отметим, что здесь происходит целых три нормировки: отдельно для каждого из поднаборов (14), (15), а затем общая нормировка перед применением формулы (13).

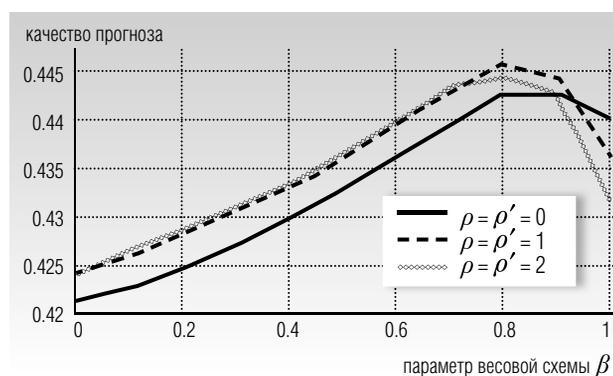


Рис. 10. Качество прогноза суммы покупок от параметра β

На рис. 10 показано, что в случае, когда мы не используем весовую схему ($\beta = 0, \rho = 0$) качество прогноза – 42.12%, использование весов повышает качество до 44.55%. При настройке параметров хорошо работает метод покоординатного спуска, поскольку, как видно на графиках, функция качества «достаточно выпуклая». При $\rho = 0.75, \rho' = 1.25, \beta = 0.8$ качество прогноза достигает 44.68%.

Конечно, очень важно настраивать параметры на «правильной выборке». Можно для всех клиентов взять первый визит на последней неделе и по этим данным производить настройку, а можно сначала запустить алгоритм прогноза дня первого визита и настройку производить только по угаданным визитам. Дальше мы на этом остановимся подробнее, пока отметим, что в данном случае (см. рис. 11) выборка не влияет на значения оптимальных параметров, хотя качество прогноза суммы покупок на угадываемых днях выше: более 48.5% (т.е. в ожидаемые дни клиент ведёт себя более предсказуемо).

Вообще, в этой задаче можно придумать много разных весовых схем. К сожалению, для нахождения лучшей их нельзя «перебрать». В некоторых

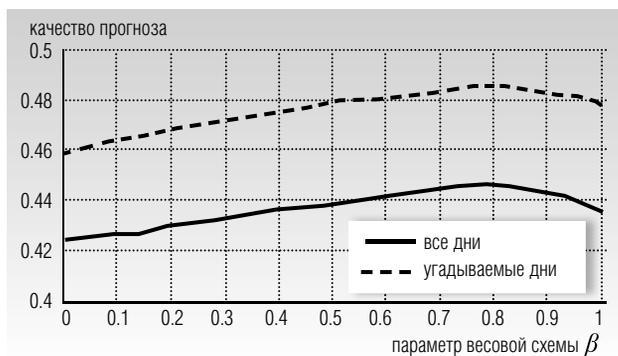


Рис. 11. Качество прогноза суммы покупок от параметра β при $\rho = 0.7, \rho' = 1.6$

прикладных задачах применялись такие «экзотические» схемы: формировался набор $s''_1, \dots, s''_{m''}$ из первых $\tilde{m} = \min(m', \omega')$ элементов набора $s'_1, \dots, s'_{m'}$ и первых $\min(m, \omega + \lfloor \sigma \tilde{m} \rfloor)$ элементов набора s_1, \dots, s_m , использовалась степенная схема весов

$$s''_i \leftrightarrow (m'' - i + 1)^\rho, i \in \{1, 2, \dots, m''\}.$$

Здесь $\omega', \omega, \sigma, \rho$ – параметры для настройки.

3. Предсказание поведения клиента

Предположим теперь, что для каждого клиента одновременно надо решить две задачи: предсказать дату визита и сумму покупок. Причём ответ считается верным, если верно решены обе задачи: правильно предсказан день, а ошибка прогноза суммы покупок не больше ε .

