

ISSN 1998-0663 (print), ISSN 2587-8166 (online)

English version: ISSN 2587-814X (print), ISSN 2587-8158 (online)

БИЗНЕС- ИНФОРМАТИКА

НАУЧНЫЙ ЖУРНАЛ НИУ ВШЭ

СОДЕРЖАНИЕ

Моделирование социальных и экономических систем

Т.-М. Брассер, А. Младенов, К. Штраус

Открытые инновации в области бизнес-моделей:
обзор литературы и направления дальнейших
исследований..... 7

*Ю.М. Акаткин, О.Э. Карпов, В.А. Коняевский,
Е.Д. Ясиновская*

Цифровая экономика: концептуальная архитектура
экосистемы цифровой отрасли 17

Информационные системы и технологии в бизнесе

З.К. Авдеева, А.А. Утробин, И.Ю. Лыков

Совершенствование процесса закупок компании
с применением интегрированной рекомендательной
системы проведения тендера..... 29

С. Митрович

Специфика интеграции технологий бизнес-интеллекта
и больших данных в процессы экономического анализа..... 40

В.И. Жуков, М.М. Комаров

Использование системы веб-аналитики как основы
для интеграции с СРА-сервисами 47

Математические методы и алгоритмы бизнес-информатики

Ж.Н. Зенкова, Е.А. Крайнова

Оценка чистой нетто-премии с использованием
дополнительной информации о квантиле
функции распределения 55

С.Ю. Лобанова, А.А. Чеповский

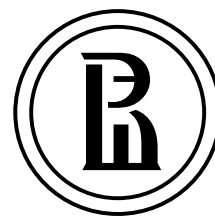
Комбинированный алгоритм выделения сообществ
в графах взаимодействующих объектов 64

Моделирование и анализ бизнес-процессов

А.Г. Мадера

Моделирование и оптимизация бизнес-процессов
и процессных систем в условиях неопределенности 74

№4(42)-2017



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*

*Заместитель главного редактора
Е.А. Кучерявый*

*Компьютерная верстка
О.А. Богданович*

*Администратор веб-сайта
И.И. Хрусталева*

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:
русскоязычная версия — 300 экз.,
англоязычная версия — 300 экз.,
онлайн-версии на русском и английском —
свободный доступ

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 школой бизнес-информатики НИУ ВШЭ. Журнал выпускается ежеквартально.

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:

- ◆ анализ данных и интеллектуальные системы
- ◆ информационные системы и технологии в бизнесе
- ◆ математические методы и алгоритмы бизнес-информатики
- ◆ программная инженерия
- ◆ Интернет-технологии
- ◆ моделирование и анализ бизнес-процессов
- ◆ стандартизация, сертификация, качество, инновации
- ◆ правовые вопросы бизнес-информатики
- ◆ принятие решений и бизнес-интеллект
- ◆ моделирование социальных и экономических систем
- ◆ информационная безопасность.

В соответствии с решением президиума Высшей аттестационной комиссии Российской Федерации 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, по следующим группам научных специальностей: 05.13.00 — информатика, вычислительная техника и управление; 05.25.00 — документальная информация; 08.00.00 — экономические науки.

Журнал входит в базу Russian Science Citation Index (RSCI) на платформе Web of Science.

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор), свидетельство ПИ № ФС77-66609 от 08 августа 2016 г.

Международный стандартный серийный номер (ISSN): 1998-0663 (на русском), 2587-814X (на английском).

Главный редактор: Голосов Алексей Олегович, кандидат технических наук, Президент компании «ФОРС — Центр разработки».

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

ГЛАВНЫЙ РЕДАКТОР

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

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА

Кучерявый Евгений Андреевич —
PhD, профессор департамента электроники и коммуникаций,
Технологический университет Тампере, Финляндия

ЧЛЕНЫ РЕДКОЛЛЕГИИ

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

Авдошин Сергей Михайлович —
кандидат технических наук, профессор, руководитель
департамента программной инженерии, Национальный
исследовательский университет «Высшая школа экономики»

Акопов Андраник Сумбатович —
доктор технических наук, профессор кафедры бизнес-аналити-
ки, Национальный исследовательский университет
«Высшая школа экономики»

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

Афанасьев Антон Александрович —
доктор экономических наук, и.о. ведущего научного сотрудника
лаборатории социального моделирования, Центральный эконо-
мико-математический институт РАН

Бабкин Эдуард Александрович —
кандидат технических наук, PhD, профессор кафедры
информационных систем и технологий, Национальный
исследовательский университет «Высшая школа экономики»

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

Баранов Александр Павлович —
доктор физико-математических наук, заместитель директора
ФГУП «Главный научно-исследовательский вычислительный
центр Федеральной налоговой службы»

Беккер Йорг —
PhD, проректор, профессор, директор Европейского исследо-
вательского центра в области информационных систем (ERCIS)
Мюнстерского университета, Мюнстер, Германия

Белов Владимир Викторович —
доктор технических наук, профессор кафедры вычислительной
и прикладной математики, 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, 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

Дискин Иосиф Евгеньевич —
доктор экономических наук, научный руководитель, председатель
Научно-экспертного совета, Всероссийский центр изучения
общественного мнения (ВЦИОМ); член Совета Общественной
палаты Российской Федерации

Зандкуль Курт —
PhD, заведующий кафедрой информационных систем
для бизнеса, Университет Росток, Германия

Ильин Николай Иванович —
доктор технических наук, член-корреспондент Академии крипто-
графии РФ, заместитель начальника Управления специальной связи,
Федеральная служба охраны Российской Федерации (ФСО России)

Исаев Дмитрий Валентинович —
кандидат экономических наук, доцент кафедры бизнес-аналитики,
Национальный исследовательский университет
«Высшая школа экономики»

Калягин Валерий Александрович —
доктор физико-математических наук, профессор, заведующий
кафедрой прикладной математики и информатики, Национальный
исследовательский университет «Высшая школа экономики»

Каменнова Мария Сергеевна —
кандидат технических наук, Генеральный директор компании
«Логика BPM»

Кравченко Татьяна Константиновна —
доктор экономических наук, профессор, заведующая кафедрой
бизнес-аналитики, Национальный исследовательский
университет «Высшая школа экономики»

Кузнецов Сергей Олегович —
доктор физико-математических наук, профессор, руководитель
департамента анализа данных и искусственного интеллекта,
Национальный исследовательский университет «Высшая школа
экономики»

Лугачев Михаил Иванович —
доктор экономических наук, профессор, заведующий кафедрой
экономической информатики, Московский государственный
университет им. М.В. Ломоносова

Мальцева Светлана Валентиновна —
доктор технических наук, профессор, заведующая кафедрой
инноваций и бизнеса в сфере информационных технологий,
Национальный исследовательский университет
«Высшая школа экономики»

Мейор Питер —
PhD, заместитель директора консультативной группы по радио-
коммуникациям, Международный телекоммуникационный союз
(ITU), заместитель руководителя Комиссии ООН по науке
и технологиям, Женева, Швейцария

Миркин Борис Григорьевич —
доктор технических наук, профессор департамента анализа данных
и искусственного интеллекта, Национальный исследовательский
университет «Высшая школа экономики»

Моттль Вадим Вячеславович —
доктор технических наук, профессор кафедры информационной
безопасности, Тульский государственный университет

Пальчунов Дмитрий Евгеньевич —
доктор физико-математических наук, заведующий кафедрой
общей информатики, 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, Национальный исследовательский
технологический университет «МИСиС»

Таратухин Виктор Владимирович —
кандидат технических наук, PhD, руководитель научной группы
Европейского исследовательского центра в области
информационных систем (ERCIS), Мюнстерский университет,
Мюнстер, Германия

Ульянов Михаил Васильевич —
доктор технических наук, профессор департамента программной
инженерии, Национальный исследовательский университет
«Высшая школа экономики»

ISSN 1998-0663 (print), ISSN 2587-8166 (online)

English version: ISSN 2587-814X (print), ISSN 2587-8158 (online)

BUSINESS INFORMATICS

HSE SCIENTIFIC JOURNAL

CONTENTS

Modeling of social and economic systems

T.-M. Brasseur, A. Mladenow, C. Strauss

Open business model innovation: Literature review
and agenda for future research 7

*Y.M. Akatkin, O.E. Karpov, V.A. Konyavskiy,
E.D. Yasinovskaya*

Digital economy: Conceptual architecture
of a digital economic sector ecosystem 17

Information systems and technologies in business

Z.K. Avdeeva, A.A. Utrobin, I.Y. Lykov

Improvement of the procurement process using
the integrated tender recommendation system..... 29

S. Mitrovic

Specifics of the integration of Business Intelligence
and Big Data technologies in the processes
of economic analysis..... 40

V.I. Zhukov, M.M. Komarov

Using a web analytics system as the basis
for integration with CPA services 47

Mathematical methods and algorithms of business informatics

Z.N. Zenkova, E.A. Krainova

Estimating the net premium using additional
information about a quantile of the cumulative
distribution function 55

S.Y. Lobanova, A.A. Chepovskiy

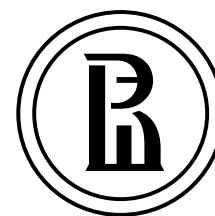
Combined method to detect communities
in graphs of interacting objects 64

Business processes modeling and analysis

A.G. Madera

Modeling and optimization of business processes
and process systems under conditions of uncertainty 74

№4(42)-2017



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

Deputy Editor-in-Chief

Y. Koucheryavy

Computer Making-up:

O. Bogdanovich

Website Administration:

I. Khrustaleva

Address:

33, Kirpichnaya Street, Moscow,
105187, Russian Federation

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

<http://bijournal.hse.ru>

E-mail: bijournal@hse.ru

Circulation:

*English version – 300 copies,
Russian version – 300 copies,
online versions in English and Russian –
open access*

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 School of Business Informatics. The journal is published quarterly.

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:

- ◆ data analysis and intelligence systems
- ◆ information systems and technologies in business
- ◆ mathematical methods and algorithms of business informatics
- ◆ software engineering
- ◆ Internet technologies
- ◆ business processes modeling and analysis
- ◆ standardization, certification, quality, innovations
- ◆ legal aspects of business informatics
- ◆ decision making and business intelligence
- ◆ modeling of social and economic systems
- ◆ information security.

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 included into Russian Science Citation Index (RSCI) database on the Web of Science platform.

International Standard Serial Number (ISSN): 2587-814X (in English), 1998-0663 (in Russian).

Editor-in-Chief: Dr. Alexey Golosov – President of FORS Development Center, Moscow, Russian Federation.

EDITORIAL BOARD

EDITOR-IN-CHIEF

Alexey Golosov —

President of FORS Development Center, Russian Federation

DEPUTY EDITOR-IN-CHIEF

Yevgeni Koucheryavy —

Professor, Department of Electronics and Communication Engineering, Tampere University of Technology, Finland

EDITORIAL BOARD

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

Sergey Avdoshin —

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

Andranik Akopov —

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

Fuad Aleskerov —

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

Anton Afanasyev —

Leading Researcher, Laboratory of Social Modeling, Central Economics and Mathematics Institute, Russian Academy of Science, Russian Federation

Eduard Babkin —

Professor, Department of Information Systems and Technologies, National Research University Higher School of Economics, Russian Federation

Alex Bayer —

Head of KAFAN FX Information Services, New York, USA

Alexander Baranov —

Deputy Head of Central Scientific and Research Computing Center, Federal Tax Service of Russia, Russian Federation

Jorg Becker —

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

Vladimir Belov —

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

Andrey Gribov —

Director General, CyberPlat Company, Russian Federation

Alexander Gromov —

Professor, Head of Department of Modeling and Optimization of Business Processes, National Research University Higher School of Economics, Russian Federation

Vladimir Gurvich —

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

Laurence Jacobs —

Professor, Medical School, University of Zurich, Switzerland

Iosif Diskin —

Academic Supervisor, Chairmen of Scientific and Expert Council, Russian Public Opinion Research Center (VCIOM); Member of the Council, The Russian Public Chamber; Russian Federation

Kurt Sandkuhl —

Professor, Head of Department of Business Information Systems, University of Rostock, Germany

Nikolay Ilyin —

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

Dmitry Isaev —

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

Valery Kalyagin —

Professor, Head of Department of Applied Mathematics and Informatics, National Research University Higher School of Economics, Russian Federation

Maria Kamennova —

Director General, BPM Logic, Russian Federation

Tatiana Kravchenko —

Professor, Head of Department of Business Analytics, National Research University Higher School of Economics, Russian Federation

Sergey Kuznetsov —

Professor, Head of School of Data Analysis and Artificial Intelligence, National Research University Higher School of Economics, Russian Federation

Mikhail Lugachev —

Professor, Head of Department of Economic Informatics, Lomonosov Moscow State University, Russian Federation

Svetlana Maltseva —

Professor, Head of Department of Innovation and Business in Information Technologies, National Research University Higher School of Economics, Russian Federation

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

Boris Mirkin —

Professor, School of Data Analysis and Artificial Intelligence, National Research University Higher School of Economics, Russian Federation

Vadim Mottl —

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

Dmitry Palchunov —

Head of Department of General Informatics, Novosibirsk State University, Russian Federation

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

Albert Silantyev —

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

Victor Taratukhin —

Managing Director, European Research Center for Information Systems (ERCIS), University of Munster, Germany

Mikhail Ulyanov —

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

Открытые инновации в области бизнес-моделей: обзор литературы и направления дальнейших исследований

Т.-М. Брассер

преподаватель департамента электронного бизнеса
Венский университет
Адрес: Austria, 1090, Vienna, Oskar-Morgenstern-Platz, 1
E-mail: tiare.brasseur@gmail.com

А. Младенов

преподаватель департамента электронного бизнеса
Венский университет
Адрес: Austria, 1090, Vienna, Oskar-Morgenstern-Platz, 1
E-mail: andreas.mladenow@univie.ac.at

К. Штраус

профессор департамента электронного бизнеса
Венский университет
Адрес: Austria, 1090, Vienna, Oskar-Morgenstern-Platz, 1
E-mail: christine.strauss@univie.ac.at

Аннотация

В сегодняшней быстро меняющейся предпринимательской среде компании вынуждены постоянно вводить новшества, чтобы сохранить конкурентоспособность. Инновации в области бизнес-моделей (business model innovation, BMI) в последнее время привлекают все больше внимания в качестве одного из перспективных подходов для достижения сравнительных преимуществ в условиях жесткой конкуренции. Однако, несмотря на огромный потенциал, BMI также подразумевают чрезвычайно сложные процессы, крайнюю неопределенность и большие финансовые риски. Благодаря развитию цифровых технологий, в последние годы BMI стали более открытыми и основанными на сотрудничестве.

Целью настоящей работы является исследование роли и последствий открытых и основанных на сотрудничестве методик в области BMI, а также анализ имеющейся литературы по данной теме. Таким образом, проведен систематический анализ литературы на стыке открытых инноваций (open innovation, OI) и BMI. Анализ литературы выявил два основных направления исследований в области открытых инноваций в области бизнес-моделей (open business model innovation, OBMI): тренды OBMI (клиентоориентированные BMI, совместное создание бизнес-моделей, оценка бизнес-моделей на ранних стадиях, виртуальное сотрудничество, дизайнерское мышление) и результаты OBMI. В целом результаты исследования подтверждают усиление тенденций сотрудничества и совместного создания в области BMI с помощью цифровых и материальных инструментов, а также указывают на то, что открытые инновации имеют непосредственное положительное влияние на успех BMI. Анализ литературы также продемонстрировал, что область OBMI все еще является недостаточно исследованной.

Ключевые слова: инновации в области бизнес-моделей, открытые инновации в области бизнес-моделей, цифровое бизнес-моделирование, открытые инновации, сотрудничество, инфраструктура взаимодействия, совместное создание, клиентские инновации.

Цитирование: Брассер Т.-М., Младенов А., Штраус К. Открытые инновации в области бизнес-моделей: обзор литературы и направления дальнейших исследований // Бизнес-информатика. 2017. № 4 (42). С. 7–16. DOI: 10.17323/1998-0663.2017.4.7.16.

Введение

Не так давно под «инновациями» при решении задач реальной сложности подразумевались более высокие инвестиции во внутренние исследования и разработки (НИОКР), чем у конкурентов [1]. Более современные и совершенные стратегии инноваций основаны на совместных подходах с участием внешних партнеров, которые дополняют инновационный потенциал компании. В работе [1] впервые введено понятие открытых инноваций (open innovation, OI) как вида инноваций, при которых компании используют внешние источники знаний и рынки с целью ускорить внутренние инновации. В последние годы OI стали объектом более активного внимания со стороны исследователей [1–10] и средств массовой информации. Общее мнение исследователей OI заключается в том, что повторное использование одних и тех же знаний имеет ограниченную ценность и может сдерживать инновационный потенциал компании [1, 2, 7]. Таким образом, парадигма OI подталкивает компании к тому, чтобы сделать свои процессы открытыми для внешнего мира и предусматривает более коллективный и менее централизованный подход к инновациям по сравнению с ранее преобладавшей парадигмой «закрытых инноваций» [1]. Понятие OI пересекается с множеством других понятий, таких как открытый исходный код, краудсорсинг, совместное создание, инновационная деятельность пользователей и другие формы внешней инновационной деятельности [11].

OI позволяют компаниям более широко использовать имеющиеся таланты и навыки, снизить стоимость НИОКР, получить дополнительные источники дохода (путем лицензирования, в рамках совместных предприятий и т.п.) и сократить сроки вывода на рынок новых продуктов [12–14]. Объединив внешние и внутренние источники инноваций, компании получают доступ к новым инновациям, покупателям, способам коммерциализации и рынкам [12, 13]. На практике такие компании, как Lego, 3M, P&G, IBM и Dell, продемонстрировали способность OI способствовать инновационной деятельности и повышению финансовых результатов.

Как и OI, BMI оказался перспективным подходом, который позволяет избежать жесткой конкуренции, особенно в условиях высокой схожести продуктов и процессов [15, 16]. В литературе утверждается, что бизнес-модели (business models, BMs) стали новой основой конкуренции [17] и, возможно, являются наиболее сложным типом инноваций. При разработке новой бизнес-модели компании

сталкиваются с высокой неопределенностью, а также с чрезвычайно сложными процессами [18]. Любое изменение в бизнес-модели может принести миллиарды долларов прибыли или привести к краху [19]. Таким образом, высокая степень риска, связанная с BMI, связана с серьезным масштабом реорганизации деятельности. В то же время BMI все чаще признается главной движущей силой сверхприбылей и главным источником устойчивых конкурентных преимуществ [1, 20]. В связи с этим проведен систематический анализ литературы, представленной на пяти ведущих научных площадках (ACM, EBSCO, IEEE, Springer Link, Wiley), для того, чтобы сформировать общий обзор имеющейся литературы о BMI и определить основные темы и подходы, обсуждаемые в данной сфере.

1. Тренды и результаты OBM

Анализ литературы выявил два основных направления исследований в области OBM: тренды и результаты OBM. В имеющейся литературе рассматриваются различные подходы и инструментарии: от теоретических моделей до (открытого) программного обеспечения для виртуального сотрудничества и от игровых методов дизайнерского мышления до передовых BMI. Кроме того, были определены пять новых тенденций в области OBM. В целом, в литературе прослеживается тенденция к формам BMI, в большей степени основанным на взаимодействии, и предлагается использование различных подходов и инструментов для интеграции внешних партнеров в деятельность, связанную с OBM (таблица 1).

Во всех исследованиях при рассмотрении участия партнеров цепочки снабжения в деятельности, связанной с OBM, клиенты включены в процесс BMI. Это неудивительно, поскольку ключевые слова, используемые для поиска литературы в систематическом каталоге, нацелены на выявление статей, изучающих участие целевой группы, то есть клиентов, в методиках OBM [2]. Кроме того, в 5 из 11 публикаций предлагается сотрудничать как с клиентами, так и с поставщиками для того, чтобы продвигать BMI. Помимо этого, в одной из работ предложено вовлечь клиентов, поставщиков и исследовательское учреждение в коллективную работу по OBM [21].

Учитывая возможную интенсивность участия партнеров по цепи поставок в деятельности, связанной с OBM, Международная ассоциация общественного участия (International Association for Public Participation) [22] разработала шкалу уча-

Таблица 1.

Анализ литературы

Источник	Тренды ОВМИ					Результаты ОВМИ	Инструментарий			
	Совместное создание бизнес–моделей	Клиентоориентированные ВМИ	Оценка бизнес–моделей на ранних стадиях	Виртуальное сотрудничество	Дизайнерское мышление		Программное обеспечение	Осязаемые инструменты	Теоретическая модель	Дополнительная информация
Berre et al. (2013) [27]	X		X	X			X			- Облачная платформа для ОВМИ - Пилотное тестирование предпринимательских платформ - Инновационное сообщество на основе социальных сетей
Buur and Gudiksen (2012) [21]	X	X			X			X		- Эксперименты с дизайном бизнес–моделей с использованием «осязаемых» моделей - Интерактивный дизайн на основе концепции пинбола (игрофикация) - Клиентская сторона бизнес–моделей
Chew (2015) [23]	X	X	X	X			X		X	- Модель интегрированного метода инноваций в сфере услуг (iSIM) для одновременного внедрения инноваций в сфере услуг и ВМИ - Созданная совместно с клиентом ценность (ценность пользователя) - Проектирование бизнес–моделей «от обратного» (начиная с клиентов)
Denicolai et al. (2014) [33]						X				- Влияние внешних источников знаний на рост компании (как критерий успеха ВМИ) - Анализ ежегодного доклада (310 европейских компаний)
Ebel et al. (2016) [24]	X	X	X	X			X		X	- Пилотное тестирование средств разработки бизнес–моделей с участием тестовых пользователей - На основе литературы по онлайн–взаимодействию - Интеграция целевой группы с экспертами и заинтересованными лицами
Gudiksen (2015) [32]	X				X			X		- Эксперименты с проектированием бизнес–моделей с использованием игровых элементов (геймификация) - Ролевые игры, использование осязаемых ресурсов - Коллективный подход к инновациям
Huang et al. (2013) [13]						X			X	- Влияние ОI на ВМИ и организационную инерцию - Влияние ВМИ на деятельность компании - Исследования среди 141 предприятий малого и среднего бизнеса (в Тайване)
Ogilvie (2015) [25]	X		X		X			X	X	- Совместное создание бизнес–моделей (клиенты и отраслевые партнеры) - Оценка бизнес–моделей и моделей получения дохода на ранних стадиях - Модель пяти «В», прототипы бизнес–моделей (макеты)
Pynnönen et al. (2012) [29]		X	X	X			X		X	- Управление клиентоориентированными бизнес–моделями - Участие клиентов в ВМИ и оценке бизнес–моделей - Инновационные сообщества на основе социальных сетей
Trimi et al. (2012) [26]		X	X						X	- Эксперименты с бизнес–моделями и гибкость бизнес–моделей - Процесс изучения клиента (оценка) - Гибкая философия (проверка гипотез бизнес–моделей)
Zolnowski et al. (2014) [28]	X	X							X	- Клиент как часть (нового) дизайна бизнес–моделей - Влияние клиента на другие параметры бизнес–моделей - Канва бизнес–моделей (точка зрения компании / клиентов / партнеров)

ствия заинтересованных сторон, чтобы продемонстрировать возможные типы (т.е. уровни) участия заинтересованных сторон. С учетом этой шкалы в литературе по ОВМІ можно выделить два основных подхода к привлечению участников: «задействованность» и «сотрудничество». В большей части имеющейся литературы (в 7 из 11 публикаций) выбор падает на подход на основе сотрудничества, который позволяет участникам цепи поставок активно привносить новые идеи в бизнес-модели. С другой стороны, в двух публикациях рассматривался подход, включавший в себя добавление, изменение или перегруппировку знаний, предоставляемых партнерами по цепи поставок.

2. Совместное создание бизнес-моделей

Вопросы совместного создания бизнес-моделей наиболее часто обсуждаются в литературе по ОВМІ [23–26]. В работе [24] совместное создание бизнес-моделей определяется как «разработка бизнес-моделей совместно с клиентами на основе сотрудничества». Другими словами, совместное создание бизнес-моделей поощряет активное участие партнеров по цепи поставок в создании новых моделей. Кроме того, описывается новая инновационная парадигма, которая включает в себя сотрудничество на основе общей платформы [26]. Соответственно, совместные инновации предусматривают создание уникальных ценностей и накопление опыта совместно с внешними заинтересованными сторонами.

Анализ литературы показал, что переход к открытому процессу ВМІ для совместного создания бизнес-моделей открывает новые возможности для бизнеса [27] и повышает качество разработанных моделей [24]. Утверждается, что использование внешних коммуникаций и сотрудничество с клиентами в целях обмена сведениями о бизнес-моделях компании позволяют протестировать коммерческую жизнеспособность новой модели до начала ее применения на практике [23, 24]. В частности, в литературе подчеркивается роль клиентов и поставщиков в качестве ценных участников процесса ВМІ. В работе [25] утверждается, что непосредственное взаимодействие с клиентами и потенциальными партнерами помогает совместно искать решения, которые приносят прибыль всем заинтересованным лицам. Авторы работы [21] приходят к заключению, что «обсуждение бизнес-моделей внутри компании, а также с поставщиками и

клиентами необходимо для того, чтобы обеспечить конкурентные преимущества компании». Аналогичным образом, в работе [24] подчеркивается, что участие целевой группы имеет решающее значение для успеха ВМІ. Рассматривая совместное создание с точки зрения клиента, автор работы [25] заявляет, что клиенты с воодушевлением принимают приглашение присоединиться к совместному созданию бизнес-моделей и с энтузиазмом относятся к идее совместного поиска новых решений [25].

Кроме того, высказывается мнение о том, что «инновации в сфере услуг и инновации в области бизнес-моделей обеспечивают рост бизнеса» [23]. Автор представляет прототип нового метода проектирования — интегрированного метода инноваций в сфере услуг (integrated service innovation method, iSIM), для одновременного внедрения инноваций в сфере услуг и проектирования ВМІ. Эта модель основана на предположении, что создание ценности совместно с клиентом является центральным элементом ВМІ. Таким образом, модель iSIM упрощает привлечение клиентов как сопроизводителей ценности. Совместное создание ВМІ повышает ценность предложения как для клиентов, так и для бизнес-экосистемы поставщиков [23, 28]. Аналогичным образом рекомендуется обсуждать и уточнять новые варианты бизнес-моделей с клиентами и поставщиками для достижения согласия между всеми заинтересованными сторонами [24].

Для практического взаимодействия компаниям предлагается создать внутреннюю культуру «открытого лидерства» и «организационного обучения» [23]. С помощью оперативных экспериментов с бизнес-моделями компании могут проверять коммерческую жизнеспособность новых моделей до инвестирования в их проектирование и внедрение [23]. В работе [24] подчеркивается определенная гибкость в отношении ресурсов и возможностей бизнес-моделей, чтобы гарантировать, что они смогут быть адаптированы к меняющемуся рынку. Авторы работы [28] подчеркивают, что для того, чтобы способствовать интеграции внешних партнеров и взаимодействию с ними важно принять «логику сетевой ценности», а не сосредотачивать все внимание на отдельной цепочке создания ценности. Кроме того, подчеркивается сочетание интеграции ресурсов и потенциала конфигурации бизнеса в целях использования внешних знаний во внутренних процессах (потенциал освоения) [23].

3. Клиентоориентированные ВМІ

Еще один аспект — необходимость в «клиентоориентированной модели» ВМІ [26]. Утверждается, что «компания не обязательно знает ценностные предпочтения своих клиентов» [29], поэтому они должны признать клиентов ценными участниками, способными выступить в различных ролях в процессе ВМІ. Авторы работы [28] считают клиента «отправной точкой» ВМІ, поместив его в верхней части бизнес-модели. Визуализация всех потенциальных точек взаимодействия с клиентом помогает определить, как клиент участвует в определении одних параметров бизнес-моделей и влияет на другие [28]. В работе [23] предлагается сначала представить, каким должен быть уникальный опыт клиента, а затем, отталкиваясь от этого, приступить к проектированию бизнес-моделей и предоставлению услуг. В свою очередь, авторы работы [29] предлагают четырехэтапную систему бизнес-процессов для анализа соответствия между текущей бизнес-моделью компании и ее ценностью для клиента [29]. Каждый из элементов бизнес-модели оценивается в соответствии с ценностными предпочтениями клиентов, для определения основных и второстепенных компонентов модели и исключить элементы, не имеющие ценности [29]. Авторы работы [26] особо подчеркивают, насколько важно получить одобрение клиента прежде, чем приступить к созданию бизнес-моделей. При этом они обращаются к идее изучения клиента. Основной принцип процесса изучения клиента заключается в том, что новый подход или идея должны быть одобрены клиентами для использования в последующих процессах, иначе они будут направлены на доработку. В отношении ВМІ одобрение клиента включает в себя проверку элементов бизнес-моделей, таких как субъективная ценность предложения, или целесообразность цен или каналов распространения [26]. Наконец, в работе [29] подчеркивается важность непрерывного участия клиентов в процессе ВМІ для того, чтобы постоянно и итеративно согласовывать бизнес-модели компании с текущими и возникающими потребностями рынка и получать в реальном времени важную информацию обо всех изменениях в предпочтениях клиентов.

4. Оценка бизнес-моделей на ранних стадиях

Рассматривая вопрос оценки бизнес-моделей на ранних стадиях, исследователи сходятся во мнении,

что использование внешних отзывов и целевой группы на ранних этапах создания модели имеет жизненно важное значение для успеха ВМІ [24, 26]. Таким образом, с помощью оценки бизнес-моделей на ранних стадиях компании могут проверить коммерческую жизнеспособность новых моделей до осуществления вложений в их дальнейшую разработку [23, 25]. Оценка бизнес-моделей на ранних стадиях позволяет компаниям снизить высокую степень неопределенности и риски, связанные с ВМІ. Считается, что разработка методов оценки бизнес-моделей на ранних стадиях является ключом к успеху [25]. Таким образом, благодаря своевременному и постоянному взаимодействию с клиентами компании могут увеличить свои шансы на успех [26]. Следовательно, бизнес-модели должны быть достаточно гибкими, чтобы дать возможность накопления новых знаний методом проб и ошибок, начиная с первых оценок. В частности, для молодых компаний решающее значение имеет развитие потенциала для быстрого тестирования и подтверждения бизнес-гипотез [26]. В работе [25] предлагается использовать методы визуализации и дизайнерского мышления при совместном создании бизнес-моделей с клиентами и потенциальными партнерами. Телематическая служба использовала визуальные прототипы бизнес-моделей (плакаты), разместив на них «ровно столько информации, сколько требовалось клиентам, чтобы понять, как, возможно, будет работать БМ», но при этом оставив достаточно места для комментариев и идей участников рабочего совещания [25]. Этот новаторский подход позволил компании оценить и переработать новую бизнес-модель и модель получения дохода при минимальных затратах [25].

5. Виртуальное сотрудничество

Виртуальное сотрудничество является еще одной важной темой в области ОВМІ, выявленной в процессе анализа имеющейся литературы. Здесь подчеркивается важная роль информационных технологий для поддержки открытых и совместных подходов к созданию новых бизнес-моделей [24]. В этом отношении цифровые платформы ОВМІ нацелены на предоставление командам разработчиков бизнес-моделей поддержки инновационного сообщества при решении задач ВМІ. Таким образом, основная идея виртуального сотрудничества для ВМІ похожа на применение краудсорсинга, который использует коллективный потенциал для решения сложных задач [30, 31]. Инструменты циф-

рового бизнес-моделирования, представленные в литературе, включают в себя множество различных компонентов. К ним, например, относятся голо-сование, оценки, рейтинг, поиск и управление доступом, элементы сообщества (страницы профиля, группы по интересам), обмен сообщениями, комментирование и другие компоненты, обеспечивающие взаимодействие, такие как обмен файлами [23, 24, 27, 29].

6. Дизайнерское мышление

Пятой важной темой, представленной в литературе, является дизайнерское мышление. Дизайнерское мышление представляет собой различные методы коллективного подхода к инновациям, которые включают в себя эксперименты, осязаемые ресурсы и инновационную деятельность в игровом формате и побуждают к творческому подходу к поиску инновационных решений [21]. Подход «осязаемого бизнес-моделирования» позволил участникам «высказаться с помощью рук» [21]. В проанализированной литературе описано применение таких осязаемых ресурсов, как листочков для заметок [25] или шариков и кубиков [32] для того, чтобы содействовать взаимодействию между междисциплинарными командами, совместно создающими бизнес-модели. Авторы работы [21] провели итеративные эксперименты, связанные с бизнес-моделями, в различных интерактивных форматах для того, чтобы дать толчок новым обсуждениям и мышлению, направленному на поиск решений. Исследователи пришли к заключению, что «мышление в игровом формате с использованием рук и тела» улучшает «качество обсуждений в целях получения новаторских результатов» [21]. Аналогичным образом автор работы [32] выбрал различные игры, связанные с проектированием бизнес-моделей, для экспериментов с новыми идеями [32]. Было продемонстрировано, что подход к бизнес-моделированию как к дизайнерской задаче позволяет взглянуть на ВМІ с другой точки зрения. Использование осязаемых ресурсов и генератора случайных чисел (например, игральные кости) позволяет генерировать новые идеи и комбинации. Резюмируя, можно сказать, что дизайнерские игры, связанные с бизнес-моделями, имеют два главных преимущества: (1) они устанавливают четкие правила, которым обязаны следовать все участники и (2) обеспечивают непринужденную и располагающую атмосферу, которая позволяет участникам выйти за рамки реальной жизни и наслаждаться «свободой играть»

и обучением, основанным на экспериментах [32]. Чтобы успешно применять дизайнерское мышление для ВМІ, компаниям необходимо создать временное пространство для творчества и поощрять экспериментальную игровую культуру, а также свободу импровизировать, играть с бизнес-моделями и тестировать их [21, 32].

7. Результаты ОВМІ

В двух публикациях по ОВМІ рассматривается, как ОІ могут быть эффективными при создании ВМІ [13, 33]. В них использован количественный подход для изучения влияния внешних источников знаний на ВМІ и деятельность компании [13], а также рост компании как критерий успеха ВМІ [33]. Используя опросы, авторы работы [13] обнаружили, что ОІ могут быть эффективными для преодоления организационной инерции и, следовательно, создания ВМІ. Организационная инерция представляет собой стремление сохранить сложившуюся организационную структуру. Кроме того, организационная инерция оказывает негативное воздействие как на ОІ, так и на ВМІ [13]. Следовательно, исключительно важно преодолеть организационную инерцию для того, чтобы иметь возможность адаптироваться к быстро меняющимся потребностям рынка [13]. Кроме того, исследователи утверждают, что ОІ имеют непосредственное положительное влияние как на ВМІ, так и на деятельность компании.

Используя показатели учета как критерии получения сведений, авторы работы [33] стремятся понять, как можно объединить внешние и внутренние источники знаний для того, чтобы они способствовали росту компании, что являлось бы показателем успеха ВМІ. Для этого ученые исследовали взаимосвязь между знаниями, накопленными внутри организации (internally developed knowledge, KINT) и знаниями, полученными извне (externally acquired knowledge assets, EINT) при создании ценности. В исследовании выявлена инвертированная U-образная взаимосвязь между затратами на внешние источники знаний и выгодами от их использования, а также внутренним потенциалом освоения [33]. Таким образом, утверждается, что увеличение инвестиций во внешние источники знаний оказывает положительное влияние на рост компании, но только до определенного момента: после этого влияние становится негативным, поскольку компания сталкивается с трудностями в освоении такого

большого объема внешних знаний [33]. Исследователи делают вывод, что существует оптимальный уровень внешних знаний, который зависит от потенциала их освоения компанией и при котором рост компании, а значит и успех ВМІ, повышается до максимума.

Заключение

Анализ представленной нами литературы показывает, что область ОВМІ все еще является недостаточно исследованной, поскольку лишь 11 публикаций соответствуют критериям включения в этот список. Было установлено, что в имеющейся литературе рассматриваются различные тренды в области методик ОВМІ, а также их результаты. Анализ имеющейся литературы показывает, что ОІ оказывают непосредственное положительное влияние на успех ВМІ. Установлено, что вовлечение внешних партнеров, особенно клиентов, в процессы ВМІ может повысить как качество созданных бизнес-моделей, так и эффективность компании, по крайней мере, до определенного предела. Что касается открытости деятельности, связанной с ВМІ, то во всех 11 работах в процесс ВМІ были вовлечены клиенты, в то время как в некоторых из них в процессе продвижения ВМІ также принимают участие поставщики и исследовательские институты. Более того, в большинстве публикаций подчеркивается коллективная работа в области ВМІ, позволяющая внешним партнерам активно привносить в бизнес-модели новые идеи. Кроме того, определены пять основных тенденций в области ОВМІ.

Во-первых, тенденция совместного создания бизнес-моделей оказалась наиболее часто затрагиваемой темой, что свидетельствует о растущей важности сотрудничества в области ВМІ. Во-вторых, анализ литературы выявил необходимость в клиентоориентированной модели ВМІ. В-третьих, результаты обзора показывают, что создание возможностей для проверки новых бизнес-моделей на ранних стадиях имеет решающее значение для успеха ВМІ. Таким образом, предложено, чтобы компании как можно раньше

привлекали к участию целевые группы, чтобы провести оценку бизнес-моделей перед их непосредственным созданием. В-четвертых, в имеющейся литературе представлены различные методы виртуального сотрудничества в рамках ВМІ. Инструменты цифрового сотрудничества, представленные в литературе, опираются на методы краудсорсинга для ВМІ задач и подчеркивают ключевую роль инфокоммуникационных технологий в содействии процессам ОВМІ. В-пятых, использование дизайнерского мышления в ВМІ является еще одним из основных выводов обзора. Здесь показано, что подход к БМ как к дизайнерской задаче с помощью осязаемых и визуальных макетов бизнес-моделей, применяемых во время экспериментальной новаторской деятельности в игровом формате, повышает эффективность взаимодействия участников и тем самым увеличивает число новаторских идей. Утверждается, что использование методов дизайнерского мышления в ОВМІ помогает компаниям взглянуть на задачи, которые перед ними ставит бизнес-моделирование, под новым углом, а также позволяет им оценить модели и собрать новые идеи для бизнес-моделей при минимальных затратах

Можно сделать вывод, что ОВМІ является важной областью исследований, которая все еще находится на начальной стадии. Настоящая работа обеспечивает всесторонний обзор знаний в этой новой области исследований. Учитывая информацию об усилении тенденций сотрудничества в области ВМІ, дальнейшие исследования будут сосредоточены на изучении совместных подходов к ВМІ. Дальнейшие исследования также могут пролить свет на такие аспекты совместного создания бизнес-моделей, как, например, их применение на разных этапах инновационного процесса и направление информационных потоков. Кроме того, поскольку ожидается, что в будущем инструменты виртуального сотрудничества и методы дизайнерского мышления будут играть все более важную роль в содействии процессам ВМІ, они также представляют собой перспективные направления дальнейших исследований. ■

Литература

1. Chesbrough H. Business model innovation: Opportunities and barriers // Long Range Planning. 2010. Vol. 43. No. 2–3. P. 354–363.
2. Brasseur T.-M., Mladenow A., Strauss C. Business model innovation to support smart manufacturing // Proceedings of American Conference on Information Systems 2017. Workshop on Smart Manufacturing. Boston, USA, 10–12 August 2017. [Электронный ресурс]: <http://aisel.aisnet.org/sigbd2017/9> (дата обращения 16.10.2017).
3. Enkel E., Gassmann O., Chesbrough H. Open R&D and open innovation: Exploring the phenomenon // R&D Management. 2009. Vol. 39. No. 4. P. 311–316.
4. Gassmann O., Enkel E., Chesbrough H. The future of open innovation // R&D Management. 2010. Vol. 40. No. 3. P. 213–221.
5. Huizingh E.K. Open innovation: State of the art and future perspectives // Technovation. 2011. Vol. 31. No. 1. P. 2–9.

6. Mladenow A., Fröschl K.A. Kooperative forschung. Frankfurt am Main: Lang, 2011.
7. Laursen K., Salter A. Open for innovation: the role of openness in explaining innovation performance among U.K. manufacturing firms // *Strategic Management Journal*. 2006. Vol. 27. No. 2. P. 131–150.
8. Lopez-Vega H., Tell F., Vanhaverbeke W. Where and how to search? Search paths in open innovation // *Research Policy*. 2016. Vol. 45. No. 1. P. 125–136.
9. Prugl R., Schreier M. Learning from leading-edge customers at The Sims: Opening up the innovation process using toolkits // *R&D Management*. 2006. Vol. 36. No. 3. P. 237–250.
10. West J., Bogers M. Leveraging external sources of innovation: a review of research on open innovation // *Journal of Product Innovation Management*. 2014. Vol. 31. No. 4. P. 814–831.
11. Mokter H. Open innovation: so far and a way forward // *World Journal of Science, Technology and Sustainable Development*. 2013. Vol. 10. No. 1. P. 30–41.
12. Alexy O., George G. Category divergence, straddling, and currency: Open innovation and the legitimization of illegitimate categories // *Journal of Management Studies*. 2013. Vol. 50. No. 2. P. 173–203.
13. Huang H.-C., Lai M.-C., Lin L.-H., Chen C.-T. Overcoming organizational inertia to strengthen business model innovation: An open innovation perspective // *Journal of Organizational Change Management*. 2013. Vol. 26. No. 6. P. 977–1002.
14. Kleemann F., Voss G.G., Rieder K. Un(der)paid innovators: The commercial utilization of consumer work through crowdsourcing // *Science, Technology & Innovation Studies*. 2008. Vol. 4. No. 1. P. 5–26.
15. Becker A., Mladenow A., Kryvinska N., Strauss C. Aggregated survey of sustainable business models for agile mobile service delivery platforms // *Journal of Service Science Research*. 2012. Vol. 4. No. 1. P. 97–121.
16. Strategic management of disruptive technologies: a practical framework in the context of voice services and of computing towards the cloud / R. Kopetzky [et al.] // *International Journal of Grid and Utility Computing*. 2013. Vol. 4. No. 1. P. 47–59.
17. Bursuk I., Mladenow A., Novak N.M., Strauss C. Online cofounder search in tech startups // *Proceedings of the 18th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2016)*. Singapore, 28–30 November 2016. P. 482–488.
18. Andries P., Debackere K. Business model innovation: Propositions on the appropriateness of different learning approaches // *Creativity and Innovation Management*. 2013. Vol. 22. No. 4. P. 337–358.
19. Gibson E., Jetter A. Towards a dynamic process for business model innovation: A review of the state-of-the-art // *Proceedings of the Infrastructure and Service Integration Conference (PICMET 2014)*. Kanazawa, Japan, 27–31 July 2014. P. 1230–1238.
20. Teece D.J. Business models, business strategy and innovation // *Long Range Planning*. 2010. Vol. 43. No. 2–3. P. 172–194.
21. Burr J., Guriksen S. Interactive pinball business // *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design (NordiCHI 2012)*. Copenhagen, Denmark, 14–17 October 2012. P. 129–138.
22. IAP2's Public Participation Spectrum. [Электронный ресурс]: https://www.iap2.org.au/Tenant/C0000004/00000001/files/IAP2_Public_Participation_Spectrum.pdf (дата обращения 01.03.2017).
23. Chew E.K. iSIM: An integrated design method for commercializing service innovation // *Information Systems Frontiers*. 2016. Vol. 18. No. 3. P. 457–478.
24. Ebel P., Bretschneider U., Leimeister J.M. Leveraging virtual business model innovation // *Information Systems Journal*. 2016. Vol. 26. No. 5. P. 519–550.
25. Ogilvie T. How to thrive in the era of collaborative services entrepreneurship // *Research Technology Management*. 2015. Vol. 58. No. 5. P. 24–33.
26. Trimi S., Berbegal-Mirabent J. Business model innovation in entrepreneurship // *International Entrepreneurship and Management Journal*. 2012. Vol. 8. No. 4. P. 449–465.
27. Open business model, process and service innovation with VDML and ServiceML / A.J. Berre [et al.] // *Proceedings of 5th International IFIP Working Conference on Enterprise Interoperability (IWEI 2013)*. Enschede, The Netherlands, 27–28 March 2013. P. 127–142.
28. Zolnowski A., Weiss C., Bohmann T. Representing service business models with the service business model canvas: The case of a mobile payment service in the retail industry // *Proceedings of the 47th Hawaii International Conference on System Science (HICSS 2014)*. Hawaii Big Island, 6–9 January 2014. P. 718–727.
29. Pynnönen M., Hallikas J., Ritala P. Managing customer-driven business model innovation // *International Journal of Innovation Management*. 2012. Vol. 16. No. 4. P. 1–13.
30. Mladenow A., Bauer C., Strauss C. Crowd logistics: the contribution of social crowds in logistics activities // *International Journal of Web Information Systems*. 2016. Vol. 12. No. 3. P. 379–396.
31. Ernst C., Mladenow A., Strauss C. Collaboration and crowdsourcing in emergency management // *International Journal of Pervasive Computing and Communications*. 2017. Vol. 13. No. 2. P. 176–193.
32. Gudiksen S. Business model design games: Rules and procedures to challenge assumptions and elicit surprises // *Creativity and Innovation Management*. 2015. Vol. 24. No. 2. P. 307–322.
33. Denicolai S., Ramirez M., Tidd J. Creating and capturing value from external knowledge: the moderating role of knowledge intensity // *R&D Management*. 2014. Vol. 44. No. 3. P. 248–264.

Open business model innovation: Literature review and agenda for future research

Tiare-Maria Brasseur

Lecturer, Department of e-Business
University of Vienna

Address: 1, Oskar-Morgenstern-Platz, Vienna, 1090, Austria

E-mail: tiare.brasseur@gmail.com

Andreas Mladenow

Lecturer, Department of e-Business
University of Vienna

Address: 1, Oskar-Morgenstern-Platz, Vienna, 1090, Austria

E-mail: andreas.mladenow@univie.ac.at

Christine Strauss

Professor, Department of e-Business
University of Vienna

Address: 1, Oskar-Morgenstern-Platz, Vienna, 1090, Austria

E-mail: christine.strauss@univie.ac.at

Abstract

In today's fast-paced business environment, firms are constantly pressured to innovate in order to remain competitive. Business model innovation (BMI) has recently attracted increasing attention as a promising approach to achieve competitive advantage in the face of fierce competition. Despite its great potential, however, BMI also entails high degrees of complexity, uncertainty and financial risk. Fueled by the rise of digital technologies, BMI has become increasingly open and collaborative in the recent past.

The aim of this paper is to investigate the role and implications of open and collaborative practices in BMI and to provide a comprehensive review of available literature in this field. Therefore, a systematic review of literature at the intersection of Open Innovation (OI) and BMI has been carried out. Our analysis of the literature identified two major research streams in open business model innovation (OBMI): OBMI trends (customer-driven BMI, BM co-creation, early BM validation, virtual collaboration, design thinking) and OBMI effects. Overall, the findings support a growing trend of collaboration and co-creation in BMI supported by digital or tangible tools, and further reveal that OI has a direct positive effect on BMI success. Analysis of the literature also shows that the field of OBMI is still an under-researched area.

Key words: business model innovation, open business model innovation, digital business modeling, open innovation, collaboration, collaborative infrastructure, co-creation, customer innovation.

Citation: Brasseur T.-M., Mladenow A., Strauss C. (2017) Open business model innovation: Literature review and agenda for future research. *Business Informatics*, no. 4 (42), pp. 7–16. DOI: 10.17323/1998-0663.2017.4.7.16.

References

1. Chesbrough H. (2010) Business model innovation: Opportunities and barriers. *Long Range Planning*, vol. 43, no. 2–3, pp. 354–363.
2. Brasseur T.-M., Mladenow A., Strauss C. (2017) Business model innovation to support smart manufacturing. Proceedings of *American Conference on Information Systems 2017. Workshop on Smart Manufacturing. Boston, USA, 10–12 August 2017*. Available at: <http://aisel.aisnet.org/sigbd2017/9> (accessed 16 October 2017).
3. Enkel E., Gassmann O., Chesbrough H. (2009) Open R&D and open innovation: Exploring the phenomenon. *R&D Management*, vol. 39, no. 4, pp. 311–316.
4. Gassmann O., Enkel E., Chesbrough H. (2010) The future of open innovation. *R&D Management*, vol. 40, no. 3, pp. 213–221.
5. Huizingh E.K. (2011) Open innovation: State of the art and future perspectives. *Technovation*, vol. 31, no. 1, pp. 2–9.
6. Mladenow A., Fröschl K.A. (2011) *Kooperative forschung*. Frankfurt am Main: Lang.
7. Laursen K., Salter A. (2006) Open for innovation: the role of openness in explaining innovation performance among U.K. manufacturing firms. *Strategic Management Journal*, vol. 27, no. 2, pp. 131–150.

8. Lopez-Vega H., Tell F., Vanhaverbeke W. (2016) Where and how to search? Search paths in open innovation. *Research Policy*, vol. 45, no. 1, pp. 125–136.
9. Prugl R., Schreier M. (2006) Learning from leading-edge customers at The Sims: Opening up the innovation process using toolkits. *R&D Management*, vol. 36, no. 3, pp. 237–250.
10. West J., Bogers M. (2014) Leveraging external sources of innovation: a review of research on open innovation. *Journal of Product Innovation Management*, vol. 31, no. 4, pp. 814–831.
11. Mokter H. (2013) Open innovation: so far and a way forward. *World Journal of Science, Technology and Sustainable Development*, vol. 10, no. 1, pp. 30–41.
12. Alexy O., George G. (2013) Category divergence, straddling, and currency: Open innovation and the legitimization of illegitimate categories. *Journal of Management Studies*, vol. 50, no. 2, pp. 173–203.
13. Huang H.-C., Lai M.-C., Lin L.-H., Chen C.-T. (2013) Overcoming organizational inertia to strengthen business model innovation: An open innovation perspective. *Journal of Organizational Change Management*, vol. 26, no. 6, pp. 977–1002.
14. Kleemann F., Voss G.G., Rieder K. (2008) Un(der)paid innovators: The commercial utilization of consumer work through crowdsourcing. *Science, Technology & Innovation Studies*, vol. 4, no. 1, pp. 5–26.
15. Becker A., Mladenow A., Kryvinska N., Strauss C. (2012) Aggregated survey of sustainable business models for agile mobile service delivery platforms. *Journal of Service Science Research*, vol. 4, no. 1, pp. 97–121.
16. Kopetzky R., Günther M., Kryvinska N., Mladenow A., Strauss C., Stummer C. (2013) Strategic management of disruptive technologies: a practical framework in the context of voice services and of computing towards the cloud. *International Journal of Grid and Utility Computing*, vol. 4, no. 1, pp. 47–59.
17. Bursuk I., Mladenow A., Novak N.M., Strauss C. (2016) Online cofounder search in tech startups. Proceedings of the *18th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2016)*. Singapore, 28–30 November 2016, pp. 482–488.
18. Andries P., Debackere K. (2013) Business model innovation: Propositions on the appropriateness of different learning approaches. *Creativity and Innovation Management*, vol. 22, no. 4, pp. 337–358.
19. Gibson E., Jetter A. (2014) Towards a dynamic process for business model innovation: A review of the state-of-the-art. Proceedings of the *Infrastructure and Service Integration Conference (PICMET 2014)*. Kanazawa, Japan, 27–31 July 2014, pp. 1230–1238.
20. Teece D.J. (2010) Business models, business strategy and innovation. *Long Range Planning*, vol. 43, no. 2–3, pp. 172–194.
21. Burr J., Guriksen S. (2012) Interactive pinball business. Proceedings of the *7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design (NordCHI 2012)*. Copenhagen, Denmark, 14–17 October 2012, pp. 129–138.
22. IAP2 (2014) IAP2's *Public Participation Spectrum*. Available at: https://www.iap2.org.au/Tenant/C0000004/00000001/files/IAP2_Public_Participation_Spectrum.pdf (accessed 01 March 2017).
23. Chew E.K. (2016) iSIM: An integrated design method for commercializing service innovation. *Information Systems Frontiers*, vol. 18, no. 3, pp. 457–478.
24. Ebel P., Bretschneider U., Leimeister J.M. (2016) Leveraging virtual business model innovation. *Information Systems Journal*, vol. 26, no. 5, pp. 519–550.
25. Ogilvie T. (2015) How to thrive in the era of collaborative services entrepreneurship. *Research Technology Management*, vol. 58, no. 5, pp. 24–33.
26. Trimi S., Berbegal-Mirabent J. (2012) Business model innovation in entrepreneurship. *International Entrepreneurship and Management Journal*, vol. 8, no. 4, pp. 449–465.
27. Berre A.J., Man H.D., Lew Y., Elvesaeter B., Ursin-Holm B.M. (2013) Open business model, process and service innovation with VDML and ServiceML. Proceedings of the *5th International IFIP Working Conference on Enterprise Interoperability (IWEI 2013)*. Enschede, The Netherlands, 27–28 March 2013, pp. 127–142.
28. Zolnowski A., Weiss C., Bohmann T. (2014) Representing service business models with the service business model canvas: The case of a mobile payment service in the retail industry. Proceedings of the *47th Hawaii International Conference on System Science (HICSS 2014)*. Hawaii Big Island, 6–9 January 2014, pp. 718–727.
29. Pynnönen M., Hallikas J., Ritala P. (2012) Managing customer-driven business model innovation. *International Journal of Innovation Management*, vol. 16, no. 4, pp. 1–13.
30. Mladenow A., Bauer C., Strauss C. (2016) Crowd logistics: the contribution of social crowds in logistics activities. *International Journal of Web Information Systems*, vol. 12, no. 3, pp. 379–396.
31. Ernst C., Mladenow A., Strauss C. (2017) Collaboration and crowdsourcing in emergency management. *International Journal of Pervasive Computing and Communications*, vol. 13, no. 2, pp. 176–193.
32. Gudiksen S. (2015) Business model design games: Rules and procedures to challenge assumptions and elicit surprises. *Creativity and Innovation Management*, vol. 24, no. 2, pp. 307–322.
33. Denicolai S., Ramirez M., Tidd J. (2014) Creating and capturing value from external knowledge: the moderating role of knowledge intensity. *R&D Management*, vol. 44, no. 3, pp. 248–264.