Сомнительным выглядит «наивный» метод решения: предсказать день, а потом в этот день предсказать сумму. Действительно, в идеальном случае (когда наши оценки вероятностей точные) вероятность правильности нашего ответа j в первой задаче это и есть максимальная вероятность \tilde{p}_j среди $\tilde{p}_1, \dots, \tilde{p}_7$ (и именно её мы максимизируем). Пусть теперь q_j –



Рис. 12. Качество предсказания поведения в зависимости от параметра h

вероятность правильности ответа во второй задаче (прогнозировании трат в j -й день). Тогда необходимо максимизировать $\tilde{p}_j q_j$ по j и q_j , а не последовательно \tilde{p}_j по j и $\tilde{p}_j q_j$ по q_j как в «наивном» методе.

Предложенный выше метод решения второй задачи позволяет оценить свою надёжность: максимальное значение плотности $f(x)$ это и есть q_j – оценка вероятности правильности ответа.

На рис. 12 показано качество предсказания поведения клиента (т.е. дня первого визита и суммы трат в этот день) в зависимости от параметра h при максимизации

$$\tilde{p}_j(q_j + h) \rightarrow \max_j. \quad (16)$$

При выборе очень большого h получается наивный метод, в этом случае качество равно 17.38, при малых $h \in [0, 0.5]$ качество около 17.48. Значение $h = 0$ соответствует методу, описанному выше. Таким образом, это параметр регулирует переход от наивного метода к «продвинутому».

Отметим, что, несмотря на простоту метода (16), в соревновании [1] все участники, кроме победителя, использовали наивный метод.

4. Заключение

В данной работе рассмотрена конкретная задача: прогнозирование следующего визита клиента и суммы покупок. Аналогичные задачи возникают практически везде, где клиенты пользуются некоторыми услугами. Например, для репозитория электронных материалов (видео, книги, софт и т.п.): для каждого пользователя предсказать следующий день прихода на сайт и объём скачанных материалов. Или для потребителей Интернет-трафика: предсказать время следующего потребления и объёмы.

Аналогичные эксперименты были проведены для одного из российских Интернет-магазинов. В качестве основной выборки были взяты постоянные клиенты (которые делают заказы не реже чем 1 раз в N дней), поскольку основная часть клиентов (более 90% зарегистрированных на сайте) делает заказы изредка (сделали лишь 1 заказ, совершают заказы только перед крупными праздниками и т.п.) Вид всех графиков и даже оптимальные значения некоторых параметров для этих данных такие же, как и для данных [1]. Отличаются лишь показатели качества.

Методы, рассмотренные в работе, зависят от большого числа параметров. Итоговый метод ре-

шения задачи прогнозирования поведения клиента зависит от следующих параметров:

- 1) параметр весов γ при оценке вероятностей визитов/первых визитов, см. (7),
- 2) параметр ансамблирования α , см. (11),
- 3) первый параметр весовой схемы ρ при оценке суммы покупок, см. (15),
- 4) второй параметр весовой схемы ρ' при оценке суммы покупок, см. (14),
- 5) третий параметр весовой схемы β при оценке суммы покупок, см. (14) – (15),
- 6) параметр h учёта правильности прогноза трат, см. (16).

Естественно, оптимизация по всем этим параметрам – отдельная большая задача. На практике она делалась покоординатным спуском (последовательно фиксировались все параметры, кроме одного, по которому проводилась оптимизация). В работе мы привели графики «срезов качества»: качества прогнозов от значения одного из параметров при фиксированных остальных. Это сделано для того, чтобы показать вид функционала качества: практически все графики изображают унимодальные функции, часто «достаточно гладкие».

Автор выражает благодарность анонимному рецензенту за ценные замечания, которые помогли существенно улучшить статью. ■