Цифровая экономика: концептуальная архитектура экосистемы цифровой отрасли

Ю.М. Акаткин

кандидат экономических наук, заведующий лабораторией социально-демографической статистики
Российский экономический университет им. Г.В. Плеханова
Адрес: 117997, г. Москва, Стремянный пер., д. 36
E-mail: u.akatkin@semanticpro.org

О.Э. Карпов

доктор медицинских наук, профессор
член-корреспондент Российской академии наук
генеральный директор Национального медико-хирургического Центра им. Н.И. Пирогова
Адрес: 105203, г. Москва, ул. Нижняя Первомайская, д. 70
E-mail: nmhc@mail.ru

В.А. Конявский

доктор технических наук, заведующий кафедрой защиты информации
Московский физико-технический институт (МФТИ)
Адрес: 117303, г. Москва, ул. Керченская, д. 1А
E-mail: konyavskiy@gospochta.ru

Е.Д. Ясиновская

старший научный сотрудник лаборатории социально-демографической статистики
Российский экономический университет им. Г.В. Плеханова
Адрес: 117997, г. Москва, Стремянный пер., д. 36
E-mail: elena@semanticpro.org

Аннотация

Основной мотив цифровой трансформации определяется стремлением «клиентов нового цифрового поколения» к своевременности, доступности, качеству и персонализации. Базовым для цифровой парадигмы является принцип «все как услуга», причем услуга, ориентированная на данные и совместное использование информационных ресурсов (в том числе государственных) с учетом требований интероперабельности и безопасности. В статье рассматриваются основные подходы к цифровой трансформации на примере наиболее активных с точки зрения инноваций секторов экономики: банковской отрасли и сферы здравоохранения; сравниваются проприетарный подход к созданию цифровых сервисов (продуктов) и возможность построения экосистемы цифровой отрасли, ориентированной на привлечение неограниченного числа участников. Определена цель создания экосистемы — предоставление населению цифровых сервисов, которые формируются «на лету», «по требованию», в реальном времени, с учетом соблюдения всех норм и регламентов, в условиях максимального доверия. Подчеркивается роль открытости для объединения усилий сообщества разработчиков, заинтересованных в развитии цифровой отрасли, развития государственно-частного партнерства и построения конкурентной среды с целью обеспечения стремительного роста числа доступных цифровых сервисов, а также улучшения их качества. Поскольку цифровая экономика основывается на экономике знаний, особенно важным авторы считают формирование семантического ядра, которое выступает носителем знаний в экосистеме цифровой отрасли. Необходимость реализации семантического ядра подтверждается кратким анализом современных семантических подходов к стандартизации взаимодействия в вышеуказанных отраслях, таких как FIBO, BIAN (банки), HL7 и UMLS (медицина). На основе проведенного исследования авторами разработана концептуальная архитектура экосистемы и сформирован ряд предложений по цифровой трансформации отрасли, в том числе по государственной поддержке инновационного развития и формированию условий для создания новых цифровых продуктов на основе принципов доступности, своевременности, персонализации, технологичности и безопасности.

Ключевые слова: цифровизация, цифровая экономика, цифровая услуга, экономический сектор, экосистема, концептуальная архитектура, совместимость, обмен информацией, семантическая интеграция, семантическое ядро.

Цитирование: Акаткин Ю.М., Карпов О.Э., Коняевский В.А., Ясиновская Е.Д. Цифровая экономика: концептуальная архитектура экосистемы цифровой отрасли // Бизнес-информатика. 2017. № 4 (42). С. 17–28. DOI: 10.17323/1998-0663.2017.4.17.28.

Введение

Взреление и активное включение в трудовую деятельность поколений, родившихся в период с 1980 по 2000 годы, существенно меняют профиль среднего потребителя, который стремится не только получать качественные и своевременные услуги (продукты), но и пользоваться при этом уже привычными мобильными и другими инновационными технологиями. Сейчас, когда речь идет о Четвертой промышленной революции [1], не остается сомнений в необходимости цифровой трансформации и интеграции различных отраслей в цифровой мир и цифровую экономику.

Хотя в отношении целей цифровой трансформации консенсус экспертов в основном достигнут, понятие «цифровая экономика» остается весьма размытым и носит скорее маркетинговый характер, следуя тенденциям рынка в интересах основных игроков. В течение последних нескольких лет информационное пространство наполнили различные обзоры и концепции, выполненные ведущими аналитическими агентствами и лидерами ИКТ-отрасли, посвященные как цифровизации экономики в целом, так и отдельным ее секторам. На основе этих материалов Аналитический центр при Правительстве Российской Федерации собрал и предоставил экспертам семь определений «цифровой экономики» [2]. Однако более полезно было бы перейти от поиска точных формулировок к реализации конкретных шагов, направленных на развитие цифровых отраслей экономики и государственного управления.

Цифровую трансформацию, которая намечается в сфере государственного управления, следует рассматривать в общем комплексе мер по переходу к цифровой экономике. Предложения по реализации перехода к цифровому правительству, сформулированные авторами в статье «Towards the digital government in Russia: Integrative approach» [3] сохраняют свою актуальность и должны быть учтены в этом комплексе мер. Речь идет, в первую очередь, об обеспечении интероперабельности, в том числе семантической, о построении открытых платформ,

обеспечивающих совместную работу специалистов предметной области и ИТ-специалистов, формировании экспертного сообщества с привлечением ведущих ученых.

В данной статье для определения основных путей и способов цифровой трансформации предложена концепция архитектуры экосистемы цифровой отрасли. Рассмотрение участников и основных составляющих экосистемы, особенностей реализации (на примере банковского сектора и здравоохранения) и роли государства в цифровой трансформации экономики позволили выработать предложения для обеспечения государственной поддержки инновационного развития на основе принципов государственно-частного партнерства, а также создания условий для разработки цифровых сервисов и продуктов, ориентированных на потребителя.

1. Текущее состояние цифровизации в некоторых секторах экономики

Используя новейшие технологии, цифровизация перекраивает картину конкуренции и размывает границы, установившиеся среди игроков того или иного сектора экономики. Переход к цифровой парадигме сегодня диктуется цифровыми лидерами — Google, Apple, Facebook и Amazon — и, соответственно, его значимость определяется уровнем ожиданий «клиентов нового цифрового поколения», которые привыкли не только к постоянной доступности услуг, но и к высокой скорости их доставки (получения). Поэтому усилия крупных ИТ-компаний, поставщиков оборудования и программного обеспечения, а также инновационных стартапов направлены в первую очередь на изменение формы, в которой клиентам предлагаются существующие продукты и услуги.

Основываясь на том, что клиенты быстро привыкают к уровню цифрового обслуживания, предлагаются решения, которые лучше удовлетворяют их потребности, повышают доступность, обеспечивают удобство использования и персонализируют услуги, учитывая индивидуальные характеристики потребителя и его интересы.

Таким образом, переход к цифровой экономике определяют:

- ♦ ориентация на потребности клиента (цифровая услуга или цифровой продукт формируются по требованию потребителя, а не по предложению производителя);

- ♦ ориентация на мобильность и скорость (принципы «здесь и сейчас», любое устройство, любой канал связи);

- ♦ ориентация на данные (получение новых данных из существующих, анализ, принятие решений).

Это приводит к трансформации бизнес-моделей, направленной на индивидуализацию предоставляемых товаров и услуг, а также возникновению принципа «все как услуга», причем услуга, ориентированная на данные. Например, уже сегодня многие рассматривают использование «car-sharing» сервисов наравне с покупкой автомобиля, не говоря об уже ставших популярными «историями успеха» цифрового рынка Uber-подобных службах.

Для удовлетворения высоких ожиданий клиентов компаниям придется ускорить цифровую трансформацию своих бизнес-процессов и выйти за рамки простой автоматизации уже существующих процессов. «Они должны переосмыслить каждый бизнес-процесс, включая сокращение числа необходимых шагов, сокращение количества документов, разработку автоматизированных решений, рассмотрение вопросов нормативного регулирования и борьбы с мошенничеством. Операционные модели, навыки, организационные структуры и роли должны быть переработаны в соответствии с новыми процессами. Модели данных должны быть скорректированы и перестроены так, чтобы обеспечить улучшение качества принятия решений, отслеживание реализации и восприятие потребностей клиентов. Цифровизация часто требует, чтобы старая «мудрость» сочеталась с новыми навыками, например, можно обучить специалистов по выкладке товаров программированию алгоритмов ценообразования. Возможно, потребуются новые роли, такие как специалист по анализу и обработке данных или/и дизайнер пользовательских интерфейсов» [4].

Банковская индустрия — один из показательных примеров стремительного цифрового развития. Консалтинговая компания Deloitte провела исследование применения новых технологий в российских банках, в результате которого эксперты

выбрали 11 инноваций, которые разделили их на несколько групп: безопасность («умная» идентификация), аналитика (Big Data, личные финансовые помощники), цифровые технологии (онлайн-кошелек, бесконтактная оплата и т.д.), автоматизация (роботы в отделениях), геймификация (игры и квесты для клиентов) и P2P-кредитование [5].

Нужно заметить, что лидеры банковского сектора уделяют специальное внимание службам развития информационных технологий — возникают цифровые фабрики.

Так, один из крупнейших банков в Канаде, банк Новой Шотландии или Scotiabank, недавно объявил о разработке новой цифровой фабрики, которая будет разрабатывать и внедрять цифровые инновации и решения для своих клиентов с философией «переосмыслить, как банковское дело служит людям»¹.

В Deutsche Bank также есть собственный центр развития цифровых банковских продуктов: Digital Factory во Франкфурте. Около 400 разработчиков программного обеспечения, ИТ-специалисты и финансовые эксперты из 14 стран совместно работают над разработкой цифровых продуктов с использованием самых современных методов [6].

Кроме того, в ноябре 2016 года Deutsche Bank предоставил разработчикам программного обеспечения из Германии и других стран возможность создавать цифровые решения для банковских клиентов, которые выходят далеко за рамки традиционных финансовых услуг. Такой подход позволил создать экосистему инноваций, объединившую три инновационных лаборатории, на базе которых банк сотрудничает со стартапами всего мира, цифровую фабрику и новый центр исследования и разработки. Банк обеспечивает доступ к проприетарной среде разработки через прикладной программный интерфейс (dbAPI), позволяющий программистам проверять свои идеи для реализации цифровых услуг будущего [7].

Логику открытых интерфейсов в части транзакций поддерживает и революционная для банковской сферы Вторая платежная Директива (Revised Payment Directive, PSD2 EU²). Она на законодательном уровне обязывает банки Европейского Союза бесплатно предоставить API для сторонних разработчиков пользовательских приложений. По поручению клиента и без необходимости заключения до-

¹ <http://digitalfactory.scotiabank.com/> (дата обращения: 19.09.2017)

² https://ec.europa.eu/info/law/payment-services-psd-2-directive-eu-2015-2366_en (дата обращения 19.09.2017)

говора с банком третья сторона сможет совершать платежи и отображать информацию о транзакциях в своих приложениях. Это еще один шаг к открытому банкингу, который с одной стороны пугает многих участников рынка, с другой — представляется новой, стратегически важной, перспективой³.

С другой стороны, довольно традиционно применяется и проприетарный подход к разработке цифровых банковских услуг. Так, в 2011 году российский Сбербанк создал ИТ-компанию «Сбертех». В настоящее время около 7000 сотрудников в 16 городах развивают 350 проектов для Сбербанка, крупнейшими из которых являются три: «Единая фронтальная система» — обслуживание клиента с любого места через любой канал, «Платформа поддержки развития бизнеса» — инструмент для создания бизнес-приложений и «Фабрика данных» — для монетизации данных о клиентах и анализа их поведения [8].

Проекты Сбертех ориентированы на модернизацию собственной банковской системы силами собственных разработчиков. Однако развитие открытых интерфейсов (Open API) также декларируется, как одно из будущих направлений, в том числе в связи с созданием так называемых маркетплейсов или агрегаторов, «когда вместе со своими партнерами начинают продавать не только финансовые, но и смежные услуги, такие как путешествия и страховки. Это особенно актуально в связи с тем, что экосистема банков разрастается и есть много желающих писать собственные приложения, связанные с банковскими сервисами» [9]. Это говорит о стремлении к открытости и к привлечению большого числа независимых ИТ-специалистов, полных ориентированных на клиента стартап-идей. Такой подход диктует парадигма цифровой экономики.

Рынок инноваций в медицине является одним из наиболее активных. Цифровая трансформация этой отрасли направлена на выявление и предотвращение большинства (80/20) угроз для жизни и здоровья за счет своевременных предварительной диагностики и мониторинга состояния здоровья, оказания первичных медицинских консультаций и услуг медицинского персонала по месту требования, инициирования услуг срочной медицинской помощи и направления на углубленные медицинские обследования в высокотехнологичные медицинские центры [10].

«The Accenture Digital Health Technology Vision 2016» [11] подчеркивает, что развитие потребительского опыта и методов предоставления медицинской помощи, а также реализация карьерного роста работников здравоохранения имеют такое же жизненно важное значение, как и необходимость идти в ногу с изменяющимися технологиями. Accenture раскрывает пять тенденций, которые подтверждают, что победа в цифровую эпоху зависит от людей:

1. Интеллектуальная автоматизация: необходимо перейти к действиям, отличающимся от старых, для новых рабочих мест, продуктов и услуг в здравоохранении;
2. Ликвидная рабочая сила: сегодняшние требования к цифровым технологиям требуют высокопрочных навыков управления здравоохранением;
3. Экономика платформ: экосистемы — новая основа цифрового здравоохранения;
4. Предсказуемый разрыв: цифровые экосистемы в обозримом будущем размывают границы здравоохранения;
5. Цифровое доверие: с ростом рисков безопасности данных в области здравоохранения возрастают и возможности заслужить доверие потребителей [11].

Один из лидеров цифровой экономики — Великобритания — активно создает инфраструктуру, обеспечивающую развитие инновационных решений, путем привлечения широкого круга разработчиков, заинтересованных в развитии цифровой отрасли.

Примером такого проекта является «Digital Catapult» (цифровая катапульта)⁴ — площадка для технических специалистов, творческих работников, представителей бизнеса и академических кругов, которые в рамках кооперации развивают новые идеи и представляют свои продукты на рынке Великобритании и в других странах. Это позволяет британским «цифровым» предприятиям внедрять инновации с большей скоростью и с меньшим риском, поэтому выход на рынок новых продуктов и услуг может быть ускорен.

Code4Health⁵ — другая инициатива, поддерживаемая NHS England и NHS Digital, которая позволяет использовать цифровые инструменты и технологии для предоставления безопасных, высококачественных, эффективных и «бережных» услуг удаленного ухода (TeleCare).

³ <https://nordeaopenbanking.com/> (дата обращения 19.09.2017)

⁴ <https://digital.catapult.org.uk/> (дата обращения: 19.09.2017)

⁵ <https://code4health.org/> (дата обращения: 19.09.2017).

Code4Health направлен на просвещение и информирование всех участников сообщества по вопросам TeleHealthCare: граждан, пациентов, опекунов, медицинских работников и специалистов в области цифровых технологий о возможностях использования цифровых технологий и оснащении их инструментами, знаниями и навыками для сотрудничества в разработке и внедрении высококачественных цифровых решений [12].

В состав проекта входят:

❖ **Платформа.** Платформа Code4Health предоставляет среду моделирования, в которой можно получить информацию о ресурсах, размещенных в создаваемой открытой цифровой экосистеме медицины. С ее использованием разработчики могут кодировать, создавать приложения, открывать и создавать контент, а также тестировать свои идеи;

❖ **Обучение.** Объяснение клиницистам сути технологий, которые они используют. Code4Health предоставляет возможности совместного обучения, чтобы помочь в разработке новых передовых сервисов для оказания медицинской помощи;

❖ **Сообщества.** Сообщества Code4Health объединяют людей с общими интересами в области цифровой медицины. Сообщества могут включать людей, заинтересованных в определенном аспекте цифровой медицины, либо объединенных географически.

Таким образом, анализ приведенных фактов и тенденций цифровой экономики в целом и в отдельных отраслях позволяет определить, что основной цифровой трансформации становятся усилия в трех взаимосвязанных направлениях:

1. Модернизация бизнес-процессов, технологий и корпоративной культуры для обслуживания клиентов, находящихся в любом месте и в любое удобное для них время в привычном цифровом формате;

2. Внедрение инновационных технологий как в отрасли, так и в экономике в целом, чтобы обеспечить присутствие на любых цифровых площадках, с использованием мобильных платформ и применением устройств Интернета вещей;

3. Создание организационной формы и нормативной базы для привлечения инноваций и эффективного сотрудничества с разработчиками цифровых решений — приложений и устройств в рамках отраслевого регулирования, обеспечивая снижение административных барьеров, стоящих на пути цифровой трансформации.

Таким образом, также необходимо оперативно решать возникающие проблемы с изменением существующего законодательства, как в отрасли, так и цифровой экономике страны в целом, сохранив разумную степень государственного регулирования и обеспечив достаточный уровень господдержки.

Представленный выше краткий обзор показывает, что сегодня цифровые лидеры используют различные подходы к реализации системы разработки и внедрения цифровых сервисов:

1. Создание проприетарной системы разработки и внедрения цифровых сервисов, а также компании-разработчика цифровых сервисов и ведение самостоятельной разработки и продвижения (например, Сбертех). Вместе с очевидными преимуществами, в этом подходе выделяются следующие минусы:

- ❖ разработчик (учредитель) проприетарной системы определяет, что хочет клиент, то есть трудно реализовать требования цифровой экономики по клиент-ориентированности;
- ❖ несмотря на повышение конкуренции в ИТ-сообществе, интеграция с отраслевым сообществом усложняется, и, стало быть, сложно добиться должного качества и своевременности появления цифровых сервисов;
- ❖ значительные инвестиции в разработку конкретных продуктов и сервисов, которые могут устаревать еще до того, как закончится разработка.

2. Создание экосистемы цифровой отрасли на основе консолидации усилий сообщества разработчиков цифровых технологий и приложений, предприятий отрасли (например, в здравоохранении — медицинских учреждений, оказывающих цифровую медицинскую помощь), а также других заинтересованных лицам.

Существующая практика показывает, что цифровая трансформация диктует необходимость открытости и привлечения широкого круга разработчиков к созданию цифровых сервисов. Поэтому наиболее перспективным является второй подход, который ниже мы рассмотрим подробнее.

2. Экосистема цифровой отрасли

Достаточно давно эксперты и консультанты [13, 14] подчеркивают, что цифровая экономика становится одним из ключевых явлений, влияющих на рост ВВП. Она также имеет важные последствия для измерения ВВП, производительности и благо-

состояния домохозяйств в секторе розничной торговли и во всех секторах экономики, связанных с оказанием услуг [15].

Вопросы цифровой экономики стали в этом году одной из основных тем Санкт-Петербургского экономического форума, на котором подчеркивалось, что России нужно наращивать кадровые, интеллектуальные и технологические преимущества в сфере цифровой экономики, и для этого власти будут действовать по направлениям, имеющим системное значение. Под этим подразумевалась и необходимость формирования гибкой нормативной базы для внедрения цифровых технологий во все сферы жизни⁶. При этом все решения должны приниматься с учетом обеспечения информационной безопасности государства, бизнеса и граждан⁷.

Цифровая трансформация поддерживается заинтересованными сторонами:

- государством — с точки зрения отраслевого регулирования, установления норм и правил, обеспечения мер государственной поддержки, а также стимулирования отраслевых лидеров к переходу на цифровую парадигму;

- ведущими предприятиями отраслей экономики — в рамках консолидации усилий экспертного сообщества, совместного использования информационных ресурсов и обеспечения доступа к ним для разработки цифровых сервисов, стандартизации и обеспечения достаточного уровня доверия и безопасности;

- инновационными компаниями — как основным драйвером появления новых цифровых услуг и продуктов, ориентированных на потребителя.

Таким образом, принципы открытости и объединения усилий сообщества для перехода к цифровой экономике принципиально важны, а формирование экосистем цифровых отраслей становится базовым решением цифровой экономики.

Экосистема цифровой отрасли — это среда, обеспечивающая условия для инновационного развития и распространения цифровых сервисов, цифровых продуктов, приложений и устройств в конкретном секторе цифровой экономики.

Цель создания экосистемы — предоставить населению цифровые сервисы, которые формируются «на лету», «по требованию», в реальном времени, с учетом соблюдения всех норм и регламентов, а также в условиях максимального доверия. Такие сервисы

позволят потребителям получать услуги и продукты, не задумываясь о том, как устроена работа отрасли в целом и как работают обеспечивающие ее информационные системы.

Экосистема создает основу для государственно-частного партнерства (ГЧП) при цифровизации различных отраслей экономики, предоставляя возможность присоединения для множества сторонних разработчиков — создателей новых приборов, продуктов, инструментов и новых цифровых услуг. Открытость для участия, государственное стимулирование и конкурентная среда сформируют условия для роста числа доступных цифровых сервисов и улучшения их качества. При этом государство может избежать затрат на массовую разработку и продвижение прикладных сервисов, создавая среду для участия малого и среднего бизнеса, поддержки российских производителей сервисов, приборов и устройств.

Рассматривая варианты реализации экосистемы, следует учесть, что ни одна технология не может одновременно решить все проблемы отрасли. Задачи отраслевой цифровизации слишком велики и сложны для какой-либо одной компании. Поэтому следует говорить не о разрозненных системах и сервисах, а о платформе, обеспечивающей возможность совместной работы разрозненных систем и организаций, как с технической, так и с коммерческой точек зрения.

Любая отрасль (подотрасль) цифровой экономики основывается на экономике знаний. Носителем знаний в экосистеме цифровой отрасли должно являться семантическое ядро, поддержка и развитие которого является крайне важным и наукоемким видом деятельности в цифровой экономике. Например, в банковском секторе, в котором семантическая интероперабельность имеет особое значение [16], в настоящее время внедряется Financial Industry Business Ontology, FIBO⁸. Онтология разработана в конце 2015 года OMG (Object Management Group) совместно с EDM (Enterprise Data Management) и является отраслевой инициативой для определения терминов, определений и синонимов финансовой индустрии, основываясь на использовании принципов Semantic Web, таких как RDF/OWL, а также широко применяемых стандартов моделирования OMG, таких как UML. Основой для создания и развития FIBO являются требования Базельского комитета по банковскому надзору (BCBS), направленные на эффективную агрегацию данных и формирование отчетности о рисках [17].

⁶ <https://sputniknews.com/business/201706021054244434-russia-economy-putin-spief-2017/> (дата обращения 19.09.2017)

⁷ <https://ria.ru/economy/20170602/1495725962.html> (дата обращения 19.09.2017)

⁸ <https://www.edmcouncil.org/financialbusiness> (дата обращения 19.09.2017)

С целью сокращения затрат на интеграцию, повышения гибкости ИТ-решений и оптимизации банковской ИТ-архитектуры создана и развивается эталонная модель, представляемая международной некоммерческой ассоциацией банков и поставщиков ИТ-решений, – Banking Industry Architecture Network e.V (BIAN)⁹. BIAN – это совместная некоммерческая экосистема, состоящая из ведущих мировых банков, поставщиков технологий, консультантов и ученых. Она может стать хорошим примером второго подхода, когда члены сообщества объединяют свои отраслевые знания для создания революционной системы банковских технологий, которая обеспечивает стандартизацию и упрощает основную банковскую архитектуру, которая зачастую может быть устаревшей и запутанной. Основываясь на принципах сервисной архитектуры, всеобъемлющая модель предоставляет банкам ориентированные на будущее решения и способствует сотрудничеству в отрасли. BIAN планирует использовать FIBO, рассматривая ее как ключевую дополняющую инициативу¹⁰.

Текущая 5-я версия BIAN включает 7 бизнес-направлений, 36 бизнес-доменов (предметных областей), около 300 сервисов в различных доменах, более 700 бизнес-сценариев и около 2000 типовых бизнес-операций в этих сервисах. В состав разработчиков входят 27 финансовых организаций (ABN AMRO Group, Credit Suisse, Societe Generale Group, Deutsche Bank, Unicredit Group, ING, Achmea, Rabobank, UBS, Banco Galicia и другие) и 43 производителя программного обеспечения (Temenos, Diasoft, Infosys, Sopra Banking Software, TCS Banks, IBM, SAP, Microsoft и другие).

Унификация информационного взаимодействия в сфере здравоохранения развивается с 70-х годов XX века в рамках стратегической медицинской инициативы, в числе направлений которой необходимо выделить Health Level Seven International (HL7) и Unified Medical Language System (UMLS). HL7 создавали как стандарт обмена данными между информационными системами медицинских учреждений. В настоящее время HL7 – это не только стандарт, но и некоммерческая, аккредитованная ANSI организация, занимающаяся разработкой стандартов, посвященных обеспечению всесторонней структуры и соответствующих стандартов для обмена, ин-

теграции, совместного использования и извлечения электронной медицинской информации, которая поддерживает клиническую практику и управление, доставку и оценку медицинских услуг¹¹.

Unified Medical Language System (UMLS) объединяет и распространяет ключевые термины, стандарты классификации и системы кодирования, а также связанные с ними ресурсы для содействия созданию более эффективных и совместимых биомедицинских информационных систем и услуг, включая электронные медицинские карты. UMLS является примером создания универсального справочника медицинских знаний в самом широком возможном смысле с использованием наиболее эффективных методов, созданных в области компьютерной обработки знаний. Разработка UMLS была начата в 1986 году в национальной Медицинской Библиотеке США. В UMLS используются такие понятия как семантическое представление и обработка медицинских знаний, созданы и постоянно развиваются онтологии для описания всех возможных предметных областей, составляющих универсум медицинских знаний¹².

В UMLS входят три инструмента, которые называют источниками знаний:

- ✦ **Метатезаурус:** термины и коды из многих словарей, включая CPT®, ICD-10-CM, LOINC®, MeSH®, RxNorm и SNOMED CT®;

- ✦ **Семантическая сеть:** расширенные категории (семантические типы) и их отношения (семантические отношения);

- ✦ **Лексикон и лексические инструменты SPECIALIST:** инструменты для обработки естественного языка.

Учитывая международные подходы в построении платформ цифровой экономики в здравоохранении и финансовом секторе, авторы разработали концептуальную архитектуру цифровой отрасли.

3. Концептуальная архитектура экосистемы цифровой отрасли

Концептуально архитектура экосистемы цифровой отрасли может быть представлена набором уровней, взаимодействующих между собой по согласованным правилам (*рисунк 1*). К этим уровням относятся:

⁹ <https://bian.org/> (дата обращения 19.09.2017)

¹⁰ <http://fibo2017.dataversity.net/sessionPop.cfm?confid=116&proposolid=10193> (дата обращения 19.09.2017)

¹¹ <http://www.hl7.org/> (дата обращения 19.09.2017)

¹² <https://www.nlm.nih.gov/research/umls/> (дата обращения 19.09.2017)

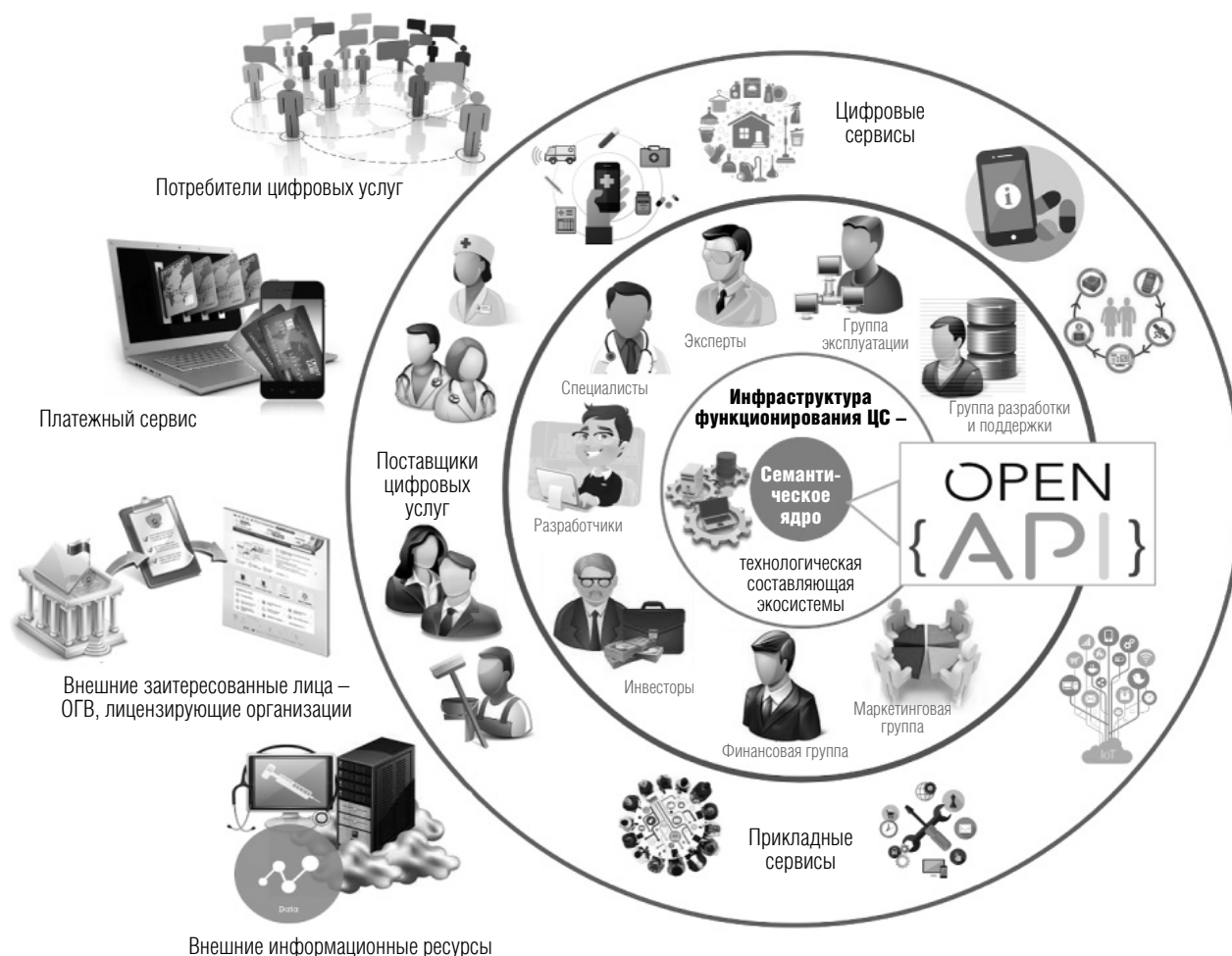


Рис. 1. Концептуальная архитектура экосистемы цифровой отрасли

1. Семантическое ядро;
2. Инфраструктура функционирования цифровых сервисов – технологическая составляющая экосистемы;
3. Пользовательские и прикладные цифровые сервисы, в том числе визуальные интерфейсы и магазин (маркетплейс) приложений, открытые инструменты для разработчиков;
4. Персонализированные приборы и устройства Интернета вещей.

Основной для создания и развития экосистемы цифровой отрасли может стать отраслевой кластер, созданный как самостоятельная единица на основе объединения ведущих предприятий, заинтересованных в развитии цифровой отрасли (например, в здравоохранении – национальных научно-медицинских центров), в рамках государственного регулирования, реализации кластерной политики и других мер государственной поддержки [18].

Ядром кластера должно стать специально организованное предприятие – институт развития, поддержки и распространения экосистемы, в том числе:

1. Для организаций, входящих в кластер – это проектный офис, центр компетенции, центр R&D, осуществляющий:
 - упреждающее проектирование и апробацию новых технологий;
 - проектирование трансформации существующих процедур и бизнес-процессов, встраивание в существующие бизнес-процессы;
 - тестирование разработок на площадке, имитирующей бизнес-процедуры цифровой отрасли для встраивания сервисов;
 - разработку платформы экосистемы, включая OpenAPI и инфраструктурные сервисы;
 - создание системы мониторинга качества цифро-

вых сервисов отрасли на всех этапах жизненного цикла.

2. Для отраслевых предприятий, учреждений и структур — это центр цифровой трансформации и адаптации продуктов и услуг к цифровому рынку, обеспечивающий:

- цифровизацию существующих процессов и внедрение лучших решений в рамках цифровой отрасли;
- легитимность существующих мобильных сервисов, объединение с другими предприятиями, организациями или учреждениями для улучшения качества услуг;
- упрощение процедур защиты информации.

3. Для потребителей — это центр доступа к приложениям, поддерживающий:

- расширенные механизмы доступа, скрывающие от клиентов сложность процедур и не требующие применения дополнительных аппаратных средств, а также упрощающие доступ к сервисам различным категориям потребителей;
- публикацию приложений сторонних разработчиков;
- внедрение и интеграцию устройств Интернета вещей и соответствующих приложений для их использования в жизни.

4. Для компаний-разработчиков — это площадка контроля качества и распространения разработок, центр компетенции и бизнес-акселератор, позволяющий:

- использовать инфраструктурные сервисы и инструменты для разработчиков;
- создавать приложения, которые могут работать с OpenAPI с другими цифровыми системами и сервисами в рамках экосистемы;
- проводить тестирование разработанных приложений, их публикацию и продажу через магазины приложений;
- облегчить поиск инвесторов и создание первичной репутации стартапов;
- участвовать в экспертном сообществе для развития семантического ядра, стандартизации требований к приложениям и сервисам.

Заключение

Таким образом, для успешного перехода к цифровой экономике в ключевых отраслях целесообразно сформировать систему государственного регулирования цифровой отрасли (подотрасли), а также:

♦ обеспечить легитимность оказания цифровых услуг и, при необходимости, регистрацию цифровых устройств (это особенно важно в отраслях с высоким уровнем государственного регулирования, таких как медицина, образование, финансы);

♦ разработать систему мер государственной поддержки цифровой отрасли (в том числе создание пилотных проектов, развитие ГЧП, формирование бизнес-инкубаторов и стимулирование внедрения цифровых сервисов на отраслевых предприятиях);

♦ обеспечить формирование и развитие экосистемы цифровой отрасли;

♦ разработать методологию формирования семантического ядра — носителя полных и непротиворечивых знаний о предметной области (гlossариев, классификаторов, тезаурусов, онтологий, моделей, стандартов, схем взаимодействия, в том числе международных) и создать экспертное сообщество;

♦ обеспечить наполнение семантического ядра на основе гармонизированных российских и международных справочников, классификаторов, тезаурусов и онтологий, для совместного использования информационных ресурсов и стандартизации взаимодействия поставщиков и потребителей цифровых услуг;

♦ разработать инфраструктуру функционирования цифровых сервисов, в том числе с использованием существующих отраслевых информационных ресурсов;

♦ создать пилотные цифровые сервисы и обеспечить их тестирование с использованием инструментальных средств инфраструктуры функционирования, а затем обеспечить их бесшовную интеграцию с существующими системами для последующей цифровой трансформации бизнес-процессов отраслевых предприятий;

♦ обеспечить трансфер инновационных решений в отраслевые предприятия и поддержку отечественных стартап-компаний в этой области.

На основе представленных подходов авторами статьи были подготовлены детальные предложения по цифровой трансформации медицинской отрасли, базовые положения которых вошли в блок «Цифровое здравоохранение» проекта программы «Цифровая экономика», которая разрабатывалась в соответствии с Перечнем поручений по реализации Послания Президента Федеральному Собранию (утв. Президентом РФ 05.12.2016 № Пр-2346) [19].

По мере расширения принятой программы «Цифровая экономика» [20] реализация предложенной отраслевой стратегии цифровой трансформации создаст значительные преимущества за счет быстрого создания новых, ориентированных на клиентов, продуктов и решений, высокой гибкости

и адаптивности к меняющимся потребностям. Эффективность стратегии в значительной мере определяется созданием экосистемы, обеспечивающей благоприятные условия для внедрения инноваций развивающихся компаний, ориентированных на цифровые сервисы. ■

Литература

1. Шваб К. Четвертая промышленная революция / Пер. с англ. М.: Эксмо, 2016.
2. Митин В. Семь определений цифровой экономики / CRN, 2017. [Электронный ресурс]: <https://www.crn.ru/news/detail.php?ID=116780> (дата обращения 19.09.2017).
3. Akatkin Y., Yasinovskaya E., Drozhzhinov V., Konyavskiy V. Towards the digital government in Russia: Integrative approach // International Conference on Electronic Governance and Open Society: Challenges in Eurasia (EGOSE'16). St. Petersburg, Russia, 22–23 November 2016. N.Y.: ACM New York, 2016. P. 41–48.
4. Markovitch S., Willmott P. Accelerating the digitization of business processes / McKinsey, 2014. [Электронный ресурс]: <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/accelerating-the-digitization-of-business-processes> (дата обращения 19.09.2017).
5. Стогней А., Седов Д. Банки для гиков: эксперты назвали лидеров финансовых инноваций / РБК, 2016. [Электронный ресурс]: <http://money.rbc.ru/news/57bb7ffe9a7947340fa28c1f> (дата обращения 19.09.2017).
6. Pertlweiser M. For the banking of the future – Deutsche Bank's digital factory / Deutsche Bank, 2016. [Электронный ресурс]: https://www.db.com/newsroom_news/16_09_29_Digital_Factory_Opening_final_Englisch.pdf (дата обращения 19.09.2017).
7. Deutsche Bank opens its data store to external software developers (October 2016). [Электронный ресурс]: https://www.db.com/newsroom_news/2016/medien/deutsche-bank-opens-its-data-store-to-external-software-developers-en-11735.htm (дата обращения 19.09.2017).
8. Кантышев П. «Сбертех» стал первым по головам // Ведомости. 2016. № 4173 (03.10.2016).
9. Сбербанк рассказал на TAdviser Summit о революционной трансформации ИТ-систем / TAdviser, 2016. [Электронный ресурс]: <http://tadviser.ru/a/318028> (дата обращения 19.09.2017).
10. Карпов О.Э., Акаткин Ю.М., Коняевский В.А., Микерин Д.С. Цифровое здравоохранение в цифровом обществе. М.: Деловой экспресс, 2016.
11. Accenture digital health vision 2016 / Accenture, 2016. [Электронный ресурс]: <https://www.accenture.com/us-en/insight-healthcare-technology-vision-2016> (дата обращения 19.09.2017).
12. Personalised health and care 2020 / GOV.UK, 2017. [Электронный ресурс]: <https://www.gov.uk/government/publications/personalised-health-and-care-2020> (дата обращения 19.09.2017).
13. The new digital economy. How it will transform business / Oxford Economics (June 2011). [Электронный ресурс]: <http://www.pwc.com/mt/en/publications/assets/the-new-digital-economy.pdf> (дата обращения 19.09.2017).
14. 21% GDP productivity growth – Why the digital economy is important! / European Parliamentary Research Service Blog, 2013. [Электронный ресурс]: <https://epthinktank.eu/2013/11/11/21-gdp-productivity-growth-why-the-digital-economy-is-important/> (дата обращения 19.09.2017).
15. Betancourt R.R. Distribution services and the digital economy: Implications for GDP measurement, productivity and household welfare. 2017. [Электронный ресурс]: <http://economics.fiu.edu/events/2017/seminar-roger-betancourt-2/betancourt2017.pdf> (дата обращения 19.09.2017).
16. Рейнгольд Л.А., Волков А.И., Копайгородский А.Н., Пустозеров Е.Ю. Семантическая интероперабельность в решении финансовых задач и способы ее измерения // Прикладная информатика. 2016. Т. 11. № 4 (64). С. 115–134.
17. Zaino J. Banking on FIBO: Financial institutions turn to semantic standard / DATAVERSITY (October 2016). [Электронный ресурс]: <http://www.dataversity.net/banking-fibo-financial-institutions-turn-standard-value-compliance/> (дата обращения 19.09.2017).
18. Скворцов Е.Н., Гуськова Н.Д. Анализ организационных структур систем управления инновационными промышленными кластерами в России / Экономический портал, 2016. [Электронный ресурс]: <http://institutions.com/innovations/2806-analiz-organizatsionnykh-struktur.html> (дата обращения 19.09.2017).
19. Перечень поручений по реализации Послания Президента Федеральному Собранию [Электронный ресурс]: <http://kremlin.ru/acts/assignments/orders/53425> (дата обращения 19.09.2017).
20. Программа «Цифровая экономика Российской Федерации». Утверждена Распоряжением Правительства Российской Федерации от 28 июля 2017 года № 1632-р. [Электронный ресурс]: <http://ac.gov.ru/files/content/14091/1632-r-pdf.pdf> (дата обращения 19.09.2017).

Digital economy: Conceptual architecture of a digital economic sector ecosystem

Yury M. Akatkin

*Head of Laboratory of Social-Demographic Statistics
Plekhanov Russian University of Economics
Address: 36, Stremyanny Lane, Moscow, 117997, Russian Federation
E-mail: u.akatkin@semanticpro.org*

Oleg E. Karpov

*Corresponding Member, Russian Academy of Sciences
General Director of Pirogov National Medical-Surgical Center
Address: 70, Nizhnyaya Pervomaiskaya Street, Moscow, 105203, Russian Federation
E-mail: nmhc@mail.ru*

Valery A. Konyavskiy

*Head of Information Security Department
Moscow Institute of Physics and Technology
Address: 1A, Kerchenskaya Street, Moscow, 117303, Russian Federation
E-mail: konyavskiy@gospochta.ru*

Elena D. Yasinovskaya

*Senior Researcher, Laboratory of Social-Demographic Statistics
Plekhanov Russian University of Economics
Address: 36, Stremyanny Lane, Moscow, 117997, Russian Federation
E-mail: elena@semanticpro.org*

Abstract

The main objective of digital transformation is to fulfill the needs of a “new digital generation customer” for on-demand delivery, quality and personalization. “Anything as a service” has become the key principle of the digital paradigm. This is about a data-oriented service which relies on sharing information resources (including public ones) and the requirements for interoperability, security and trust. This paper presents the main approaches to digital transformation based on the example of the most innovatively active sectors such as banking and healthcare. We compare the proprietary development of digital services (products) to the building of a digital sector ecosystem aimed at attracting an unlimited number of participants. We defined the purpose of creating an ecosystem that is to provide the population with digital services formed on demand, in real time, in compliance with legislation and regulations, as well as in the context of maximum trust. We emphasize the role of openness for uniting the efforts of the community interested in the development of a digital industry, extension of public-private partnerships and building a competitive environment in order to ensure the rapid growth of available digital services, as well as to improve their quality. Since the knowledge economy is the basis for the digital economy, the authors consider it especially important to form a semantic core which acts as the carrier of knowledge in a digital sector ecosystem. We confirmed the necessity to implement the semantic core by a brief analysis of modern semantic approaches to standardization of information sharing in the above-mentioned industries, such as FIBO, BIAN (banking), HL7 and UMLS (health). The research carried out allowed the authors to design the conceptual architecture of the ecosystem and to suggest several proposals for digital transformation of an industry. The proposals express the necessity of state support for innovation and providing the conditions for the entry of new digital products based on the following principles: accessibility, timeliness, personalization, adaptability and security.

Key words: digitalization, digital economy, digital service, economic sector, ecosystem, conceptual architecture, interoperability, information sharing, semantic integration, semantic core.

Citation: Akatkin Y.M., Karpov O.E., Konyavskiy V.A., Yasinovskaya E.D. (2017) Digital economy: Conceptual architecture of a digital economic sector ecosystem. *Business Informatics*, no. 4 (42), pp. 17–28. DOI: 10.17323/1998-0663.2017.4.17.28.

References

1. Schwab K. (2016) *Chetvertaya promyshlennaya revolyutsiya* [The fourth industrial revolution]. Moscow: Eksmo (in Russian).
2. Mitin V. (2017) *Sem' opredeleniy tsifrovoy ekonomiki* [Seven definitions of digital economy]. Available at: <https://www.crn.ru/news/detail.php?ID=116780> (accessed 19 September 2017) (in Russian).
3. Akatkin Y., Yasinovskaya E., Drozhzhinov V., Konyavskiy V. (2016) Towards the digital government in Russia: Integrative approach. Proceedings of *International Conference on Electronic Governance and Open Society: Challenges in Eurasia (EGOSE'16)*. St. Petersburg, Russia, 22–23 November 2016. N.Y.: ACM New York, pp. 41–48.
4. Markovitch S., Willmott P. (2014) *Accelerating the digitization of business processes*. Available at: <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/accelerating-the-digitization-of-business-processes> (accessed 19 September 2017).
5. Stogney A., Sedov D. (2016) *Banki dlya gikov: eksperty nazvali liderov finansovykh innovatsiy* [Banks for geeks: experts listed the leaders of finance innovations]. Available at: <http://money.rbc.ru/news/57bb7ffe9a7947340fa28c1f> (accessed 19 September 2017) (in Russian).
6. Pertlweiser M. (2016). *For the banking of the future – Deutsche Bank's digital factory*. Available at: https://www.db.com/newsroom_news/16_09_29_Digital_Factory_Opening_final_English.pdf (accessed 19 September 2017).
7. Deutsche Bank (2016) *Deutsche Bank opens its data store to external software developers* (October 2016). Available at: https://www.db.com/newsroom_news/2016/medien/deutsche-bank-opens-its-data-store-to-external-software-developers-en-11735.htm (accessed 19 September 2017).
8. Kantyshev P. (2016) «Sbertekh» stal pervym po golovam [Sbertech became a headhunting leader]. *Vedomosti*, no. 4173 (3 October 2016) (in Russian).
9. TAdviser (2016) «Sberbank rasskazal na TAdviser SummIT o revolyutsionnoy transformatsii IT-sistem [Sberbank told on TAdviser SummIT about revolutionary transformation of IT systems]. Available at: <http://tadviser.ru/a/318028> (accessed 19 September 2017) (in Russian).
10. Karpov O., Akatkin Y., Konyavsky V., Mikerin D. (2016) *Tsifrovoe zdavookhraneniye v tsifrovom obshchestve* [Digital healthcare in digital society]. Moscow: Delovoy Express (in Russian).
11. Accenture (2016) *Accenture digital health vision 2016*. Available at: <https://www.accenture.com/us-en/insight-healthcare-technology-vision-2016> (accessed 19 September 2017).
12. GOV.UK (2017) *Personalised health and care 2020*. Available at: <https://www.gov.uk/government/publications/personalised-health-and-care-2020> (accessed 19 September 2017).
13. Oxford Economics (2011) *The new digital economy. How it will transform business*. Available at: <http://www.pwc.com/mt/en/publications/assets/the-new-digital-economy.pdf> (accessed 19 September 2017).
14. European Parliamentary Research Service Blog (2013) *21% GDP productivity growth – Why the digital economy is important!* Available at: <https://epthinktank.eu/2013/11/11/21-gdp-productivity-growth-why-the-digital-economy-is-important/> (accessed 19 September 2017).
15. Betancourt R.R. (2017) *Distribution services and the digital economy: Implications for GDP measurement, productivity and household welfare*. Available at: <http://economics.fiu.edu/events/2017/seminar-roger-betancourt-2/betancourt2017.pdf> (accessed 19 September 2017).
16. Reingold L.A., Volkov A.I., Kopaygorodsky A.N., Pustozarov E.Y. (2016) Semanticheskaya interoperabel'nost' v reshenii finansovykh zadach i sposoby ee izmereniya [Semantic interoperability in solving financial problems and ways of measuring it]. *Applied Informatics*, vol. 11, no. 4 (64), pp. 115–134 (in Russian).
17. Zaino J. (2016) *Banking on FIBO: Financial institutions turn to semantic standard*. DATAVERSITY (October 2016). Available at: <http://www.dataversity.net/banking-fibo-financial-institutions-turn-standard-value-compliance/> (accessed 19 September 2017).
18. Skvortsov E.N., Guskova N.D. (2016) *Analiz organizatsionnykh struktur sistem upravleniya innovatsionnymi promyshlennymi klasterami v Rossii* [Analysis of organizational structures of management systems for innovative industrial clusters in Russia]. Economics Portal (14 June 2016). Available at: <http://institutiones.com/innovations/2806-analiz-organizacionnyx-struktur.html> (accessed 19 September 2017) (in Russian).
19. The President of Russia (2016) *Perechen' porucheniy po realizatsii Poslaniya Prezidenta Federal'nomu Sobraniyu* [Assignment list for realization of the President's Message to the Federal Assembly]. Available at: <http://kremlin.ru/acts/assignments/orders/53425> (accessed 19 September 2017) (in Russian).
20. The Government of the Russian Federation (2017) *Programma «Tsifrovaya ekonomika Rossiyskoy Federatsii»* [The Digital Economy of the Russian Federation Program]. Approved by the Government of the Russian Federation in its resolution No. 1632-r, 28 July 2017. Available at: <http://ac.gov.ru/files/content/14091/1632-r-pdf.pdf> (accessed 19 September 2017) (in Russian).

Совершенствование процесса закупок компании с применением интегрированной рекомендательной системы проведения тендера

З.К. Авдеева

кандидат технических наук
старший научный сотрудник

Институт проблем управления им. В.А. Трапезникова РАН;
доцент кафедры инноваций и бизнеса в сфере информационных технологий
Национальный исследовательский университет «Высшая школа экономики»
Адрес: 117997, г. Москва, ул. Профсоюзная, д. 65
E-mail: avdeeva@hse.ru

А.А. Утробин

бакалавр бизнес-информатики
Национальный исследовательский университет «Высшая школа экономики»
Адрес: 101000, г. Москва, ул. Мясницкая, д. 20
E-mail: alexanderutrobin11@gmail.com

И.Ю. Лыков

студент магистратуры, факультет мировой экономики и мировой политики
Национальный исследовательский университет «Высшая школа экономики»
Адрес: 101000, г. Москва, ул. Мясницкая, д. 20
E-mail: lykov.ilya@gmail.com

Аннотация

В настоящее время процедуры проведения конкурса занимают важное место в работе отдела закупок компании. Большинство современных рекомендательных сервисов функционируют на веб-площадках. Внедрение систем проведения тендеров на предприятии может повысить уровень зрелости процесса закупок и не потребует внесения глобальных изменений в структуру работы процессов. Данная статья посвящена исследованию структуры интегрированной рекомендательной системы проведения тендеров. В Российской Федерации существует несколько федеральных законов, которые определяют порядок процедуры тендера. Также существуют государственные и коммерческие электронные торговые площадки. Рекомендательные сервисы агрегируют конкурсы с разных площадок и предоставляют дополнительные услуги пользователям.

Основная цель исследования — разработать эффективную модель интегрированной рекомендательной системы проведения тендеров. В описательной части работы представлена информация об особенностях процедуры тендера в Российской Федерации и рассмотрены современные рекомендательные сервисы. Выявлены функциональные преимущества интегрированной системы по сравнению с веб-площадкой. Спроектирована структура системы с применением нескольких подходов. С помощью методологии IDEF0 разработана и описана функциональная модель системы, отражающая работу процессов. С помощью спроектированных диаграмм методологии DFD проанализирована работа основной системы и подсистем. Описана математическая модель динамической фильтрации тендеров для создания рекомендаций пользователям. Основываясь на базовых принципах коллаборативной фильтрации, с помощью соответствующих алгоритмов интегрированная система выдает рекомендации пользователям и определяет вероятность успеха в конкретном тендере. Применение данной технологии проведения торгов возможно в компаниях разного масштаба. Разработанная структура интегрированной системы и методы фильтрации для рекомендаций основаны на базовых принципах нового международного направления — e-tendering.

Ключевые слова: рекомендательный сервис, тендер, электронная торговая площадка, интегрированная система, динамическая фильтрация.

Цитирование: Авдеева З.К., Утробин А.А., Лыков И.Ю. Совершенствование процесса закупок компании с применением интегрированной рекомендательной системы проведения тендера // Бизнес-информатика. 2017. № 4 (42). С. 29–39. DOI: 10.17323/1998-0663.2017.4.29.39.

Введение

С развитием информационных технологий большинство стран мира перешли на электронные площадки по проведению торгов. Тендер – это неотъемлемая составляющая государственного и коммерческого секторов экономики. Существует множество электронных торговых площадок (ЭТП) и рекомендательных веб-сервисов, которые собирают тендеры разных категорий и предоставляют информацию посредством запросов. Такие площадки и сервисы также могут предлагать предоставление дополнительных услуг.

Эффективное функционирование отдела закупок предприятия находится в прямой зависимости от автоматизации процесса проведения тендера. Сотрудники регулярно допускают ошибки при составлении заявок, либо при формировании пакета документов в «рутинных» тендерах, либо могут пропустить важные изменения в условиях или сроках. Для того, чтобы избежать ошибок и повысить эффективность работы отдела закупок необходимо внедрить на предприятии интегрированную систему проведения тендера. В результате сотрудники смогут загружать документы в систему и проводить все операции по тендеру, начиная от планирования и заканчивая подписанием договора с заказчиком. Интегрированная система проведения тендера будет совмещать в себе функции классических ЭТП и рекомендательных веб-сервисов, а также предоставит пользователю возможность автоматизации «рутинных» тендеров и выдаст рекомендуемые конкурсы. Система также всегда будет предупреждать об изменениях в условиях заявки, сроков и состава документов.

Объектом исследования являются современные способы проведения тендера, предметом исследования – функционирование интегрированной рекомендательной системы проведения тендера.

Данная работа имеет следующую структуру. В первом разделе рассмотрены особенности проведения конкурсов в Российской Федерации. Во втором разделе проанализирована работа современ-

ных рекомендательных сервисов. В третьем разделе представлена модель интегрированной рекомендательной системы проведения тендера, спроектирована эффективная структура системы, определены функции и подсистемы. В данном разделе также представлена система динамической фильтрации, которая основана на известных математических методах, применяемых в рекомендательных сервисах. Наконец, в четвертом разделе представлены результаты исследования.

1. Особенности электронного конкурса в РФ

Тендер представляет собой способ конкурентной борьбы, происходящей при соблюдении государственных правил и при определенных условиях, зависящих от внутренних правил торговой площадки, в результате которого, осуществляется процесс заключения договора между заказчиком и победителем конкурса. При данном виде торговой операции заказчик делает выбор из множества предложений компаний, в результате чего выбранная компания обязуется выполнить работу, услугу или поставку товара заказчику. Если обратиться к российским правовым актам (<http://zakupki-tendery.ru/zakonodatel-stvo>), то понятие «тендер» официально не употребляется, вместо него используется термин «конкурс».

Современные веб-площадки по проведению торгов представлены в трех направлениях: государственный заказ, муниципальный заказ и коммерческий заказ. Каждому направлению соответствует организатор и свойственные специализированные критерии отбора. В качестве организаторов могут выступать государственные организации, органы местного самоуправления и коммерческие организации. К числу специализированных критериев относятся соотношение цена-качество, информация о заказах, выполненных ранее участником, а также дополнительные условия, предоставляемые участниками в заявке. На трех направлениях представлены различные формы проведения конкурса. Наиболее часто используемые формы – открытый

конкурс, закрытый конкурс, специализированные торги закрытого типа, запрос котировок, закупка у одного поставщика, запрос предложений.

С января 2011 года все государственные заказчики обязаны регистрировать и проводить конкурс на одной из пяти федеральных государственных электронных площадках: Сбербанк АСТ, Росэторг, ОСЭТ, ЭТП ММБВ «Госзакупки». С января 2014 г. вступил в силу Федеральный закон от 05.04.2013 г. № 44-ФЗ «О контрактной системе в сфере закупок товаров, работ, услуг для обеспечения государственных и муниципальных нужд». Теперь, помимо процедуры тендера, на электронных торговых площадках проводится анализ предыдущих торгов, размещается статистика по итогам конкурсов и предоставляется информация о планировании заказов. С помощью Федерального закона от 18.07.2011 г. № 223-ФЗ «О закупках товаров, работ, услуг отдельными видами юридических лиц» регулируются действия юридических лиц — государственных корпораций, монополий, автономных учреждений, хозяйственных обществ, компаний среднего и малого бизнеса. До появления закона действия юридических лиц и организация системы конкурса регулировались с помощью множества правовых актов.