Литература

1. Международный конкурс *Dunnhumby's Shopper Challenge* / Kaggle [Электронный ресурс]: <http://www.kaggle.com/c/dunnhumbychallenge> (дата обращения 24.11.2013)
2. Воронцов К.В. Метрические алгоритмы классификации // Электронные лекции [Электронный ресурс]: www.ccas.ru/voron/download/MetricAlgs.pdf (дата обращения 24.11.2013)
3. Воронцов К.В. Алгоритмические композиции // Электронные лекции [Электронный ресурс]: <http://www.ccas.ru/voron/download/Composition.pdf> (дата обращения 24.11.2013)
4. Дьяконов А.Г. Решение задач анализа данных, основанное на линейной комбинации деформаций // *Машинное обучение и анализ данных*. 2013. Т. 1. № 5. С. 543–554.
5. Воронцов К.В. Комбинаторная теория надёжности обучения по прецедентам. Дис. ... д-ра физ.-мат. наук. М., 2010. 271 с.
6. Расин Д. Непараметрическая эконометрика: вводный курс // *Квантиль*. 2008. №4. С. 7-56.
7. Parzen E. On Estimation of a Probability Density Function and Mode // *Annals of Mathematical Statistics*. 1962. No. 33. P. 1065–1076.
8. Rosenblatt M. Remarks on Some Nonparametric Estimates of a Density Function // *Annals of Mathematical Statistics*. 1956. No. 27. P. 832-837.

SUPERMARKETS CLIENTS BEHAVIOUR FORECASTING BY WEIGHTED METHODS OF PROBABILITY AND DENSITY ESTIMATIONS

Alexander D'YAKONOV

Professor, Department of Mathematical Methods of Forecasting, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University;

Senior Researcher, Dorodnitsyn Computing Center, Russian Academy of Sciences

Address: 1, build. 52, Lomonosov Moscow State University, Leninskie Gory, GSP-1, Moscow, 119991, Russian Federation

E-mail: djakonov@mail.ru

We consider two tasks in describing a supermarkets clients' behavior: prediction of a client's next visit date and prediction of his/her spends. The first problem is equal to estimating visit probability, and the second – to estimating density for visitor spends. To solve these problems, we propose using weighed methods: real non-negative value (weight) is assigned to every event. Weights allow considering additional information, for example history (earlier visits have smaller weights). We consider several weighted schemes (methods of assigning weights to events) and weights optimization (performance optimization by changing weight parameters). The paper shows that weighted methods don't lead to overfitting, i.e. learning on a training set doesn't decrease performance on an independent test set. We can see, that assemblers of different methods can increase performance (we consider linear combination of probabilities estimated by different methods). All experiments are made on real data of large International competition on data mining. The last span of statistics does not contain holidays, which allows concentrating only on statistical methods of problems solving while solving these tasks. Besides, we also considered construction of algorithm to solve the problems (next visit date and spends prediction) simultaneously. It can be seen that the problems not always can be solved independently. We propose a function to estimate solutions of both problems and optimization method for this function.

Key words: forecasting, probability estimation, density estimation, non-parametric methods, applied problems.

References

1. Kaggle (2011) Dunnhumby's Shopper Challenge (electronic resource). Available at: <http://www.kaggle.com/c/dunnhumbychallenge> (accessed 24 November 2013).
2. Vorontsov K.V. Metricheskie algoritmy klassifikacii [Metrical Algorithms of Classification] (electronic lecture). Available at: www.ccas.ru/voron/download/MetricAlgs.pdf (accessed 24 November 2013). (in Russian)
3. Vorontsov K.V. Algoritmicheskie kompozicii [Algorithmic Compositions] (electronic lecture). Available at: <http://www.ccas.ru/voron/download/Composition.pdf> (accessed 24 November 2013). (in Russian)
4. D'yakonov A.G. (2013) Reshenie zadach analiza dannyh, osnovannoe na linejnoj kombinacii deformatsij [Data Mining Problems Solving by Using of Linear Combinations of Deformations]. *Machine learning and data analysis*, vol. 1, no. 5, pp. 543–554. (in Russian)
5. Vorontsov K.V. (2010) Kombinatornaja teorija nadjozhnosti obuchenija po precedentam [The Combinatory Theory of Machine Learning Reliability] (PhD Thesis), Moscow: Dorodnitsyn Computing Centre of RAS. (in Russian)
6. Rasin D. (2008) Neparametricheskaja ekonometrika: vvodnyj kurs [Non-parametric Econometrics: Introduction]. *Kvantil*, no. 4, pp. 7-56. (in Russian)
7. Parzen E. (1962) On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, no. 33, pp. 1065-1076.
8. Rosenblatt M. (1956) Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, no. 27, pp. 832-837.

**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.

Представляемая для публикации статья должна быть актуальной, обладать новизной, отражать постановку задачи (проблемы), описание основных результатов исследования, выводы, а также соответствовать указанным ниже правилам оформления.