2. Современные рекомендательные сервисы по поиску и планированию тендера

Для того, чтобы найти потенциальных партнеров по бизнесу, компании могут использовать поисковые системы, веб-агрегаторы государственных площадок, отраслевые и универсальные ЭТП, платные и бесплатные тендерные площадки, форумы, ЭТП по продаже имущества должника. С развитием информационных технологий большинство современных ЭТП предоставляет своим пользователям дополнительные функции по планированию и проведению тендера. Однако к настоящему времени только платные системы представляют собой полноценные сервисы по поиску и планированию тендера. Данные сервисы решают ряд проблем, среди которых — огромное количество онлайн-площадок проведения конкурсов, специфический интерфейс систем, слабая отдача от поиска, сложность мониторинга изменений условий тендера, вопросы учета и систематизации документов.

Обратимся к результатам опроса крупных компаний Австралии, Великобритании, Китая, Соеди-

ненных Штатов Америки и Японии, представленных в работе [1]. Существуют три вида архитектуры современных электронных торговых площадок. Самая популярная на сегодняшний день — это архитектура, основанная на принципе. В данном случае в роли заказчика и организатора системы выступает государство. Второй вид — архитектура, основанная на независимой третьей стороне. Яркий пример такой системы — коммерческая электронная торговая площадка. Третий вид — архитектура, основанная на распределенных обязанностях между независимыми третьими сторонами. В таких системах функции выполняются разными юридическими лицами, к примеру, одна компания занимается безопасностью системы, а другая отвечает за электронную цифровую подпись. К настоящему времени рекомендательные системы по поиску и планированию конкурса представлены в последних двух видах архитектуры.

3. Интегрированная рекомендательная система проведения тендера

Совершенствование работы отдела закупок компании может быть осуществлена с помощью интегрированной системы проведения тендеров. Для современной России характерно, что рекомендательные сервисы по прохождению конкурсов предоставляют свои услуги посредством веб-сайта и облачных вычислений. Интегрированная система имеет ряд преимуществ: минимизируется вероятность сбоя системы в неподходящий момент, сохраняется конфиденциальность переписки и контроль доступа к информации о конкурсах, минимизируются финансовые и юридические риски, связанные с третьими лицами, используется эффективная система фильтрации тендеров, применяется индивидуальный подход к каждой зарегистрированной в системе компании.

В целом развитие онлайн-систем проведения конкурсов для крупных компаний направлена на повышение эффективности процесса закупок через открытые тендеры. Для среднего бизнеса — это возможность расширения партнерских связей и поиска новых партнеров по бизнесу. Для малого бизнеса — это удобная система поиска и эффективная система рекомендаций, используя которую, компании могут определять тендеры, которые они могут выиграть с наибольшей вероятностью, ознакомиться со статистикой и построить особую стратегию работы отдела закупок.

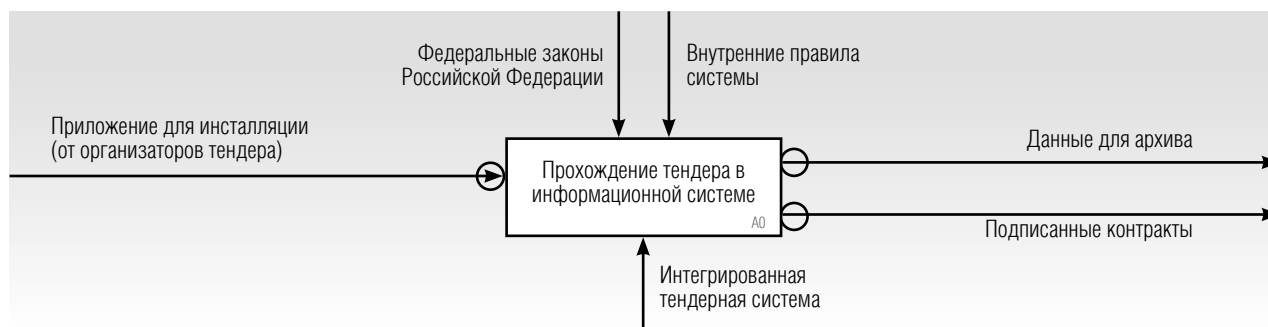


Рис. 1. Контекстная диаграмма процесса прохождения тендера в интегрированной системе

3.1. Процесс прохождения тендера в интегрированной системе проведения тендера

Формализация процесса прохождения тендера строится на основе методологии IDEF0. На *рисунке 1* представлена контекстная диаграмма процесса прохождения тендера в интегрированной системе и соответствующая дополнительная информация о нем. Входной объект – заявки от исполнителей, управление – ФЗ РФ и внутренние правила системы, механизм – интегрированная тендерная система, выходные объекты – данные для архива и подписанные контракты.

Основные функции, которые предоставляет интегрированная система – регистрация в системе, поиск вручную, автоматический поиск, оказание до-

полнительных услуг, учет прохождения конкурса, архивация (*рисунок 2*).

Рассмотрим диаграмму поэтапно. Регистрация в системе подразумевает работу с одобренными заявками от компаний, эта функция выполняется с помощью интегрированной системы и опирается на внутренние правила компании. В результате функция выдает информацию для архива и идентификационный номер пользователя. Зарегистрированные в системе компании могут осуществлять ручной поиск через соответствующие запросы. Автоматический поиск подразумевает рекомендации пользователям, предоставление информации о важных тендерах для компании, а также автоматическое отправление заявок на регулярные конкурсы и настройку электронной подписи в личном ка-

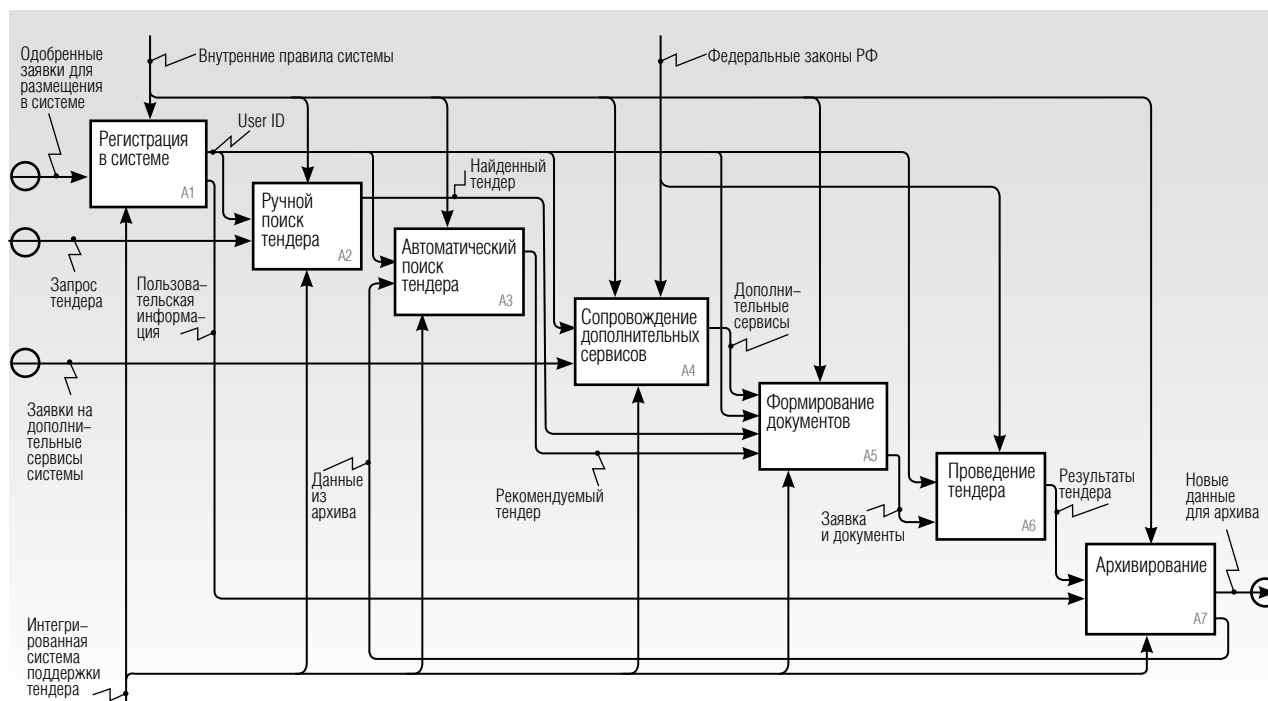


Рис. 2. Диаграмма первого уровня процесса прохождения тендера в интегрированной системе

бинете. Функция оказания дополнительных услуг подразумевает индивидуальный подход к пользователям, при необходимости – помощь при составлении заявки на конкурс, напоминания о сроках и изменений условий, а также подготовку пакета документов. Функция формирования пакета документов использует результаты предыдущих функций – результаты ручного и автоматического поиска и оказанные дополнительные услуги. Функция прохождения тендера подразумевает работу с заявками и пакетами документов от поставщиков. С помощью функции архивации осуществляется анализ данных, который необходим для работы функции автоматического поиска, кроме того, архивируются данные, которые будут использованы в дальнейшей работе.

3.2. Анализ потоков данных в интегрированной системе проведения тендера

Интегрированная система подразумевает работу с исполнителями будущих и текущих конкурсов, а также взаимодействие с руководством системы. Соответствующая контекстная диаграмма, составленная по методологии DFD, представлена на рисунке 3. От зарегистрированных пользователей система получает информацию о компании, после чего на основе последующей совместной работы в личном кабинете формируется статистика, которая позволяет работать рекомендательной системе. От системы компании получают необходимую информацию по тендерам. От руководства система получает внутренние правила, а руководство получает данные для архива.

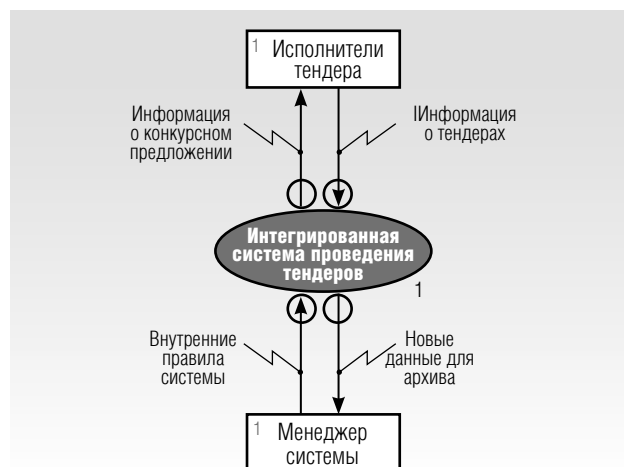


Рис. 3. Контекстная диаграмма интегрированной системы прохождения тендера

Обратимся к диаграмме нулевого уровня, составленной по методологии DFD, на которой изображены все подсистемы интегрированной системы проведения тендера (рисунки 4). Подсистема архива подразумевает хранение информации по сделкам и клиентам, а также ее поиск в хранилище данных. Подсистема поиска предоставляет пользователям возможность осуществлять поиск и сортировку тендеров. При этом поиск может быть осуществлен с помощью ручного метода, также возможен отбор конкурсов по необходимым критериям в рекомендуемых системой тендерах. Подсистема дополнительных услуг предоставляет пользователю возможность выбрать необходимые для него услуги, с помощью которых вероятность выигрыша в тендере возрастет. Подсистема прохождения тендера осуществляет работу по взаимодействию установленной программы с государственными и коммерческими площадками. Посредством эффективно спроектированного интерфейса пользователи могут проходить тендеры прямо в интегрированной системе. Подсистема динамической фильтрации осуществляет подбор рекомендательных тендеров для каждой компании в индивидуальном порядке, основываясь на математических методах и статистике, сформированной в системе архива. Подсистема рекомендации предоставляет пользователю результаты работы подсистемы динамической фильтрации. Если пользователь разрешает проводить автоматическое составление и отправку заявки на регулярные тендеры, то подсистема рекомендаций выполняет эту работу с помощью заложенного алгоритма и электронной цифровой подписи.

3.3. Подсистема динамической фильтрации

Зрелость процесса закупок компании определяется возможностью автоматизации части процессов, что позволяет экономить ресурсы предприятия [2]. Электронный формат заключения сделок позволяет избежать транзакционных издержек, возникающих при поиске подходящего исполнителя. Кроме того, формат тендера открывает ряд возможностей для компаний-подрядчиков в части поиска заказчика. Однако процесс поиска заказчика может оказаться достаточно длительным и требовать сбора большого количества информации. На этом этапе возрастает вероятность для обеих сторон сделать неправильный выбор. В данной части работы предложена модель, позволяющая производить автоматический подбор тендера для исполнителя. В основе модели лежит использование базовых принципов рекомен-

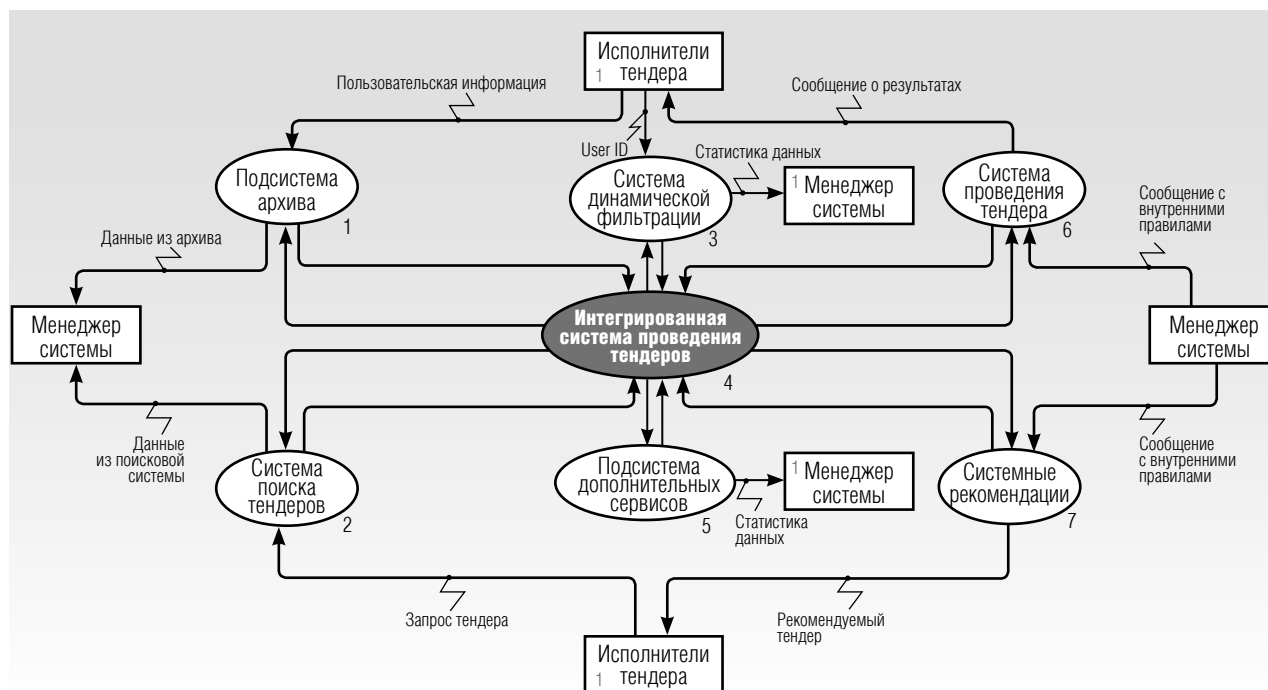


Рис. 4. Диаграмма нулевого уровня интегрированной системы прохождения тендера

дательного сервиса. Ключевое отличие — в том, что метод применяется для подбора тендера, в котором компания-подрядчик может принять участие. В основе лежит принцип коллаборативной фильтрации: для рекомендаций используется история оценок как самого пользователя, так и других пользователей.

Для начала работы рекомендательной системы, пользователям необходимо оценить тендеры, в которых они приняли участие через данную программу. Это условие является обязательным, поскольку на основании этих данных будут строиться дальнейшие рекомендации. Для оценки различных тендеров пользователям предложен определенный набор характеристик. Оценки за каждый параметр располагаются в диапазоне от 1 до 5.

К числу рассматриваемых характеристик относятся:

- ◆ стабильность состояния документов;
- ◆ стабильность состояния условий заявки;
- ◆ стабильность сроков;
- ◆ уровень обратной связи с заказчиком;
- ◆ степень прозрачности сделки;
- ◆ корректность условий заявки.

На основе оценок каждой из этих характеристик формируется общая оценка тендера, которая вычисляется по формуле среднего взвешенного:

$$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{n},$$

где $\bar{x}$ – оценка, присвоенная пользователем определенному тендеру;

x_i – балл, выставленный за i -ю характеристику,
 $i = 1, \dots, n$;

w_i – вес i -й характеристики, $i = 1, \dots, n$;

n – количество характеристик.

Цель алгоритма – сгенерировать рекомендацию для каждого подрядчика. Рекомендация представляет собой предложение тендера, которое может быть интересно данному исполнителю. Система делает рекомендацию, предсказывая оценку, которую пользователь поставил бы определенному тендеру на основании его предыдущих оценок. Важным является то, что при регистрации в системе пользователь указывает сферу своей деятельности для подбора соответствующего предложения.

Простейшая метрика, позволяющая оценить близость двух элементов – евклидово расстояние между векторами. Очевидно, что чем меньше эта величина, тем ближе между собой находятся элементы и, следовательно, имеют больше сходств. На основании этого могут быть построены некоторые рекомендательные подборки. Однако данный метод является не единственным. Более качественный подбор воз-

можен при использовании более сложных расчетов. Рассмотрим несколько возможных алгоритмов предсказания оценки.

Прежде всего, рассмотрим более продвинутые способы фильтрации. Имеется множество исполнителей ($u \in U$) и множество тендеров, подходящих данному подрядчику ($t \in T$). На основании обработанной информации, система дает рекомендацию e_{ut} .

Первый алгоритм связан с кластеризацией пользователей. Сначала необходимо ввести функцию схожести $\text{sim}(u, v)$, где u – выбранный пользователь системы, v – сопоставляемый с ним пользователь. Эта функция показывает, насколько у пользователей схожа история выставления оценок.

Данную функцию можно задать несколькими способами. Например, ее можно охарактеризовать как долю тендеров, в которых приняли участие обе компании, или долю тендеров, для которых у компаний совпали оценки. Далее пользователи разбиваются на кластеры, по принципу «похожие с похожими»: $u \rightarrow G(u)$.

Предсказываемая оценка пользователя для определенного тендера вычисляется как средняя оценка всего кластера, к которому относится данный пользователь:

$$\widehat{e}_{ut} = \frac{1}{|G(u)|} \sum_{v \in G(u)} e_{vt},$$

где $\widehat{e}_{ut}$ – предсказанная оценка тендера t пользователем u , $t \in T$, $u \in G(u)$;

$|G(u)|$ – число пользователей в кластере, к которому относится пользователь u ;

e_{vt} – оценка тендера t пользователем v (другим пользователем кластера), $t \in T$, $v \in G(u)$;

T – множество тендеров;

$G(u)$ – множество пользователей кластера.

Рассмотрим еще один алгоритм, основанный на принципе «user-based», который можно использовать для создания рекомендаций [4]. Приведем формулу, по которой делается предсказание оценки. Здесь также потребуются функция схожести, о которой упоминалось ранее:

$$\widehat{e}_{ut} = \bar{e}_u + \frac{\sum_{v \in U} \text{sim}(u, v) (e_{vt} - \bar{e}_v)}{\sum_{v \in U} \text{sim}(u, v)},$$

где $\widehat{e}_{ut}$ – предсказанная оценка тендера t пользователем u , $t \in T$, $u \in G(u)$;

$\bar{e}_u$ – средняя типичная оценка, выставляемая пользователем u , $u \in G(u)$;

$\bar{e}_v$ – средняя оценка схожего пользователя v , $v \in G(u)$;

e_{vt} – фактическая оценка тендера t схожим пользователем v , $t \in T$, $v \in G(u)$;

$\text{sim}(u, v)$ – функция сходства оценок пользователей u и v , $u \in G(u)$, $v \in G(u)$;

T – множество тендеров;

$G(u)$ – множество пользователей кластера.

Еще один способ оценки основан на идее сингулярного разложения [4]:

$$\widehat{e}_{ut}(\beta) = p_u^T q_t, \\ \beta = \{p_u; q_t \mid u \in U, t \in T\},$$

где $\widehat{e}_{ut}(\beta)$ – предсказанная оценка тендера t пользователем u , $t \in T$, $u \in G(u)$;

p_u – возможная оценка пользователя u , $u \in U$;

q_t – возможная оценки тендера t , $t \in T$;

β – пара оценок (возможная оценка пользователя p_u и возможная оценка тендера q_t);

T – множество тендеров;

U – множество пользователей.

В данном случае p_u и q_t являются параметрами, которые необходимо оценить. Необходимо оптимизировать качество будущих предсказаний:

$$(\widehat{e}_{ut}(\beta) - e_{ut})^2 \rightarrow \min, \quad (1)$$

где $\widehat{e}_{ut}(\beta)$ – предсказанная оценка тендера t пользователем u , $t \in T$, $u \in G(u)$;

e_{ut} – фактическая оценка тендера t пользователем u , $t \in T$, $u \in G(u)$;

β – множество возможных оценок пользователя u и тендера t ;

T – множество тендеров;

U – множество пользователей.

Система располагает некоторым количеством прошлых оценок, следовательно, можно говорить о наличии обучающей выборки. Перепишем (1), проведя регуляризацию:

$$\sum_{u,t} (\widehat{e}_{ut}(\beta) - e_{ut})^2 + \lambda \sum_{\varphi \in \beta} \varphi^2 \rightarrow \min, \quad (2)$$

где $\widehat{e}_{ut}(\beta)$ – предсказанная оценка тендера t пользователем u , $t \in T$, $u \in G(u)$;

e_{ut} – фактическая оценка тендера t пользователем u и тендера t ;

λ – параметр регуляризации, $\lambda > 0$;

φ – стабилизирующий функционал;

β – множество возможных оценок пользователя u и тендера t ;

T – множество тендеров;

U – множество пользователей.

Свойством параметра регуляризации λ является то, что при $\lambda \rightarrow 0$ и наличии ошибок в исходных данных решение корректной задачи стремится к истинному решению соответствующей некорректной задачи. Назначение стабилизирующего функционала φ – обеспечить устойчивость нахождения решения задачи при неточных исходных данных.

Перепишем (2) для целей поставленной задачи:

$$\sum_{u,t} (p_u^T q_t - e_{ut})^2 + \lambda \left(\sum_u \|p_u\|^2 + \sum_t \|q_t\|^2 \right) \rightarrow \min,$$

где β – множество возможных оценок пользователя u и тендера t ;

p_u – возможная оценка пользователя u , $u \in U$;

$\|p_u\|$ – норма оценок пользователя u , $u \in U$;

q_t – возможная оценка тендера t , $t \in T$;

$\|q_t\|$ – норма оценок тендера t , $t \in T$;

e_{ut} – фактическая оценка тендера t пользователем u , $t \in T$, $u \in G(u)$;

λ – параметр регуляризации, $\lambda > 0$;

T – множество тендеров;

U – множество пользователей.

Таким образом, задача свелась к нахождению минимума данного функционала. Это позволяет найти наиболее корректные оценки для предсказаний.

Отдельным преимуществом разработанной системы является не только автоматический подбор тендера для подрядчика, но и оценка вероятности победы в нем. Для вычисления вероятности система анализирует несколько критериев, выставляя им оценку от 1 до 10. К числу критериев относятся:

- ♦ степень загруженности исполнителя тендера;
- ♦ опыт работы пользователя в системе;
- ♦ соответствие тематической области тендера;
- ♦ себестоимость производства;
- ♦ географическая удаленность (если данный показатель не играет роли для конкретной ситуации, то система использует первые четыре критерия для оценки вероятности выигрыша).

Вероятность вычисляется по формуле:

$$P(T) = \frac{\sum_{i=1}^n A_i}{\sum_{i=1}^n B_i} \cdot \frac{1}{n}, \quad (3)$$

где $P(T)$ – вероятность победы в тендере;

A_i – балл который присваивается системой конкретному тендеру, исходя из представленной выше оценки;

B_i – максимальный балл;

$\sum_{i=1}^n A_i$ – сумма баллов, присвоенных конкретному тендеру, исходя из критериев;

$\sum_{i=1}^n B_i$ – максимально возможная сумма баллов;

n – количество участников тендера.

Помимо расчета вероятности в процентах, предусматривается разбиение результата вероятности выигрыша по категориям (таблица 1).

Таблица 1.

Категории вероятности выигрыша в тендере

Категории	Название категории	Вероятность выигрыша (%)
Категория 1	Наиболее высокая вероятность выигрыша	90 – 100
Категория 2	Высокая вероятность выигрыша	70 – 89
Категория 3	Средняя вероятность выигрыша	50 – 69
Категория 4	Низкая вероятность выигрыша	30 – 49
Категория 5	Наименее низкая вероятность выигрыша	0 – 29

Таким образом, в данной части работы рассмотрены следующие методы фильтрации: оценка важных параметров тендера и расчет рекомендации через среднее взвешенное; простейшая метрика через евклидово расстояние; алгоритм, связанный с кластеризацией пользователей; алгоритм, основанный на принципе «user-based»; алгоритм, основанный на сингулярном разложении матрицы. Вероятность выигрыша в конкурсе оценивается с помощью критериев и расчета оценок по формуле (3).

4. Результаты исследования

В ходе исследования разработана эффективная интегрированная рекомендательная система по проведению тендера. За счет агрегирования и отбора данных с разных ЭТП из совокупности альтернативных тендеров исключаются так называемые «ложные» тендеры (лоты с единственным поставщиком, лоты со специальными ограничениями и т.п.). С помощью автоматизированных функций и рекомендаций уменьшается уровень транзакционных издержек со стороны исполнителей тендера. Качественная интегрированная система подразумевает возможность хранения и обработки огромного массива данных.

Данная система может быть использована компаниями разного масштаба производства и разного рода деятельности. С помощью автоматизации

процессов в отделе закупок, связанных с проведением тендера, предприятия могут экономить свои денежные средства. Эффективность данной системы зависит от проработанной структуры подсистем, от окружения основных функций системы, от базовых алгоритмов, заложенных в систему фильтрации. Разработка структуры интегрированной системы опиралась на базовые принципы современного направления e-tendering. Основные черты данного направления определены в книге [5], к ним относятся минимизация юридических и финансовых рисков, минимизация ошибок при составлении заявок, эффективная система поиска тендеров.

В ближайшем будущем системы проведения тендеров смогут выйти на более высокий уровень качества. Многие зарубежные авторы рассуждают о возможном скором изменении структуры тендера (например, в работе [6]), а также исследуют вопрос об изменении конструкции проведения тендера. С развитием технологий рекомендательных сервисов появятся интегрированные системы, которые смогут автоматизировать часть процессов в отделе закупок, в результате чего сократятся расходы на обеспечение и проведения закупок через тендер, увеличится уровень зрелости процесса закупок, увеличится поток сделок, расширится база партнеров и автоматизируются «рутинные» тендеры.

4.1. Пример работы системы фильтрации

Приведем некоторые расчеты, иллюстрирующие работу системы фильтрации. Для примера выберем алгоритм, основанный на методе «user-based». Предположим, что имеется два пользователя *A* и *B*. Есть информация о том, как они оценили тендеры, в которых они приняли участие (таблицы 2 и 3). Необходимо предсказать, какую оценку поставит пользователь *A* тендеру 6.

Таблица 2.

Оценки пользователя А

Тендер	Оценка
Тендер 1	5
Тендер 2	3
Тендер 3	4
Тендер 4	2
Тендер 5	5

Таблица 3.

Оценки пользователя В

Тендер	Оценка
Тендер 1	5
Тендер 2	3
Тендер 4	4
Тендер 5	5
Тендер 6	1
Тендер 8	2

Введем функцию сходства как долю тендеров, для которых у исполнителей совпали оценки. Оба подрядчика приняли участие в тендерах 1, 2, 4 и 5. Однако, оценки совпали лишь для тендеров 1, 2, и 5. Таким образом доля тендеров, для которых совпали оценки для пользователя *A*, составляет 3/5.

Далее, применяя формулу предсказания оценки, вычислим предсказание оценки тендера 6 пользователем *A*:

$$\widehat{e}_{ut} = \bar{e}_u + \frac{\sum_{v \in U} \text{sim}(u, v)(e_v - \bar{e}_v)}{\sum_{v \in U} \text{sim}(u, v)},$$

где $\widehat{e}_{ut}$ – предсказанная оценка тендера *t* пользователем *u*, $t \in T$, $u \in G(u)$;

$\bar{e}_u$ – средняя типичная оценка, выставляемая пользователем *u*, $u \in G(u)$;

$\bar{e}_v$ – средняя оценка схожего пользователя *v*, $v \in G(u)$;

e_v – фактическая оценка тендера *t* схожим пользователем *v*, $t \in T$, $v \in G(u)$;

$\text{sim}(u, v)$ – функция сходства оценок пользователей *u* и *v*, $u \in G(u)$, $v \in G(u)$;

T – множество тендеров;

$G(u)$ – множество пользователей кластера.

Среднюю типичную оценку пользователя ($\bar{e}_u$) возьмем как среднее арифметическое ряда оценок, выставленных системой. Если мы опасаемся зависеть оценку, то можно брать среднее геометрическое, поскольку оно всегда меньше среднего арифметического, кроме случаев, когда все оценки равны. В данном случае это 3,8, для случая использования среднего геометрического это 3,59. Для пользователя В средняя оценка ($\bar{e}_v$) составляет 3,3. Проведя вычисления, мы получаем 3,83. Это означает, что пользователь *A*, вероятнее всего, поставит тендеру 6 оценку 4.

4.2. Пример расчета

Для оценки вероятности победы в тендере система использует предложенную нами методологию, присваивая каждому из параметров значение от 1 до 10. Здесь стоит обратить внимание на то, что разные параметры по-разному учитываются системой. Еще раз обратимся к характеристикам, используемым при подсчете вероятности и приведем некоторые комментарии по методологии присвоения оценок:

♦ степень загруженности исполнителя тендера (в данном случае связь обратная: чем больше загружен данный исполнитель, тем меньший балл за данную характеристику он получает);

♦ опыт работы пользователя в системе (здесь связь прямая: чем более опытен пользователь, тем выше оценку ему присваивается);

♦ соответствие сферы деятельности предметной области закупки (прямая связь);

♦ себестоимость производства (в данном случае связь обратная: чем ниже себестоимость производства у данного производителя, тем большую оценку он получает. Чем ниже себестоимость, тем более низкую цену может запросить данный исполнитель и, следовательно, он имеет больше шансов получить заказ, что должно найти отражение в формуле);

♦ географическая удаленность (если данный показатель не играет роли для конкретной ситуации, то система использует первые четыре критерия для оценки вероятности выигрыша. В данном случае связь тоже прямая: чем ближе находится исполнитель к заказчику, тем больше вероятность выигрыша).

Предположим, что в тендере участвуют две фирмы – *А* и *В*. Ниже приведена *таблица 4* с оценками, выставленными системой.

Таблица 4.

Оценки, выставленные системой

Параметр	Фирма А	Фирма В
Степень загруженности	8	2
Опыт работы	7	4
Соответствие	10	7
Себестоимость	8	8
Географическая удаленность	10	5

Используя формулу (3) для нахождения вероятности выигрыша, получим, что для фирмы *А* вероятность выигрыша составляет 43 %, а для фирмы *В* – 26 %. Очевидно, что у фирмы *А* шансов на победу больше, причем с ростом числа участников вероятность победы будет снижаться.

Заключение

Таким образом, в данной статье рассмотрена эффективная интегрированной система проведения тендера. Проанализировано текущее состояние федеральных законов РФ, выделены особенности прохождения конкурсов. Сформулированы проблемы, которые решает данная система для компаний различного масштаба производства. Разработана структура рекомендательного сервиса с помощью методологий IDEF0 и DFD. Рассмотрены известные математические методы и алгоритмы, которые могут быть применимы в рамках подсистемы динамической фильтрации тендеров. ■

Литература

- 1 Du R., Foo E., Nieto J.G., Boyd C. Designing secure e-tendering systems // Trust, privacy, and security in digital business / S. Katsikas, J. Lopez, G. Pernul (eds). Lecture Notes in Computer Science, vol. 3592. Springer, Berlin, Heidelberg: TrustBus, 2005. P. 70–79.
- 2 Модель зрелости закупок. Анализ функции закупок в российских компаниях / KPMG, 2011. [Электронный ресурс]: <https://home.kpmg.com/ru/ru/home/insights/2011/06/model-of-procurement-maturity.html> (дата обращения 01.09.2017).
- 3 Sarwar B.M., Karypis G., Konstan J.A., Reidl J. Item-based collaborative filtering recommendation algorithms // Proceedings of the 10th International World Wide Web Conference. Hong Kong, 1–5 May 2001. P. 285–295.
- 4 Джонс Т.М. Рекомендательные системы. Часть 1. Введение в подходы и алгоритмы. [Электронный ресурс]: <https://www.ibm.com/developerworks/ru/library/os-recommender1/> (дата обращения: 15.03.2017).
- 5 E-tendering – Security and legal issues: Research report / E.P. Dawson [et al.]. Brisbane: CRC Construction Innovation, 2005.
- 6 Lavelle D., Bardon A. E-tendering in construction: time for a change? // Northumbria Working Paper Series: Interdisciplinary Studies in the Built and Virtual Environment. 2009. No. 2 (2). P. 104–112.

Improvement of the procurement process using the integrated tender recommendation system

Zinaida K. Avdeeva

Senior Researcher, Trapeznikov Institute of Control Sciences;
Associate Professor, Department of Innovation and Business in Information Technologies
National Research University Higher School of Economics
Address: 65, Profsoyuznaya Street, Moscow, 117997, Russian Federation
E-mail: avdeeva@hse.ru

Alexander A. Utrobin

Bachelor in Business Informatics
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: alexanderutrobin11@gmail.com

Ilya Y. Lykov

MSc Program Student, Faculty of World Economy and International Affairs
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: lykov.ilya@gmail.com

Abstract

Currently, tender procedures occupy an important place in the work of the procurement department of any company. Most of the modern recommendation services operate on web-platforms. Implementation of a tendering system in a company can increase the level of maturity of the procurement process and will not require serious changes in the structure of the processes. This article is devoted to the study of the structure of an integrated tender recommendation system. The integrated tender recommendation system is based on procedures prescribed by federal laws, aggregated tenders from different e-trade sites (presenting state and commercial platforms); it offers to its users additional services.

The main purpose of the study is to develop an effective model of an integrated tender recommendation system. In the description part of this article, we present information on peculiarities of the tender procedure in the Russian Federation and modern advisory services are considered. The functional advantages of an integrated system in comparison with the web platform are set out. The structure of the system is designed using several approaches. Using the IDEF0 methodology, a functional model of the system that reflects the work of processes has been developed and described. The operation of the main system and subsystems has been analyzed using the projected diagrams of the DFD methodology. A mathematical model of dynamic filtering of tenders for creating recommendations to users is described. Relying on the basic principles of collaborative filtering and, with the help of appropriate algorithms, the integrated system gives recommendations to users and determines the probability of success in a particular tender. Application of such technology of tenders is possible in companies of different scale. The developed structure of the integrated system and filtering methods for recommendations are based on the basic principles of a new international trend – e-tendering.

Key words: recommendation services, tender, electronic trading platform, integrated system, dynamic filtration.

Citation: Avdeeva Z.K., Utrobin A.A., Lykov I.Y. (2017) Improvement of the procurement process using the integrated tender recommendation system. *Business Informatics*, no. 4 (42), pp. 29–39. DOI: 10.17323/1998-0663.2017.4.29.39.

References

1. Du R., Foo E., Nieto J.G., Boyd C. (2005) Designing secure e-tendering systems. *Trust, privacy and security in digital business* / S. Katsikas, J. Lopez, G. Pernul (eds). Lecture Notes in Computer Science, vol. 3592. Springer, Berlin, Heidelberg, pp. 70–79.
2. KPMG (2011) *Model' zrelosti zakupok. Analiz funktsii zakupok v rossiyskikh kompaniyakh* [Procurement maturity model. Analysis of the procurement function in Russian companies]. Available at: <https://home.kpmg.com/ru/ru/home/insights/2011/06/model-of-procurement-maturity.html> (accessed 01 September 2017) (in Russian).
3. Sarwar B.M., Karypis G., Konstan J.A., Reidl J. (2001) Item-based collaborative filtering recommendation algorithms. *Proceedings of the 10th International World Wide Web Conference. Hong Kong, 1–5 May 2001*. P. 285–295.
4. Jones T.M. (2014) *Rekomendatel'nye sistemy. Chast' 1. Vvedenie v podkhody i algoritmy* [Recommendation systems. Part 1. Introduction to approaches and algorithms]. Available at: <https://www.ibm.com/developerworks/ru/library/os-recommender1/> (accessed 15 March 2017) (in Russian).
5. Dawson E.P., Duncan W.D., Christensen S.A., Black P.J., Foo E., Du R., Boyd C.A., Nieto J. (2005) *E-tendering – Security and legal issues: Research report*. Brisbane: CRC Construction Innovation.
6. Lavelle D., Bardon A. (2009) E-tendering in construction: time for a change? *Northumbria Working Paper Series: Interdisciplinary Studies in the Built and Virtual Environment*, no. 2 (2), pp. 104–112.

Специфика интеграции технологий бизнес-интеллекта и больших данных в процессы экономического анализа

С. Митрович

доктор экономических наук, Университет г. Нови-Сад, Республика Сербия;

докторант экономического факультета

Московский государственный университет им. М.В. Ломоносова

Адрес: 119991, г. Москва, Ленинские горы, д. 1

E-mail: Mitrovic.Stanislav@hotmail.com

Аннотация

Объем данных, используемых в процессах экономического анализа деятельности организаций, ежегодно возрастает. Несмотря на то, что вся необходимая для экономического анализа информация, по сути, присутствует в различных источниках, такие данные фактически часто превращаются в бесполезный, с точки зрения экономического потенциала, источник анализа.

Цель данного исследования заключается в определении основ для интеграции технологий бизнес-интеллекта (Business Intelligence, BI) и больших данных (Big Data) в процессы экономического анализа. Теоретическую и методологическую базу исследования составляют научные исследования, методологические и практические разработки отечественных и зарубежных авторов по вопросам применения ИТ-решений в экономическом анализе.

Как показывают результаты исследования, современные информационные технологии, в частности, системы бизнес-интеллекта и большие данные, значительно изменили возможности совершенствования экономического анализа и сокращения времени принятия решений. С методических позиций, многие аспекты интеграции решений в области BI и Big Data и их внедрения в процессы экономического анализа российских организаций остаются недостаточно исследованными. В отличие от российского, зарубежный рынок применения современных информационных технологий для аналитической обработки экономической информации имеет более давнюю историю и развивается более быстрыми темпами.

Основные выводы исследования указывают на то, что современным организациям, функционирующим в условиях высококонкурентного рынка, следует исходить из понимания, что накопление больших данных не всегда приводит к получению ожидаемой бизнес-выгоды. В данном контексте также можно сделать вывод, современной компании не следует ставить перед собой задачу обработки всех имеющихся данных с целью повышения качества результатов экономического анализа. Более значимым является использование всего объема данных для их сегментации, что позволяет эффективно строить большое количество моделей для небольших кластеров, решая конкретные задачи экономического анализа на базе применения современных информационных технологий.

Ключевые слова: экономический анализ, современные информационные технологии, бизнес-интеллект, большие данные, озера данных, бизнес-аналитика, анализ больших данных.

Цитирование: Митрович С. Специфика интеграции технологий бизнес-интеллекта и больших данных в процессы экономического анализа // Бизнес-информатика. 2017. № 4 (42). С. 40–46.
DOI: 10.17323/1998-0663.2017.4.40.46.

Введение

В отличие от российского, зарубежный рынок современных информационных технологий, в частности систем бизнес-интеллекта (Business Intelligence, BI) и больших данных (Big Data), применяемых для аналитической обработки экономической информации, имеет более давнюю историю и развивается более быстрыми темпами. Согласно прогнозу компании IDC (март 2017 года), к концу 2017 года мировой доход на рынке больших данных и бизнес-аналитики (Big Data analytics, BDA) достигнет 150,8 млрд. долл., увеличившись на 12,4% относительно предшествующего года. Ожидается, что коммерческие закупки оборудования, программного обеспечения и услуг, связанных с бизнес-аналитикой больших данных, будут демонстрировать среднегодовой темп роста (в сложных процентах — Compound Annual Growth Rate, CAGR¹) 11,9% в период до 2020 года, когда доход составит более 210 млрд. долл. [1].

Самый значительный объем инвестиций в технологии бизнес-анализа больших данных в 2017 году ожидается в банковской сфере, сегментах непрерывного производства, федеральных (центральных) органах государственной власти, а также в сфере профессиональных услуг. В совокупности предполагается, что эти сегменты потратят 72,4 млрд. долл. на развитие методологии Big Data к концу 2017 года. Они останутся крупнейшими сегментами расходов и в 2020 году, когда их общие инвестиции составят 101,5 млрд. долл. При этом самые высокие темпы роста расходов продолжит демонстрировать банковский сектор (CAGR 13,3%) и здравоохранение, страхование, сегмент операций с ценными бумагами и услуг управления инвестициями, а также отрасль телекоммуникационных услуг (каждый — с CAGR 12,8%) [2].

1. Применение систем бизнес-интеллекта и больших данных для совершенствования экономического анализа в российской и мировой практике

По оценкам исследователей Gartner Research, в современных банковских структурах бизнес-аналитика больших данных позволяет лучше контролировать информацию внутри компании и выявлять признаки мошенничества в более сжатые сроки, чем это было возможно ранее, обеспечивая сохранность

финансовых средств. Экономический анализ больших данных, как свидетельствуют данные вышеупомянутого исследования, способствует «более эффективному решению практически всех ключевых задач банков, в числе которых: привлечение клиентов, повышение качества услуг, оценка заемщиков и др. Кроме того, повышая скорость и качество формирования отчетности, увеличивая глубину экономического анализа данных и его охват, участвуя в противодействии мошенническим операциям, отмыванию незаконных средств, эти технологии помогают банкам соответствовать требованиям регулирующих органов» [3].

Преимущественно технологии BDA+BI применяются в банковском секторе для анализа базы клиентов, включая их поведение, предпочтения и ожидания. Например, банк PNC (США) осуществляет поведенческий анализ своих клиентов на различных сайтах (в первую очередь, в онлайн-кабинетах и социальных сетях), обрабатывает информацию о совершенных ими покупках и образе жизни и использует эти данные для гибкого начисления процентных ставок. В свою очередь Commonwealth Bank of Australia (CBA) осуществляет анализ всех транзакций своих вкладчиков, дополняя его поведенческим анализом клиентов на основе изучения данных социальных сетей. Соединяя и обрабатывая совместно эти информационные потоки, в результате банку удалось добиться значительного снижения процента неуплаты по кредитам [4].

В российских условиях также имеется положительный опыт. Например, Уральский банк реконструкции и развития несколько лет назад начал проект по обработке больших данных с помощью технологии BI, работая с данными клиентской базы для формирования персонализированных кредитных предложений, вкладов и других услуг. При этом только за один год применения инновационных ИТ-решений розничный кредитный портфель банка увеличился более чем на 50% [4].

Несмотря на то, что страховые компании также находятся среди лидеров сегмента и заинтересованы в том, чтобы применять технологии Big Data в интеграции с бизнес-интеллектом в экономическом анализе, согласно оценкам агентств Bravura Solutions и Financial Services Council, пока еще лишь немногие структуры (по сравнению с банковским сектором) начали работать в этом направлении [5]. Согласно

¹ Compound Annual Growth Rate (CAGR) — средний темп, с которым инвестиции растут в течение периода более одного года

результатам проведенного опроса, 67% страховых компаний считают, что у них есть лишь ограниченный доступ к пользовательским данным. По мнению респондентов, этих данных достаточно, чтобы персонализировать взаимодействие с клиентами, но недостаточно, чтобы считать их большими данными и прогнозировать на их основе клиентское поведение. Тем не менее, около 30% опрошенных страховых компаний за рубежом уже сегодня используют технологии Big Data и BI в процессе экономического анализа для того, чтобы предугадывать нужды клиентов и создавать персонализированные предложения. Однако для 23,7% организаций вопрос о подобных инновационных решениях информационного характера пока не стоит [5].

Активное развитие технологий анализа больших данных в розничной торговле, — еще одного лидера данного сегмента, — протекает в разных направлениях. Перечислим возможности практического применения, которые можно рассматривать как положительный опыт для российских условий. Так, в мировой практике оффлайн-розничный сектор применяет экономический анализ больших данных с помощью технологий BI для анализа поведения покупателей, планирования маршрутов следования по торговому залу, группировки и расстановки продукции в соответствии с предпочтениями покупателей, планирования закупок, что, в конечном итоге, приводит к повышению объема продаж. В онлайн-розничном секторе на результатах экономического анализа больших данных (даже в случае использования самых минимальных, базовых возможностей таких технологий) может быть построен сам механизм продажи товаров в магазинах: пользователям предлагают товары на основе их предшествующего выбора и персональных предпочтений, информация о которых собирается, например, в социальных сетях, по результатам прошлых покупок и заказов в личных кабинетах онлайн-магазинов. В обоих случаях интеллектуальный экономический анализ Big Data помогает сократить издержки, повысить лояльность клиентов и охватить большую аудиторию, стимулируя покупательский спрос [6].

В подобных условиях экспертами закономерно прогнозируется, что инвестиции в аналитическое программное обеспечение для интеллектуального анализа больших данных продолжат свой рост — до более чем 70 млрд. долл. в 2020 году, при этом основная их часть будет приходиться на обработку опросов конечных пользователей, инструменты отчетности и анализа, а также средства управления хранилища-

ми данных [5]. Следует особо отметить тот факт, что спрос на технологии и сервисы в сфере работы с большими данными и бизнес-аналитикой на мировом рынке стимулируют такие факторы, как увеличение объемов данных, развитие новых технологий и культурный сдвиг в сторону принятия решений, связанных с информацией [2]. Согласно данным агентства PwC, такой стремительный рост популярности применения технологий бизнес-анализа в области Big Data и развития методики в данной области объясняется тем, что этот сегмент по отношению к технологиям бизнес-интеллекта можно назвать следующим уровнем развития: он позволяет организациям пользоваться расширенным функционалом приложений (со встроенными функциями BI), не требуя при этом от ответственных за процесс экономического анализа сотрудников дополнительных специализированных навыков, в то время, как подобный ему сегмент CRM-систем требует иного методического подхода к работе и уровня квалификации персонала [5].

Сегодня в мировой практике наблюдается переход технологий обработки больших данных на более высокий уровень развития, по сравнению с российскими условиями. В опубликованной компанией Gartner последней версии «Цикла зрелости технологий» [7], оценивающей степень развития методологии и практики внедрения, большие данные уже покинули «пик чрезмерных ожиданий, который занимали в 2016 г., и перешли в стадию «избавления от иллюзий», то есть стадию активной апробации и развития методики. Финальной стадии зрелости, по оценкам экспертов, большие данные достигнут через 5–10 лет. Параллельно своего пика достиг и анализ данных, аккумулирующий весь комплекс подходов работы с большими данными, в том числе с использованием технологий бизнес-интеллекта [7].

2. Перспективы взаимодействия систем бизнес-интеллекта и больших данных в экономическом анализе

Несмотря на то, что больше всего инвестиций в развитие методики бизнес-анализа больших данных по-прежнему приходится на США, остальные мировые игроки проявляют в этом направлении большую активность, чем ранее. Так, по данным IDC, на страны Европы, Ближнего Востока, Африки и Азиатско-тихоокеанского региона (без учета Японии) приходится около 45% рынка оборудования, программного обеспечения и услуг бизнес-анализа в сфере больших данных [1].

Растет спрос на применение аналитических инструментов, выполняющих сложную прогнозную аналитику в области больших данных. Одновременно возрастает и вариативность сегмента типовых решений в области бизнес-анализа Big Data — их уже предлагают такие крупнейшие компании, как Microsoft, IBM, Teradata, Oracle и SAP, что способствует более качественному развитию рынка в данном направлении.

78% организаций в мире уверены, что бизнес-анализ больших данных принесет организации очевидную прибыль и преимущества в различных областях в ближайшие 1–3 года. 70% крупных организаций на мировом рынке уже не просто используют свои большие данные для бизнес-аналитики, но и приобретают внешние данные (при этом прогнозируется, что к 2019 году этот показатель вырастет до 100%) [1]. Предполагается также, что все больше организаций начнут извлекать доход из своих данных, продавая их или предоставляя ценный контент на их основе. Следует отметить, что для России данные перспективы пока носят достаточно отдаленный характер. Однако уже появился новый класс предложений — «данные как услуга» (от поисковой системы Google и других), что делает выход на рынок Big Data для многих компаний в мире, в том числе российских (включая представителей малого бизнеса) гораздо более простым и возможным уже в самой ближайшей перспективе.

Согласно прогнозу компании IDC, к 2018 году инвестиции в системы бизнес-анализа больших данных, удобные для самостоятельного применения всеми пользователями, станут обязательными для всех предприятий [1]. При этом прогнозируется, что бюджеты организаций на менее финансово и организационно затратные облачные решения в области бизнес-анализа больших данных будут расти втрое быстрее, чем на локальные решения.

По мнению экспертов Gartner, можно заключить, что рынок Big Data и BI находится на последних стадиях многолетнего перехода от систематического составления отчетности (ИТ-подход) в сторону средств автоматического анализа (бизнес-подход). В результате к современным платформам бизнес-аналитики и анализа больших данных (BI&BDA) уже начинают предъявляться новые требования в части доступности, гибкости и глубине аналитического понимания [7]. Это стимулирует развитие в направлении более простых, гибких и быстрых решений бизнес-аналитики, способных расширить бизнес-возможности за счет глубокого понимания различных источни-

ков данных. В связи с этим одним из перспективных ключевых направлений в зарубежной методике Big Data остается «облачная» бизнес-аналитика и хранение больших данных (например, в облаке Amazon). Эта возможность актуальна для многих компаний, особенно для малого бизнеса, и может быть рекомендована в качестве перспективного (в том числе тестового) варианта для российских организаций. Ее преимущества состоят в том, что это исключает довольно высокие финансовые затраты на приобретение или аренду серверных мощностей для хранения больших данных (которые часто делают невозможным бизнес-анализ больших данных для компаний малого бизнеса), учитывая, что данное оборудование не требуется организации на постоянной основе.

Еще одним перспективным инструментом бизнес-анализа больших данных рассматривается так называемая потоковая (streaming) аналитика, которая позволяет анализировать поступающие данные в реальном времени. Следует также отметить, что для мирового рынка характерно разделение направления Big Data, которое в российских условиях пока воспринимается в общем виде, на множество самостоятельных направлений (например, хранение, обработка, фильтрация данных, построение моделей, изменяемых в режиме реального времени (например, по интересам потребителя), визуализация данных, поиск глубинных связей и т.д.), решающих более узкие специфические задачи экономического анализа. С таким разделением также связано развитие готовых отраслевых решений для малого и среднего бизнеса, работающих как автономные приложения и по моделям SaaS или BDaaS (Big Data as Service) [8].

Говоря о взаимосвязи двух инновационных информационных технологий в экономическом анализе, — BigData и BI, — необходимо отметить, что они развиваются не параллельно и не в соподчинении друг другу, как указывают некоторые специалисты: с 2016 года развитие методики бизнес-анализа больших данных происходит в тесной интеграции с развитием методики бизнес-интеллекта, формируя двухуровневую модель экономического анализа. Первый уровень представляет собой «традиционную» аналитику Big Data, когда большие массивы данных подвергаются экономическому анализу не в режиме реального времени. Новый, второй уровень обеспечивает возможность экономического анализа относительно больших объемов данных в реальном времени, в основном за счет технологий аналитики в памяти (in-memory), включенных сегодня в большинство BI-решений. Следует отметить, что именно второй

уровень открывает для сегодняшнего российского бизнеса совершенно новые возможности, которые могут быть извлечены из результатов экономического анализа накапливаемых организацией данных. Это связано, в частности, с появлением и использованием такого явления, как «озера данных», которые выступают еще одним эволюционным шагом для совершенствования процессов бизнес-аналитики «на лету», с целью влияния на события в то время, когда они происходят. Охарактеризуем кратко данное инновационное явление с позиций его применения в сфере экономического анализа и сформулируем ряд методических рекомендаций в данной области для современных российских компаний.

Озера данных представляют собой репозиторий, хранящий большие объемы данных в сыром, необработанном виде — «до востребования». При этом они обязательно должны быть связаны с аналитической системой, действующей в организации для осуществления экономического анализа. Озера данных позволяют анализировать различные несопоставимые данные в их оригинальном формате. Для современных организаций это является существенным преимуществом, так как не требует больших затрат на хранение и обработку, а данные из «озера» могут быть доступны для бизнес-аналитики любым пользователям в компании в любой промежуток времени [9].

Экономические данные поступают и хранятся в «озере», потому что заранее неизвестно, на какие вопросы они отвечают и какую ценность могут принести для развития бизнеса. Однако не следует исходить из ошибочного мнения (что свойственно многим организациям, приступающим к внедрению в данной области), что «озеро» может заменить собой традиционное хранилище данных. Оно должно обеспечить организации новые аналитические возможности, способствуя одновременно оптимизации расходов на обработку и хранение данных в ходе экономического анализа в организации. В связи с этим для многих современных компаний, внедряющих инструменты экономического анализа больших данных, встает вопрос, будут ли оправданы организационные усилия и финансовые затраты на формирование озер данных, если компания уже реализовала или находится в процессе многолетней реализации хранилища данных для интеграции технологий BI, обрабатывающих Big Data. По оценкам исследователей компании Teradata [9], озеро данных можно использовать для более дешевого хранения информации, в отличие от комплексных и финансово затратных хранилищ, к тому же при интеграции эти техноло-

гии дают больший кумулятивный эффект в области экономического анализа. Важное отличие «озера» — то, что оно может увеличиваться до экстремальных объемов данных при по-прежнему низких издержках, а также может обеспечивать мгновенный доступ к любому данным для всех пользователей. В отличие от традиционных хранилищ, где данные структурированы, «озера» позволяют дешево хранить любые их типы, поступающие из социальных сетей, с различных устройств, в том числе в аудио и видео форматах [9].

В то же время компании должны учитывать то, что накопление «озер», используемых для совершенствования результатов экономического анализа в организации, связано с определенными рисками. Например, в «озеро» персоналом компании помещаются любые данные, и это происходит практически бесконтрольно, без возможности определить их качество. Очень редко решаются задачи обеспечения безопасности данных и контроля доступа к ним. Эффективно работать с данными из «озера» может, в первую очередь, бизнес-аналитик, хотя в то же время они доступны всем бизнес-пользователям в компании. Наконец, главный риск — что со временем эти «озера» рискуют превратиться в «болота» (если их периодически не обрабатывать и не сортировать), непригодные даже для простейших процессов бизнес-аналитики, откуда уже невозможно будет извлечь никакую ценность для бизнеса [10]. Следует отметить, что концепция Big Data исходит из понимания того, что никакие данные организации нельзя «выбрасывать», особенно если цена хранения является достаточно низкой. Любые данные организации изначально рассматриваются, как «умные» (smart data). Если выбросить данные сразу, то об их ценности для бизнеса уже никогда нельзя будет узнать. В то же время необходимо учитывать, что определить потенциальную ценность таких данных достаточно сложно. Именно поэтому современная организация не обязательно должна создавать и держать такое «озеро» у себя. Сегодня уже существует практика аутсорсинга, и различные компании предлагают соответствующие инструменты (в частности, уже отмеченные в статье «облачные» инструменты) для работы с ними.