Текст должен быть тщательно вычитан автором, который несет ответственность за научно-теоретический уровень публикуемого материала.

Материалы представляются в электронном виде по адресу:
bijournal@hse.ru.

ТРЕБОВАНИЯ К ОФОРМЛЕНИЮ СТАТЕЙ

ТЕКСТ СТАТЬИ представляется в редакцию в электронном виде (в формате MS Word, версия 2003 или выше).

ОБЪЕМ. Ориентировочный объем статьи составляет 20-25 тысяч знаков (с пробелами).

ШРИФТ, ФОРМАТИРОВАНИЕ, НУМЕРАЦИЯ СТРАНИЦ

ШРИФТ – Times New Roman, кегль набора – 12 пунктов, полуторный интервал, форматирование по ширине. Нумерация страниц – сверху по центру, поля: левое – 2,5 см, верхнее, нижнее и правое – по 1,5 см.

НАЗВАНИЕ СТАТЬИ приводится на русском и английском языках. Название статьи должно быть информативным и раскрывать содержание статьи.

СВЕДЕНИЯ ОБ АВТОРАХ приводятся на русском и английском языках и включают следующие элементы:

- ◆ фамилия, имя, отчество всех авторов полностью
- ◆ должность, звание, ученая степень каждого автора
- ◆ полное название организации – места работы каждого автора в именительном падеже, полный почтовый адрес каждой организации (включая почтовый индекс)
- ◆ адрес электронной почты каждого автора.

АННОТАЦИЯ К СТАТЬЕ представляется на русском и английском языках.

- ◆ Объем – 200-300 слов.
- ◆ Аннотация должна быть информативной (не содержать общих слов).
- ◆ Аннотация должна отражать основное содержание статьи и быть структурированной (следовать логике описания результатов в статье).
- ◆ Структура аннотации: предмет, цель, метод или методологию проведения исследования, результаты исследований, область их применения, выводы.
- ◆ Метод или методологию проведения исследований целесообразно описывать в том случае, если они отличаются новизной или представляют интерес с точки зрения данной работы. В аннотациях статей, описывающих экспериментальные работы, указывают источники данных и характер их обработки.
- ◆ Результаты работы описывают предельно точно и информативно. Приводятся основные теоретические и экспериментальные результаты, фактические данные, обнаруженные взаимосвязи и закономерности. При этом отдается предпочтение новым результатам и дан-

ным долгосрочного значения, важным открытиям, выводам, которые опровергают существующие теории, а также информации, которая, по мнению автора, имеет практическое значение.

- ◆ Выводы могут сопровождаться рекомендациями, оценками, предложениями, гипотезами, описанными в статье.
- ◆ Сведения, содержащиеся в названии статьи, не должны повторяться в тексте аннотации. Следует избегать лишних вводных фраз (например, «автор статьи рассматривает...»).
- ◆ Исторические справки, если они не составляют основное содержание документа, описание ранее опубликованных работ и общеизвестные положения, в аннотации не приводятся.
- ◆ В тексте аннотации следует употреблять синтаксические конструкции, свойственные языку научных и технических документов, избегать сложных грамматических конструкций.
- ◆ В тексте аннотации следует применять значимые слова из текста статьи.

КЛЮЧЕВЫЕ СЛОВА приводятся на русском и английском языках. Количество ключевых слов (словосочетаний) – 6-10. Ключевые слова или словосочетания отделяются друг от друга точкой с запятой.

ФОРМУЛЫ. При наборе формул, как выключных, так и строчных, должен быть использован редактор формул MS Equation. В формульных и символических записях греческие (русские) символы, а также математические функции записываются прямыми шрифтами, а переменные аргументы функций в виде английских (латинских) букв – наклонным курсивом (пример “cos a”, “sin b”, “min”, “max”). Нумерация формул – сквозная (по желанию авторов допускается двойная нумерация формул с указанием структурного номера раздела статьи и, через точку, номера формулы в разделе).

РИСУНКИ (графики, диаграммы и т.п.) могут быть оформлены средствами MS Word или MS Excel. Ссылки на рисунки в тексте обязательны и должны предшествовать позиции размещения рисунка. Допускается использование графического векторного файла в формате wmf/emf или cdr v. 10. Фотографические материалы предоставляются в формате TIF или JPEG, с разрешением изображения не менее 300 точек на дюйм. Нумерация рисунков – сквозная.

ТАБЛИЦЫ оформляются средствами MS Word или MS Excel. Нумерация таблиц – сквозная.

СПИСОК ЛИТЕРАТУРЫ составляется в соответствии с требованиями ГОСТ 7.0.5-2008. Библиографическая ссылка (примеры оформления размещены на сайте журнала <http://bi.hse.ru/>). Нумерация библиографических источников – в порядке цитирования. Ссылки на иностранную литературу – на языке оригинала без сокращений.

СПИСОК ЛИТЕРАТУРЫ ДЛЯ АНГЛОЯЗЫЧНОГО БЛОКА оформляется в соответствии с требованиями SCOPUS (примеры оформления размещены на сайте журнала <http://bi.hse.ru/>). Для транслитерации русскоязычных наименований можно воспользоваться сервисом <http://translit.ru/>.

ЛИЦЕНЗИОННЫЙ ДОГОВОР

Для размещения полнотекстовых версий статей на сайте журнала с авторами заключается лицензионный договор о передаче авторских прав.

Плата с авторов за публикацию рукописей не взимается.

AUTHORS GUIDELINES

Articles should be topical and original, should outline tasks (issues), describe key results of the author's research and appropriate conclusions.

Manuscripts are submitted via e-mail: bijournal@hse.ru.

MANUSCRIPT REQUIREMENTS

TEXT FILES should be submitted in electronic form, as a MS Word document (version 2003 or higher).

LENGTH. Articles should be between 20 and 25 thousand characters (incl. spaces).

FONT, SPACING, MARGINS. The text should be in Times New Roman 12 pt, 1.5 spaced, fit to the width, margins: left – 25 mm, all other – 15 mm.

TITLE of the article should be submitted in native language and English.

AUTHORS' DETAILS are presented in native language and English. The details include:

- ◆ Full name of each author
- ◆ Position, rank, academic degree of each author
- ◆ Affiliation of each author, at the time the research was completed
- ◆ Full postal address of each affiliation (incl. postcode / ZIP)
- ◆ E-mail address of each author.

ABSTRACT are presented in native language and English.

- ◆ The abstract should be between 200 and 300 words.
- ◆ The abstract should be informative (no general words), original, relevant (reflects your paper's key content and research findings); structured (follows the logics of results' presentation in the paper)
 - ◆ The recommended structure: purpose (mandatory), design / methodology / approach (mandatory), findings (mandatory), research limitations / implications (if applicable), practical implications (if applicable), originality / value (mandatory).
- ◆ It is appropriate to describe the research methods/methodology

if they are original or of interest for this particular research. For papers concerned with experimental work the data sources and data procession technique should be described.

- ◆ The results should be described as precisely and informatively as possible. Include your key theoretical and experimental results, factual information, revealed interconnections and patterns. Give special priority in the abstract to new results and long-term impact data, important discoveries and verified findings that contradict previous theories as well as data that you think have practical value.

- ◆ Conclusions may be associated with recommendations, estimates, suggestions, hypotheses described in the paper.

- ◆ Information contained in the title should not be duplicated in the abstract. Authors should try to avoid unnecessary introductory phrases (e.g. "the author of the paper considers...").

- ◆ Authors should use the language typical of research and technical documents to compile your abstract and avoid complex grammatical constructions.

- ◆ The text of the abstract should include key words of the paper.

KEYWORDS are presented in native language and English. The number of key words / words combinations are from 6 to 10 (separated by semicolons).

FORMULAE should be prepared using MS Equation tool.

FIGURES should be of high quality, legible and numbered consecutively with Arabic numerals. All figures (charts, diagrams, etc.) should be submitted in electronic form (photo images – in TIF or JPEG formats, minimum resolution 300 dpi). Appropriate references in the text are required.

REFERENCES should be presented in Harvard style and carefully checked for completeness, accuracy and consistency.

The publication is free of charge.