Заключение

Анализируя рассматриваемую проблему, следует отметить, что с методических позиций многие аспекты интеграции решений в области BI и BigData и их внедрения в процессы экономического анализа сегодня все еще остаются недостаточно проработанными.

Один из таких вопросов заключается в поиске алгоритмов и методических принципов, а также факторов и условий, которые позволят современным компаниям осуществлять экономический анализ больших данных с использованием возможностей бизнес-интеллекта не просто с максимальной бизнес-пользой, но и с преимуществами, позволяющими повысить конкурентоспособность компании на рынке. Как показывает проведенный анализ, методически этот вопрос освещен в научных трудах и практических кейсах компаний гораздо меньше, чем аспекты хранения, аналитической обработки данных и технологии управления Big Data.

Основным выводом данного исследования является понимание того, что современным организациям, функционирующим в условиях высококонкурентного рынка, следует исходить из понимания того, что накопление больших данных не во всех случаях приведет к последующему получению бизнес-выгоды. Как и для любого проекта по оптимизации организационной или коммерческой деятельности, при реализации проекта внедрения бизнес-интеллекта для обработки больших данных компании важно начать с проблемы хранения данных, их систематизации и необходимости для тех или иных целей бизнеса, а именно с вопросов: «как должны в идеале выглядеть результаты?», «как я могу измерить успех?» и «какая информация более полезна для достижения идеаль-

ного результата?». Ответы на эти вопросы могут привести к реализации хранилища Big Data, однако во многих случаях могут и не привести к этому, например, в случае выбора облачных решений как для использования больших данных, так и их обработки для принятия стратегических управленческих решений.

В данном контексте очевиден еще один важный вывод: несмотря на концепцию BI & Big Data, современной компании не следует ставить перед собой задачу обработки всех имеющихся данных с целью повышения качества результатов экономического анализа. Необходимо учитывать, что самый эффективный результат экономического анализа больших данных с помощью возможностей бизнес-интеллекта и других инструментов современных информационных технологий приносит при построении прогнозных моделей, точность которых связана с качеством выборки. В результате реальная значимость Big Data в интеграции с BI в экономическом анализе организации состоит не в том, чтобы построить прогностические модели на основе всех данных (от этого точность моделей не станет выше). Более значимым является использование всего объема данных для сегментации и кластеризации данных, что позволит эффективно строить большое количество моделей для небольших кластеров, решая конкретные задачи экономического анализа на основе применения современных информационных технологий. ■

Литература

1. Большие данные (Big Data): мировой рынок / Tadviser, 2017. [Электронный ресурс]: [http://www.tadviser.ru/index.php/Статья:Большие_данные_\(Big_Data\)_мировой_рынок](http://www.tadviser.ru/index.php/Статья:Большие_данные_(Big_Data)_мировой_рынок) (дата обращения: 14.08.2017).
2. Double-digit growth forecast for the worldwide Big Data and business analytics market through 2020 led by banking and manufacturing investments, according to IDC / IDC, 2015. [Электронный ресурс]: <http://www.idc.com/getdoc.jsp?containerId=prUS41826116> (дата обращения: 14.07.2017).
3. Инфографика: Что могут большие данные в банках? / CNews, 2014. [Электронный ресурс]: http://www.cnews.ru/articles/infografika_что_могут_большие_данные (дата обращения: 14.07.2017).
4. Как крупнейшие банки используют большие данные / CNews, 2014. [Электронный ресурс]: http://www.cnews.ru/articles/kak_kрупнейшие_банки_используют_большие_данные (дата обращения: 14.07.2017).
5. Теория и практика Больших данных в отраслях / Tadviser, 2017. [Электронный ресурс]: http://www.tadviser.ru/index.php/Статья:Теория_и_практика_Больших_данных_в_отраслях#cite_note-4 (дата обращения: 18.08.2017).
6. Инфографика: Как розница использует большие данные / CNews, 2014. [Электронный ресурс]: http://www.cnews.ru/articles/infografika_kak_roznitsa_ispolzuet (дата обращения: 13.07.2017).
7. Gartner hype cycles 2016: Major trends and emerging technologies / Gartner, 2016. [Электронный ресурс]: <https://www.gartner.com/webinar/3377422> (дата обращения: 18.08.2017).
8. Бизнес-аналитика и большие данные в России 2016 / CNews, 2016. [Электронный ресурс]: http://www.cnews.ru/reviews/bi_bigdata_2016 (дата обращения: 14.08.2017).
9. Insights and tools from Big Data and analytics innovators / Teradata, 2017. [Электронный ресурс]: <http://bigdata.teradata.com/> (дата обращения: 16.08.2017).
10. Как развернуть «озеро данных»: 5 советов / CNews, 2015. [Электронный ресурс]: http://www.cnews.ru/articles/kak_razvernut_ozero_dannyh_5_sovetov (дата обращения: 16.08.2017).

Specifics of the integration of Business Intelligence and Big Data technologies in the processes of economic analysis

Stanislav Mitrovic

University of Novi Sad, Republic of Serbia;

Doctoral Student of the Faculty of Economics

Lomonosov Moscow State University

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

E-mail: Mitrovic.Stanislav@hotmail.com

Abstract

The volume of data used for economic analysis of the activities of organizations is growing every year. Despite the fact that all information required for economic analysis is available from various sources, such data are very often useless for analysis from the point of view of their economic potential.

The purpose of this study is to outline a foundation for integrating Business Intelligence and Big Data into economic analysis processes. The theoretical and methodological basis of this study is provided by scientific research, methodological and practical developments of domestic and foreign authors on the application of IT solutions in economic analysis.

According to the results of the research, modern information technologies, in particular, the Business Intelligence and Big Data systems have considerably changed the possibilities for improving economic analysis and reducing decision-making time. From the methodological point of view, many aspects of integration of BI and Big Data solutions and their implementation in the economic analysis processes in Russia's companies remain insufficiently developed. The foreign market of modern information technologies for business analytics has a longer history and is being developed more rapidly.

The main conclusions of the study indicate that modern organizations operating on a highly competitive market should understand that the accumulation of Big Data does not always lead to the expected business benefits. In this context, the conclusion is that a modern company should not set as its goal to process all the available data in order to improve the quality of its economic analysis. It is more significant to use the entire volume of data for segmentation, which allows effective construction of a large number of models for small clusters, solving specific problems of economic analysis based on the application of modern IT systems.

Key words: economic analysis, modern information technologies, Business Intelligence, Big Data, Data Lake, business analytics, Big Data analytics.

Citation: Mitrovic S. (2017) Specifics of the integration of Business Intelligence and Big Data technologies in the processes of economic analysis. *Business Informatics*, no. 4 (42), pp. 40–46. DOI: 10.17323/1998-0663.2017.4.40.46.

References

1. Tadviser (2017) *Bol'shie dannye (Big Data): mirovoy rynek* [Big Data: world market]. Available at: [http://www.tadviser.ru/index.php/Статья:Большие_данные_\(Big_Data\)_мировой_рынок](http://www.tadviser.ru/index.php/Статья:Большие_данные_(Big_Data)_мировой_рынок) (accessed 14 August 2017) (in Russian).
2. IDC (2015) *Double-digit growth forecast for the worldwide Big Data and business analytics market through 2020 led by banking and manufacturing investments, according to IDC*. Available at: <http://www.idc.com/getdoc.jsp?containerId=prUS41826116> (accessed 14 July 2017).
3. CNews (2014) *Infografika: Chto mogut bol'shie dannye v bankakh?* [Infographics: What Big Data can do in banks?]. Available at: http://www.cnews.ru/articles/infografika_chno_mogut_bolshie_dannye (accessed 14 July 2017) (in Russian).
4. CNews (2014) *Kak krupneyshie banki ispol'zuyut bol'shie dannye* [How the largest banks use Big Data]. Available at: http://www.cnews.ru/articles/kak_krupneyshie_banki_ispolzuyut_bolshie (accessed 14 July 2017) (in Russian).
5. Tadviser (2017) *Teoriya i praktika Bol'shikh dannykh v otraslyakh* [Theory and practice of Big Data in industries]. Available at: http://www.tadviser.ru/index.php/Статья:Теория_и_практика_Больших_данных_в_отраслях#cite_note-4 (accessed 18 August 2017) (in Russian).
6. CNews (2014) *Infografika: Kak roznitsa ispol'zuet bol'shie dannye* [Infographics: How retailers use Big Data]. Available at: http://www.cnews.ru/articles/infografika_kak_roznitsa_ispolzuet (accessed 13 July 2017) (in Russian).
7. Gartner (2016) *Gartner hype cycles 2016: Major trends and emerging technologies*. Available at: <https://www.gartner.com/webinar/3377422> (accessed 18 August 2017).
8. CNews (2016) *Biznes-analitika i bol'shie dannye v Rossii 2016* [Business analytics and Big Data in Russia 2016]. Available at: http://www.cnews.ru/reviews/bi_bigdata_2016 (accessed 14 August 2017) (in Russian).
9. Teradata (2017) *Insights and tools from Big Data and analytics innovators*. Available at: <http://bigdata.teradata.com/> (accessed 16 August 2017).
10. CNews (2016) *Kak razvernut' «zero dannykh»: 5 sovetov* [How to rollout a data lake: 5 recommendations]. Available at: http://www.cnews.ru/articles/kak_razvernut_zero_dannyh_5_sovetov (accessed 16 August 2017) (in Russian).

Использование системы веб-аналитики как основы для интеграции с CPA-сервисами

В.И. Жуков

ассистент кафедры инноваций и бизнеса в сфере информационных технологий

Национальный исследовательский университет «Высшая школа экономики»

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

E-mail: vzhukov@hse.ru

М.М. Комаров

кандидат технических наук, PhD

доцент кафедры инноваций и бизнеса в сфере информационных технологий

Национальный исследовательский университет «Высшая школа экономики»

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

E-mail: mkomarov@hse.ru

Аннотация

В современном электронном бизнесе существует большое число способов привлечения потенциальных клиентов на сайт, с помощью как «оффлайн», так и «онлайн» методов. В «онлайн» пространстве бизнес обычно использует несколько каналов привлечения клиентов, отличающихся местом размещения рекламы, моделями оплаты и другими параметрами.

Один из популярных онлайн-методов привлечения клиентов — использование CPA-сетей (cost per action), которые позволяют веб-мастерам размещать на своих сайтах ссылки на рекламируемый сайт и получать вознаграждение за клиента, который приобрел услугу, перейдя по этой ссылке. CPA-сети работают на основе оплаты за достижения целевых действий. Целевое действие может происходить как «онлайн», так и «оффлайн». При этом важнейшей задачей является связь источника привлечения (обычно это определенные UTM-метки) с целевым действием клиента, поскольку вознаграждение для CPA-сети должно осуществляться только за заказы клиентов, источником привлечения которых является эта CPA-сеть. Соответственно, возникают проблемы фиксации таких целевых действий, связывания их с источником привлечения, а также хранения и предоставления доступа CPA-сети к этим данным.

В данной работе дан краткий обзор существующих подходов к решению данной проблемы: анализ логов, использование маркетинговых пикселей и инструментов веб-аналитики. Приведены преимущества и недостатки рассматриваемых методов для решения задачи фиксации целевых действий клиентов и обеспечения доступа CPA-сети к данным. Также описан практический кейс интеграции с CPA-сетью на основе использования возможностей сквозной веб-аналитики. Описаны преимущества, недостатки и ограничения предложенного метода.

Ключевые слова: веб-аналитика, Интернет-маркетинг, CPA-сети, сквозная аналитика, Google-аналитика, CRM.

Цитирование: Жуков В.И., Комаров М.М. Использование системы веб-аналитики как основы для интеграции с CPA-сервисами // Бизнес-информатика. 2017. № 4 (42). С. 47–54.
DOI: 10.17323/1998-0663.2017.4.47.54.

Введение

При ведении бизнеса в онлайн-пространстве обычно возникает задача привлечения потенциальных клиентов с использованием глобальной сети Интернет. Эта задача решается как «оффлайн», так и «онлайн» методами. «Онлайн» методы предполагают, что реклама в той или иной форме размещается на сторонних источниках, со ссылкой на сайт компании. Для этого используются баннеры, контекстная реклама, контекстно-медийная реклама, ссылки в статьях и другие подходы.

Существуют несколько моделей определения вознаграждения за размещение рекламы на стороннем источнике: оплата за размещение на определенный период, оплата за количество показов (cost per view, CPV) или часто применяемая схема оплаты за тысячу показов (cost per millennium, CPM), оплата за клики (cost per click, CPC), оплата за целевые действия (cost per action, CPA) [1]. Также существует модель оплаты за привлеченных клиентов (производящих оплату), которая является частным случаем CPA-модели. Каждая из моделей имеет свои преимущества и недостатки: в случае CPC или CPV результат не гарантирован, в отличие от CPA, где рекламодатель платит только за привлеченных платящих клиентов. Таким образом, использование CPA позволяет значительно снизить риск возможных финансовых потерь [2]. Однако следует отметить, что CPA-модель требует от рекламодателя значительных технических доработок, что не требуется при оплате за показы или клики, поскольку площадка, где будет размещена реклама, обычно предоставляет удобные способы ее публикации.

На основе CPA-модели действуют CPA-сети, которые являются агрегаторами — посредниками между владельцами сайтов (чаще с относительно небольшой посещаемостью) и бизнесом, который размещает рекламу на этих сайтах. Веб-мастера публикуют на своих сайтах рекламу, по которой переходят пользователи, и в случае совершения этими пользователями целевых действий веб-мастер получает оплату за привлеченного клиента. Тип действия и цена за него определяются рекламодателями и называются предложением (или «офером»). CPA-сеть, по сути, является промежуточным звеном между большим количеством веб-мастеров с их сайтами и компаниями, которые хотят привлечь аудиторию с этих сайтов на свою веб-площадку с целью продажи товаров или услуг.

Обычно рекламодатель использует несколько каналов привлечения покупателей в сети Интернет.

Для определения источника привлечения клиента в ссылку добавляют специальные параметры (метки). Обычно это UTM-метки [3], для реферальных программ часто используются специальные идентификаторы, которые присваиваются клиентам [4].

Перед совершением целевого действия клиент может заходить на сайт через несколько каналов (например, с контекстной рекламы, а потом из поиска). Для того, чтобы это учесть, при анализе эффективности источника привлечения определяют модель атрибуции — правила отнесения целевого действия к источнику привлечения или правила распределения ценности конверсии между источниками. В случае CPA-сетей обычно используется атрибуция по последнему значимому клику [5], без учета бесплатных источников трафика (прямые переходы, переходы через органическую выдачу поиска), т.е. атрибуцию по последнему платному клику (last payment click).

Поскольку вознаграждение базируется на количестве и типах целевых действий, возникает проблема фиксации этих действий на стороне рекламодателя и передача достоверных данных о них в CPA-сеть для определения объема вознаграждения, причитающегося веб-мастерам.

1. Подходы к решению проблемы

Существует несколько методов фиксирования целевых действий пользователей на сайте, перешедших из определенного канала и обеспечения доступа к ним для CPA-сети: сервис анализа логов веб-сервера, вызов специальных пикселей (web beacons), а также использование веб-аналитических систем учета (трекеров).

1.1. Сервис анализа логов веб-сервера

В данном случае вся «механика» атрибуции и достижения целевых действий разрабатывается на сервере в виде сервиса анализа логов [6], в которые записываются все обращения к веб-страницам, отправка форм транзакции и другие активности пользователей. В частности, сохраняются реферал и UTM-метка, по которой пользователь перешел на сайт. Также записывается IP клиента, его идентификатор и другая служебная информация (браузер и его версия, операционная система, время обращения и др.) [7]. Данные логи обычно записывают в одном из форматов [8]:

- ◆ NCSA Common Log;
- ◆ NCSA Combined Log;

- ◆ NCSA Separate Log;
- ◆ W3C Extended Log.

Затем разрабатывается сервис для построения отчетов по этим данным, в том числе логика модели атрибуции. При необходимости эти данные объединяют с другими источниками данных компании, например, с CRM-системами. Данные связывают на основе идентификатора клиента, который во всех системах для одного и того же клиента должен быть одинаков. Для интеграции с СРА-сетью реализуется специальное API для доступа к данным об оплатах клиентами, привлеченным из СРА-сети (определяется на основе UTM-меток).

Из достоинств такого подхода можно отметить возможность достижения абсолютной точности в подсчетах и реализации большего количества разных типов целевых действий. Также, за счет возможности интеграции с другими внутренними источниками данных компании, имеется возможность зафиксировать действия, которые произошли за пределами сайта, например, определить пользователей, которые оплатили услугу или продукт наличными деньгами, через терминалы оплаты или банковским переводом.

Из недостатков отметим значительные расходы на разработку такого сервиса (или внедрение готового решения) и его дальнейшую поддержку. Следует упомянуть и проблему доверия, так как бизнес может предоставлять недостоверные данные для СРА-сети, и это крайне сложно проверить.

1.2. Маркетинговые пиксели (web beacons)

Данный подход основан на вызове внешних пикселей при наступлении определенных событий на сайте [9, 10]. На сайте компании размещаются специальные пиксели от СРА-сети, которые вызываются при наступлении целевых действий. Обычно это gif 1x1 px. [11], расположенный на сервере СРА-сети, с вызовом которого передаются параметры и куки пользователя. Эти пиксели вызываются только для тех клиентов, которые были атрибутированы на основе UTM-меток как привлеченные из СРА-сети. Таким образом, СРА-сеть получает информацию о достижении целевых действий. Обычно такое решение применяется для относительно несложных целевых действий, которые происходят на самом сайте (просмотр конкретной статьи, нажатие

на определенные кнопки или отправка заявки).

Достоинствами этого подхода является открытость (поскольку всю логику вызова пикселей может проверить СРА-сеть), а также простое внедрение на сайт клиента.

Недостатком является низкая масштабируемость подхода: при увеличении количества фиксируемых целевых действий каждое из них в отдельности нужно помечать отдельным пикселем с другими параметрами вызова. Стоит упомянуть и трудности, связанные с модернизацией внешней части сайта: необходимо сохранять бизнес-логику вызова этих пикселей. Также существует проблема сложности фиксации событий, которые могут происходить за пределами сайта, например, при оплате услуги клиентом. Еще одним недостатком является неполная точность данных из-за потерь, связанных с отключенным у пользователя JavaScript, средствами блокировки рекламы и др.

1.3. Использование целей в системах веб-аналитики

В этом случае используются уже настроенные цели в системах веб-аналитики (например, Google Analytics¹, Яндекс.Метрика², Piwik³ и др.) и возможности этих систем по установке атрибутов конверсий по источникам. Для специалистов СРА-сети открывается доступ к трекерам веб-аналитики с фильтрацией на соответствующий СРА-сети источник трафика (обычно это характерная метка в UTM-параметрах). На основе этих данных строятся отчеты и определяется объем вознаграждения.

К достоинствам этого метода также относится открытость, поскольку СРА-сеть всегда может проверить корректность сбора целей и атрибуции. Также отметим относительную дешевизну интеграции: обычно счетчики уже стоят на сайте клиента, а цели в них уже настроены и используются для задач анализа внутри компании.

Недостатком являются потери данных, причины которых совпадают с подходом «web beacons».

2. Описание кейса

2.1. Постановка задачи

Компания «А», работающая в области продажи доменов и предоставления хостинговых услуг, приняла

¹ <https://www.google.com/analytics/>

² <https://metrika.yandex.ru/>

³ <https://piwik.org/>

решение использовать CPA-сеть для увеличения объема продаж. Для решения поставленной задачи была выбрана CPA-сеть «В». Компания «А» определила в качестве целевого действия «оплаченный заказ». При этом, поскольку компания продает большой перечень различных услуг с разной маржинальностью, вознаграждение («офер») за покупку привлеченным из CPA-сети клиентом должно также варьироваться в зависимости от типа услуги. Также важно, что большая доля оплат происходит не через сайт, а через терминалы оплаты, банковский перевод и др.

В качестве технического условия интеграции (базы взаиморасчетов) CPA-сеть требовала наличия API для доступа к данным об оплатах услуг пользователей, привлеченных из этой CPA-сети, с передачей категории приобретенной услуги; модель атрибуции — последний платный клик (last payment click). В компании «А» такой сервис отсутствовал, а его реализации требовала значительных ресурсов разработки, что сводило на нет положительный эффект от использования CPA-сети.

Ранее для задач более точного анализа эффективности рекламных кампаний и для глубокого анализа поведения клиента от момента привлечения на сайт до реальной оплаты в компании «А» был внедрен подход сквозной веб-аналитики [3] на основе интеграции Google Analytics с внутренней самостоятельно разработанной системой биллинга через Measurement Protocol (MP API) [12, 13]. Это позволяло фиксировать целевые действия пользователей сайта (оплаты), совершенные за пределами сайта.

Для задачи технической интеграции с CPA-сетью в целях экономии ресурсов было принято решение использовать уже существующую в компании систему «сквозной веб-аналитики», доработав ее.

2.2. Описание интеграции Google Analytics с биллингом

При оформлении клиентом заказа в качестве параметров передается Client ID (cid) из системы Google Analytics (на его основе Google Analytics идентифицирует отдельного пользователя сайта), а при оплате заказа в Google Analytics из биллинга отправляется специализированный запрос — расширенной электронной торговли (Enhanced Ecommerce)⁴ с привязкой к этому клиенту. Такое объединение данных позволяет средствами Google Analytics отследить полный путь клиента от перехо-

да на сайт до реальной оплаты конкретной услуги и прибыли от продажи этой услуги. В хит оплаты передается следующая информация:

- ◆ номер заказа;
- ◆ SKU услуги;
- ◆ категория услуги;
- ◆ название услуги;
- ◆ выручка;
- ◆ прибыль;
- ◆ clientID.

2.3. Эмуляция «last payment click»

Для задач регулярной веб-аналитики в компании система Google Analytics была настроена по стандартной модели атрибуции — последнего непрямого перехода. Однако это не подходит для интеграции с CPA-сетью, так как помимо платных источников может возникнуть доминирование трафика из бесплатных ресурсов, в основном из поискового и реферального трафиков, которые будут «затирать» платный источник привлечения. Такое происходит, если клиент перешел по платному источнику, а дальше, например, перешел по органической выдаче и приобрел товар, при этом в поиске он вводил просто название компании.

Чтобы реализовать модель «last payment click» для интеграции с CPA-сетью был разработан дополнительный Javascript-модуль на фронтенд-стороне сайта, который эмулировал логику «last payment click» через специальные пользовательские параметры Google Analytics, установленные на уровне пользователя (сохраняются в течение всего времени жизни clientID). В эти параметры записывались UTM-метки и отправлялись в Google Analytics. Данные были доступны для анализа и построения отчетов, при этом не нарушая процессы и системы, уже используемые в компании для анализа поведения клиента на сайте.

2.4. API для доступа к данным

Описанная организация позволила сделать Google Analytics общим хранилищем данных о полном треке взаимодействия клиента с сайтом: от источника привлечения до реальной оплаты услуги. При этом, поскольку в Google Analytics уже были реализованы инструменты для получения данных по API, необходимость разрабатывать сервис доступа к этим данным отсутствовала.

⁴ <https://developers.google.com/analytics/devguides/collection/analyticsjs/enhanced-ecommerce?hl=ru>

Для CPA-сети было настроено отдельное представление Google Analytics, включающее CPA-сети источника трафика, который передавался через UTM-метки. Данные из него были доступны по Google Analytics Core API [14].

2.5. Схема и этапы взаимодействия

Схема интеграции с CPA-сетью, представленная на *рисунке 1*, предусматривает выполнение следующих шагов.

Шаг 1. Клиенты с ресурсов веб-мастеров переходят на сайт клиента по URL, со специальными UTM-параметрами, которые содержат данные о самой сети и идентификатор веб-мастера.

Шаг 2. При загрузке сайта, помимо стандартных данных о просмотре страницы (cid и др.) в Google Analytics также отправляются данные об UTM-метках, в специальных параметрах на уровне пользователя.

Шаг 3. При заказе услуги на сайте в биллинг отправляется cid (Client ID, идентификатор клиента в сервисе Google Analytics).

Шаг 4. После оплаты услуги биллинг передает в Google Analytics через MP API хит с составом оплаченных услуг, а также cid, который был привязан к заказу. После получения данных об оплате Google Analytics соотносит их с конкретным пользователем (на основе cid), данные о посещении которого были зафиксированы на шаге 2.

Шаг 5. С помощью Google Analytics Core API CPA-сеть выгружает данные об оплатах услуг и их категориях, совершенных клиентами, которые перешли на сайт компании «А» через ресурсы веб-мастеров, работающих с этой CPA-сетью. На основе этих данных выполняются последующие финансовые взаиморасчеты между компанией, CPA-сетью и веб-мастерами.

3. Преимущества и ограничения предложенного метода

У описанного метода есть ряд преимуществ и ограничений, которые нужно принимать во внимание при решении задачи интеграции с CPA-сетью.

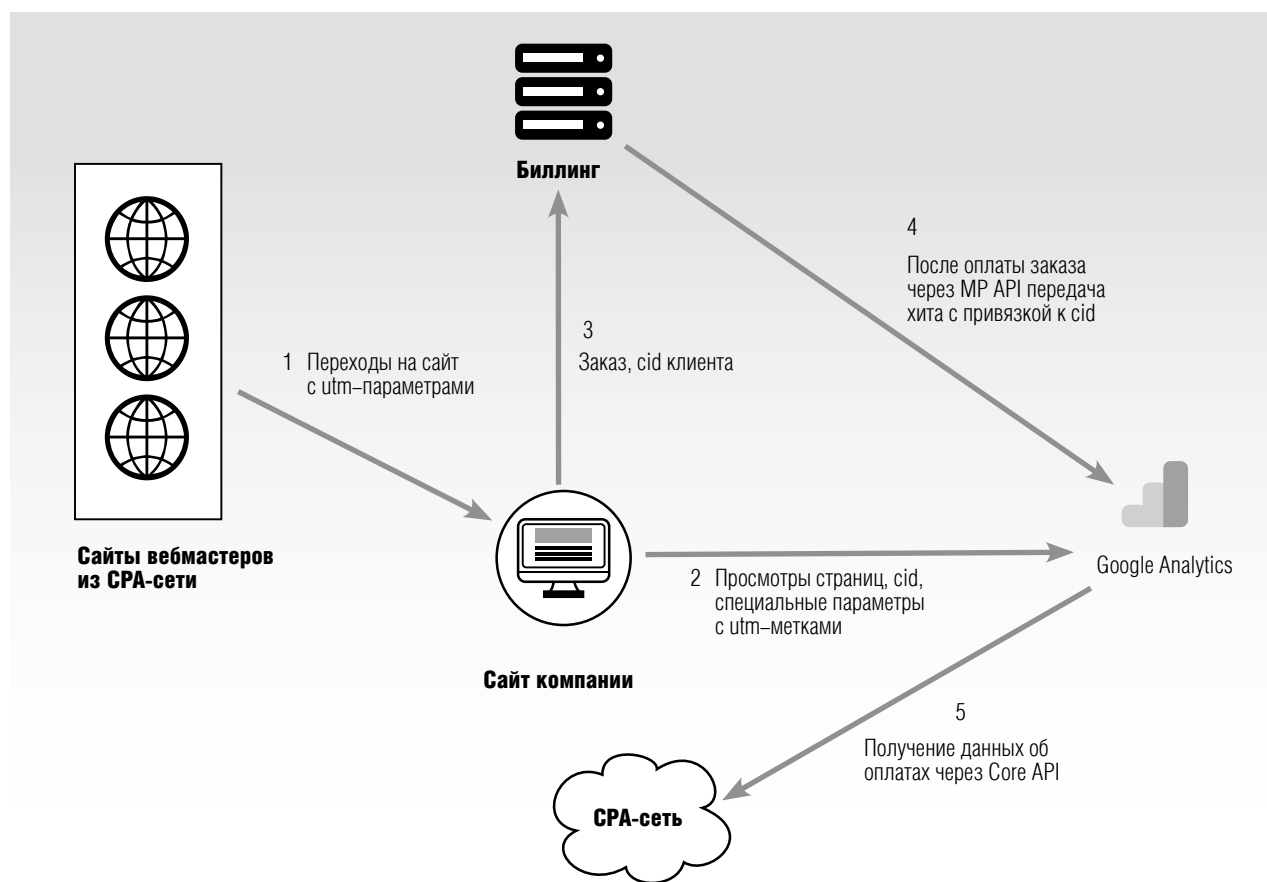


Рис. 1. Схема интеграции с CPA-сетью

Одним из главных преимуществ метода является то, что незначительная доработка инфраструктуры, разработанной для аналитических задач, позволяет решать задачу интеграции. Для компаний, которые используют сквозную веб-аналитику в своей работе, описанный метод позволит:

- ◆ значительно сократить время и затраты на решение задачи интеграции с CPA-сетью;
- ◆ максимально быстро протестировать этот канал для решения маркетинговых задач роста клиентской базы, увеличения прибыли или вывода на рынок новой услуги;
- ◆ в случае неэффективности канала продвижения быстро прекратить работу с ним;
- ◆ возможность учитывать категории услуги при определении вознаграждения, причитающегося веб-мастеру.

В то же время у рассматриваемого метода есть несколько технических ограничений, приводящих к частичной потере данных.

Во-первых, Google Analytics может попадать под аномайзеры и блокировщики рекламы, например, Adblock⁵, что может приводить к потере данных об оплатах привлеченных из CPA-сети клиентов, которые использует подобные блокировщики. Эту проблему полностью можно решить, только перейдя на подход с анализом логов, описанный выше, но из-за дороговизны его разработки и поддержки, обычно применяют простое повышение вознаграждения веб-мастеров на определенный процент. На относительно небольших объемах заказов данный подход оказывается выгодней, чем реализация и поддержка системы анализа логов.

Во-вторых, некоторые клиенты могут узнать об услуге (быть привлеченными на сайт) с одного устройства (например, с мобильного телефона), а приобрести услугу уже с другого. В данной схеме

такая ситуация также приведет к потере данных об оплате. Для ее решения можно использовать один из методов идентификации клиента с разных устройств [13], передавая в Google Analytics дополнительно еще userID из биллинга или CRM. Такое решение сработает, если клиент авторизовался на сайте на каждом из своих устройств.

Заключение

Поскольку данные о реальных оплатах услуг пользователей на сайте связывались и с их данными в системе Google Analytics и эти данные были доступны для получения CPA-сетью, задача интеграции была решена. При этом передача категории товара и его названия позволила определить разные суммы вознаграждения для веб-мастера, в зависимости от услуги, оплаченной привлеченным клиентом.

Благодаря встроенным возможностям Google Analytics (в том числе, в части интеграции с другими источниками данных компании и возможности доступа к этим данным по API), стало возможным решить сразу две задачи бизнеса: связать историю посещения клиента сайтом с данными об оплатах услуг для более глубокого и точного анализа поведения клиента (сквозная веб-аналитика) и обеспечить интеграцию с CPA-сетью.

Описанный метод расширяет метод интеграции с CPA-сетью на основе систем веб-аналитики, обогащая его за счет подхода сквозной веб-аналитики. Метод позволяет фиксировать действия клиента за пределами сайта, но в то же время связывать эти действия с историей клиента и источником привлечения. Данный подход также может использоваться для фиксирования других типов целевых действий, происходящих в «офлайн» среде, которым предшествует посещение сайта. ■

Литература

1. Rzemieniak M. Measuring the effectiveness of online advertising campaigns in the aspect of e-entrepreneurship // *Procedia Computer Science*. 2015. No. 65. P. 980–987.
2. Hu Y., Shin J., Tang Z. Performance-based pricing models in online advertising: Cost per click versus cost per action. Atlanta: Georgia Institute, 2012.
3. Коробков С.А. Основные компоненты системы сквозной маркетинговой Интернет-аналитики для малого бизнеса // *Международный научно-исследовательский журнал*. 2016. № 11–1 (53). С. 48–50.
4. Brown B.C. The complete guide to affiliate marketing on the web: How to use and profit from affiliate marketing programs. Ocala, Florida: Atlantic Publishing Company, 2009.
5. Berman R. Beyond the last touch: Attribution in online advertising. March 5, 2015. Available at: <http://ron-berman.com/papers/attribution.pdf> (accessed 01 December 2017).

⁵ <https://adblockplus.org/ru/android>

6. Grace L.K.J., Maheswari V., Nagamalai D. Analysis of web logs and web user in web mining // International Journal of Network Security and its Applications. 2011. Vol. 3. No. 1. P. 99–110.
7. Goel N., Jha C.K. Analyzing users behavior from web access logs using automated log analyzer tool // International Journal of Computer Applications. 2013. Vol. 62. No. 2. P. 29–33.
8. Booth D., Jansen B.J. A review of methodologies for analyzing websites // Handbook of research on web log analysis / B.J. Jansen, A. Spink, I. Taksa (eds.). Hershey, PA: IGI Global, 2009. P. 143–164.
9. Waisberg D., Kaushik A. Web Analytics 2.0: Empowering customer centricity // Search Engine Marketing Journal. 2009. Vol. 2. No. 1. P. 5–11. Available at: <https://pdfs.semanticscholar.org/f804/b8c0f66b28220d64060e27892dbe0a7a3baa.pdf> (accessed 01 December 2017).
10. Mayer J.R., Mitchell J.C. Third-party web tracking: Policy and technology // 2012 IEEE Symposium on Security and Privacy (SP 2012). San Francisco, California, 20–23 May 2012. P. 413–427.
11. Dwyer C. Behavioral targeting: A case study of consumer tracking on Levis.com // Fifteenth Americas Conference on Information Systems (AMCIS 2009). San Francisco, California, 6–9 August 2009. P. 1–10.
12. Waisberg D. Google Analytics integrations. John Wiley & Sons, 2015.
13. Weber J. Practical Google Analytics and Google Tag Manager for developers. Apress, 2015.
14. Clifton B. Advanced web metrics with Google Analytics. John Wiley & Sons, 2012.

Using a web analytics system as the basis for integration with CPA services

Vyacheslav I. Zhukov

*Assistant Professor, Department of Innovation and Business in Information Technologies
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: vzukov@hse.ru*

Mikhail M. Komarov

*Associate Professor, Department of Innovation and Business in Information Technologies
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: mkomarov@hse.ru*

Abstract

In modern e-business, there are many ways to attract potential customers to the site both with the help of offline and online methods. Companies usually use several channels to attract customers, differing in the placement of advertising, payment models and other parameters.

One of the most popular online methods is using CPA networks, which allow webmasters to place on their websites links to the advertised website and earn rewards for customers who purchased a service by clicking on the link. CPA networks work on the basis of payment for achieving targeted goals. A targeted goal can occur both online and offline. The most important task is to link the source of attraction (usually certain UTM tags) to the target goal of the client, since remuneration for the CPA network should only occur for orders from customers who are drawn to the CPA network. There are problems fixing operations, linking to source of attraction, storing and providing access to these data.

In this paper, we give a brief overview of various approaches to solving the problem: log analysis, use of marketing pixels and web analytics tools. We have analyzed the benefits and challenges of these methods, which were given to solve the task of fixing the target actions of clients and provided access to the data for the CPA network. Also in this article, we have described a practical case of integration with the CPA network based on the use of end-to-end web analytics. The advantages, disadvantages and limitations of the proposed method are set out in this paper.

Key words: web analytics, Internet marketing, CPA networks, end-to-end web analytics, Google Analytics, CRM.

Citation: Zhukov V.I., Komarov M.M. (2017) Using a web analytics system as the basis for integration with CPA services. *Business Informatics*, no. 4 (42), pp. 47–54. DOI: 10.17323/1998-0663.2017.4.47.54.

References

1. Rzemieniak M. (2015) Measuring the effectiveness of online advertising campaigns in the aspect of e-entrepreneurship. *Procedia Computer Science*, no. 65, pp. 980–987.
2. Hu Y., Shin J., Tang Z. (2012) *Performance-based pricing models in online advertising: Cost per click versus cost per action*. Atlanta: Georgia Institute, 2012.
3. Korobkov S.A. (2016) Osnovnye komponenty sistemy skvoznoy marketingovoy Internet-analitiki dlya malogo biznesa [Main components of a system of end-to-end marketing Internet-analytics for small business]. *International Scientific and Research Journal*, no. № 11–1 (53), pp. 48–50 (in Russian).
4. Brown B.C. (2009) *The complete guide to affiliate marketing on the web: How to use and profit from affiliate marketing programs*. Ocala, Florida: Atlantic Publishing Company.
5. Berman R. (2015) *Beyond the last touch: Attribution in online advertising*. Available at: <http://ron-berman.com/papers/attribution.pdf> (accessed 01 December 2017).
6. Grace L.K.J., Maheswari V., Nagamalai D. (2011) Analysis of web logs and web user in web mining. *International Journal of Network Security and its Applications*, vol. 3, no. 1, pp. 99–110.
7. Goel N., Jha C.K. (2013) Analyzing users behavior from web access logs using automated log analyzer tool. *International Journal of Computer Applications*, vol. 62, no. 2, pp. 29–33.
8. Booth D., Jansen B.J. (2009) A review of methodologies for analyzing websites. *Handbook of research on web log analysis* (eds. B.J. Jansen, A. Spink, I. Taksa). Hershey, PA: IGI Global, pp. 143–164.
9. Waisberg D., Kaushik A. (2009) Web Analytics 2.0: Empowering customer centricity. *Search Engine Marketing Journal*, vol. 2, no. 1, pp. 5–11. Available at: <https://pdfs.semanticscholar.org/f804/b8c0f66b28220d64060e27892dbe0a7a3baa.pdf> (accessed 01 December 2017).
10. Mayer J.R., Mitchell J.C. (2012) Third-party web tracking: Policy and technology. Proceedings of *2012 IEEE Symposium on Security and Privacy (SP 2012)*. San Francisco, California, 20–23 May 2012, pp. 413–427.
11. Dwyer C. (2009) Behavioral targeting: A case study of consumer tracking on Levis.com. Proceedings of the *Fifteenth Americas Conference on Information Systems (AMCIS 2009)*. San Francisco, California, 6–9 August 2009, pp. 1–10.
12. Waisberg D. (2015) *Google Analytics integrations*. John Wiley & Sons.
13. Weber J. (2015) *Practical Google Analytics and Google Tag Manager for developers*. Apress.
14. Clifton B. (2012) *Advanced web metrics with Google Analytics*. John Wiley & Sons.

Оценка чистой нетто-премии с использованием дополнительной информации о квантиле функции распределения

Ж.Н. Зенкова

кандидат физико-математических наук

доцент института прикладной математики и компьютерных наук

Национальный исследовательский Томский государственный университет;

СПО «Ассоциация профессиональных актуариев»

Адрес: 634050, г. Томск, пр. Ленина, д. 36

E-mail: zhanna.zenkova@mail.tsu.ru

Е.А. Крайнова

аспирант института прикладной математики и компьютерных наук

Национальный исследовательский Томский государственный университет

Адрес: 634050, г. Томск, пр. Ленина, д. 36

E-mail: lanshakoal@gmail.com

Аннотация

В работе рассмотрена задача повышения точности оценивания чистой нетто-премии за счет использования дополнительной информации о квантиле заданного уровня функции распределения понесенных убытков в страховании ином, чем страхование жизни. Дополнительная информация привлекается путем проектирования эмпирической функции распределения в класс функций с заданным квантилем определенного уровня. Далее методом подстановки модифицированной эмпирической функции распределения в интеграл, определяющий математическое ожидание, получается модифицированная оценка математического ожидания с учетом дополнительной информации об известном квантиле, которая является несмещенной, при этом ее асимптотическая, нормированная на объем наблюдений дисперсия, и, следовательно, среднеквадратическая ошибка не превышает дисперсии классического выборочного среднего, т.е. модифицированная оценка является более точной по сравнению с традиционной для больших объемов наблюдений.

Также исследовано влияние квантиля на значение дисперсии для случаев равномерного, треугольного и нормального распределений. Сделано предположение, что для симметричных распределений минимальное значение дисперсии достигается тогда, когда квантиль совпадает с медианой (центром симметрии). На примере треугольного распределения показано, что априорная информация является наиболее ценной именно при отсутствии симметрии, т.к. позволяет более существенно повлиять на дисперсию.

Модифицированная оценка применялась для расчета чистой нетто-премии на примере реальных данных об убытках по добровольному медицинскому страхованию, понесенных страховой компанией. Показано, что классический расчет привел к ее недооценке, что априори увеличивало риск банкротства компании. В результате перерасчета чистой нетто-премии с учетом знания квантиля риск банкротства снизился.

Работа носит практически значимый характер, по ее результатам страховой компании даны рекомендации.

Ключевые слова: нетто-премия, выборочное среднее, дополнительная информация, квантиль функции распределения, модифицированная оценка среднего, точность оценивания, среднеквадратическая ошибка, страхование иное, чем страхование жизни.

Цитирование: Зенкова Ж.Н., Крайнова Е.А. Оценка чистой нетто-премии с использованием дополнительной информации о квантиле функции распределения // Бизнес-информатика. 2017. № 4 (42). С. 55–63. DOI: 10.17323/1998-0663.2017.4.55.63.

Введение

Страхование — это наукоемкая сфера деятельности, активно использующая современные статистические методы и модели. В данной работе рассматривалось страхование иное, чем страхование жизни [1], при котором выплаты осуществляются только по факту наступления страхового случая и практически не являются детерминированными. В результате их случайного характера страховые компании в основе расчета цены страхового полиса используют так называемую чистую нетто-премию [2], зависящую от размера среднего убытка и вероятности наступления страхового случая. Далее чистая нетто-премия корректируется на повышающий коэффициент, который вычисляется с учетом заданной актуарием вероятности разорения компании, затем добавляются надбавки, позволяющие получить прибыль и осуществить агентские выплаты.

В целях оптимизации своей деятельности страховые компании всегда стремятся найти такую цену полиса, которая, с одной стороны, привлекала бы клиентов, а с другой — максимально точно учитывала случайный характер рассматриваемого явления. При этом используются разнообразные статистические методы и модели, в том числе методы, привлекающие дополнительную информацию различного характера [3–14].

В статье рассмотрена задача улучшения качества оценивания чистой нетто-премии путем привлечения дополнительной информации о квантиле функции распределения понесенных убытков за счет использования более точной модифицированной оценки математического ожидания выплат страховщика по одному страховому случаю. Новая методика расчета апробировалась на реальных данных об убытках в добровольном медицинском страховании и позволила переоценить стоимость полиса, снизив при этом риск разорения страховой компании.

Работа является весьма актуальной и практически значимой, поскольку данный подход ранее не применялся при расчетах чистой нетто-премии.

1. Оценка средней выплаты с учетом дополнительной информации о квантиле функции распределения

Чистая нетто-премия p играет ключевую роль при расчете тарифов в страховании ином, чем страхование жизни [1], и определяется следующим образом [2]:

$$p = z \cdot EX, \quad (1)$$

где X — размер выплаты, случайная величина с функцией распределения $F(x) = P(X < x)$;

EX — математическое ожидание случайной величины;

z — вероятность наступления страхового случая.

В практике страхования ни z , ни EX точно не известны, и их нужно оценить по имеющимся данным о выплатах. При этом нередко исследователь знает некоторые дополнительные сведения об изучаемой случайной величине и ее распределении. Источником таких сведений могут быть условия эксперимента, накопленный опыт эксперта и т.д. Известно, что привлечение дополнительной информации положительно сказывается на свойствах модифицированных статистик, повышая точность оценивания [3–13].

В данной работе рассмотрена дополнительная информация об известном квантиле функции распределения x_q заданного уровня q , т.е. изучалась ситуация, когда

$$F(x_q) = q. \quad (2)$$

Модифицированную оценку математического ожидания с учетом свойства (2) получим с помощью метода подстановки [15] модифицированной эмпирической функции распределения в функционал EX :

$$\bar{X}^q = \int_{-\infty}^{+\infty} x dF_N^q(x).$$

При этом модифицированная эмпирическая функция распределения $F_N^q(x)$ строится путем проектирования обычной эмпирической функции распределения в априорный класс [9] по формуле:

$$F_N^q(x) = \begin{cases} q \cdot \left(\frac{F_N(x)}{F_N(x_q)} \wedge 1 \right) + \\ + (1-q) \cdot \left(\left(\frac{F_N(x) - F_N(x_q)}{1 - F_N(x_q)} \right) \vee 0 \right) \wedge 1, & \text{если } F_N(x_q) \in (0; 1); \\ F_N(x), & \text{если } F_N(x_q) = 0 \text{ или } F_N(x_q) = 1, \end{cases} \quad (3)$$

$$\text{где } F_N(x) = \frac{1}{N} \sum_{i=1}^N C(x - X_{(i)}) - \quad (4)$$

классическая эмпирическая функция распределения, символы $\wedge$ и $\vee$ обозначают максимум (ми-

нимум) из двух величин, $C(y) = \{0 : y \leq 0; 1 : y > 0\}$ – С-функция, $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(N)}$ – вариационный ряд, построенный по независимой выборке $\{X_1, X_2, \dots, X_N\}$ объема N .

Тогда

$$\begin{aligned}\bar{X}^q &= \frac{q}{N \cdot F_N^q(x_q)} \cdot \sum_{i=1}^N X_i \cdot I_{(X_i < x_q)} + \\ &+ \frac{(1-q)}{N(1-F_N^q(x_q))} \cdot \sum_{i=1}^N X_i \cdot I_{(X_i \geq x_q)} = \\ &= \frac{q \cdot N}{N \cdot r} \cdot \sum_{i=1}^r X_{(i)} + \frac{N \cdot (1-q)}{N \cdot (N-r)} \cdot \sum_{i=r+1}^N X_{(i)} = \\ &= \frac{q}{r} \cdot \sum_{i=1}^r X_{(i)} + \frac{1-q}{N-r} \cdot \sum_{i=r+1}^N X_{(i)}.\end{aligned}$$

Здесь случайная величина r определяется как

$$F_N(x_q) = \frac{r}{N},$$

при этом она имеет биномиальное распределение $Bi(N, g)$ [15].

В итоге модифицированная с учетом знания квантиля оценка математического ожидания определяется формулой

$$\bar{X}^q = \begin{cases} \frac{q}{r} \cdot \sum_{i=1}^r X_i \cdot I_{(X_i < x_q)} + \frac{1-q}{N-r} \cdot \sum_{i=1}^N X_i \cdot I_{(X_i \geq x_q)}, & \text{если } r = \overline{1, N-1}; \\ \bar{X}, & \text{если } r = 0 \text{ или } N, \end{cases} \quad (5)$$

где
$$\bar{X} = \frac{1}{N} \cdot \sum_{i=1}^N X_i - \quad (6)$$

классическое выборочное среднее. Также для случая, когда $r = \overline{1, N-1}$ формула (5) может быть определена через порядковые статистики:

$$\bar{X}^q = \frac{q}{r} \cdot \sum_{i=1}^r X_{(i)} + \frac{1-q}{N-r} \cdot \sum_{i=r+1}^N X_{(i)}.$$

Заметим, что формула (5) может быть рассмотрена как обобщение асимптотически нормальной оценки, предложенной в работе [8], где исследовался более широкий случай оценивания функционала с привлечением знания нескольких точно известных значений математических ожиданий, однако не учитывались ситуации, когда модификация не может быть определена. Любопытно, что упомянутая оценка была получена не с помощью метода подстановки, а путем проектирования в априорный класс с помощью расстояния Кульбака-Лейблера.

Найдем математическое ожидание оценки (5):

$$\begin{aligned}E\bar{X}^q &= E\{E(\bar{X}^q | r)\} = E\left\{E\left\{\frac{q}{r} \cdot \sum_{i=1}^N X_i \cdot I_{(X_i < x_q)} + \right.\right. \\ &\quad \left.\left. + \frac{1-q}{N-r} \cdot \sum_{i=1}^N X_i \cdot I_{(X_i \geq x_q)} \mid r = \overline{1, N-1}\right\}\right\} + \\ &+ E\left\{E\left\{\frac{1}{N} \cdot \sum_{i=1}^N X_i \mid r = 0 \text{ или } N\right\}\right\} = (1-q^N - (1-q)^N) \cdot \\ &\cdot E\left\{\frac{q}{r} \cdot r \cdot \int_{-\infty}^{x_q} x d\left(\frac{F(x)}{q}\right) + \frac{1-q}{N-r} \cdot (N-r) \cdot \int_{x_q}^{+\infty} x d\left(\frac{F(x)-q}{1-q}\right) \mid r = \right. \\ &\quad \left. = \overline{1, N-1}\right\} + EX \cdot (q^N + (1-q)^N) = EX,\end{aligned}$$

т.к. выборочное среднее (6) является несмещенной оценкой математического ожидания.

Таким образом, модифицированная оценка (5) является несмещенной.

Найдем дисперсию оценки (5), используя формулу

$$D\bar{X}^q = D(E(\bar{X}^q | r)) + E(D(\bar{X}^q | r)),$$

при этом

$$D\{E(\bar{X}^q | r)\} = D\{EX\} = 0.$$

Поскольку выборочные элементы $X_i, i = \overline{1, N}$ независимы,

$$\begin{aligned}D\left\{\frac{q}{r} \cdot \sum_{i=1}^N X_i I_{(X_i < x_q)} + \frac{1-q}{N-r} \cdot \sum_{i=1}^N X_i I_{(X_i \geq x_q)} \mid r = \overline{1, N-1}\right\} = \\ = \overline{1, N-1} \left\{ \frac{q^2}{r} D\{X_i I_{(X_i < x_q)}\} + \frac{(1-q)^2}{N-r} D\{X_i I_{(X_i \geq x_q)}\} \right\}.\end{aligned}$$

Воспользуемся тем, что, согласно свойствам математического ожидания,

$E\left(\frac{1}{r}\right) \leq \frac{1}{Er}$ в силу выпуклости функции $f(r) = \frac{1}{r}$, при этом сходимость

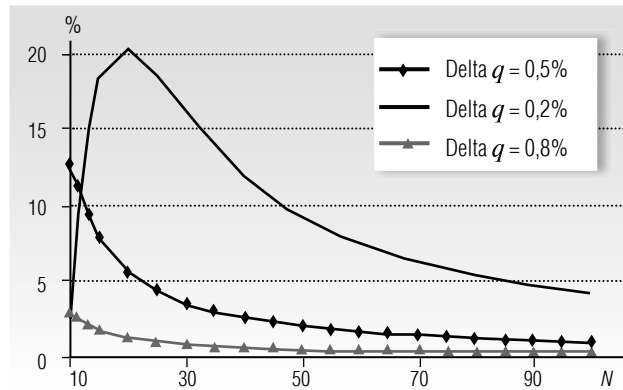
$$E\left(\frac{1}{r} \mid r > 0\right) \xrightarrow{N \rightarrow \infty} \frac{1}{E(r \mid r > 0)}$$

была подтверждена численно для $N = \overline{10, 100}$.

На рисунке 1 приведены графики относительного отклонения

$$\Delta = \frac{\left|E\left(\frac{1}{r} \mid r > 0\right) - \frac{1}{E(r \mid r > 0)}\right|}{E\left(\frac{1}{r} \mid r > 0\right)} \cdot 100\%$$

в зависимости от объема выборки N для $q = 0,2, 0,5$ и $0,8$.


Рис. 1. Отклонение Δ для $q = 0,2, 0,5, 0,8$

Так как $q \in (0,1)$, то $P(r=1, N-1) = 1 - q^N - (1-q)^N \xrightarrow{N \rightarrow \infty} 1$, $P(r=0) = q^N \xrightarrow{N \rightarrow \infty} 0$, $P(r=N) = (1-q)^N \xrightarrow{N \rightarrow \infty} 0$. Для получения асимптотического значения дисперсии достаточно рассмотреть случай $r=1, N-1$. В итоге получим:

$$\begin{aligned} \sigma_q^2 &= \lim_{N \rightarrow \infty} ND\bar{X}^q = \frac{q^2}{q} D\{X_i \cdot I_{(X_i < x_q)}\} + \\ &+ \frac{(1-q)^2}{(1-q)} D\{X_i \cdot I_{(X_i \geq x_q)}\} = q \cdot D\{X_i \cdot I_{(X_i < x_q)}\} + \\ &+ (1-q) \cdot D\{X_i \cdot I_{(X_i \geq x_q)}\} = q \cdot \int_{-\infty}^{x_q} x^2 d\frac{F(x)}{q} - q \cdot \left(\int_{-\infty}^{x_q} x d\frac{F(x)}{q} \right)^2 + \\ &+ (1-q) \cdot \int_{x_q}^{+\infty} x^2 d\left(\frac{F(x)-q}{1-q} \right) - (1-q) \cdot \left(\int_{x_q}^{+\infty} x d\left(\frac{F(x)-q}{1-q} \right) \right)^2 = \\ &= \int_{-\infty}^{x_q} x^2 dF(x) - \frac{1}{q} \cdot L^2 + \int_{x_q}^{+\infty} x^2 dF(x) - \frac{1}{1-q} \cdot R^2 = \\ &= DX + \left(\int_{-\infty}^{x_q} x dF(x) \right)^2 - \frac{1}{q} \cdot L^2 - \frac{1}{1-q} \cdot R^2. \end{aligned}$$

Здесь $L = \int_{-\infty}^{x_q} x dF(x)$, $R = \int_{x_q}^{+\infty} x dF(x)$.

Таким образом, асимптотическая, нормированная на объем выборки дисперсия модифицированной оценки математического ожидания определяется формулой:

$$\sigma_q^2 = DX + (EX)^2 - \frac{1}{q} \cdot \left(\int_{-\infty}^{x_q} x dF(x) \right)^2 - \frac{1}{1-q} \cdot \left(\int_{x_q}^{+\infty} x dF(x) \right)^2. \quad (7)$$

Так как $(EX)^2 = L^2 + 2 \cdot L \cdot R + R^2$,

$$(EX)^2 - \frac{1}{q} \cdot L^2 - \frac{1}{1-q} \cdot R^2 = -\frac{1-q}{q} \cdot L^2 +$$

$$\begin{aligned} &+ 2 \cdot \sqrt{\frac{1-q}{q}} \cdot \sqrt{\frac{q}{1-q}} \cdot L \cdot R - \frac{q}{1-q} \cdot R^2 = \\ &= -\left(\sqrt{\frac{1-q}{q}} \cdot L - \sqrt{\frac{q}{1-q}} \cdot R \right)^2 \leq 0, \end{aligned}$$

из чего следует, что формула (7) может быть представлена в виде:

$$\sigma_q^2 = DX - \left(\sqrt{\frac{1-q}{q}} \cdot \int_{-\infty}^{x_q} x dF(x) - \sqrt{\frac{q}{1-q}} \cdot \int_{x_q}^{+\infty} x dF(x) \right)^2,$$

или, что то же самое [8],

$$\sigma_q^2 = \sigma^2 - \frac{1}{q(1-q)} \cdot \left(\int_{-\infty}^{x_q} x dF(x) - q \cdot EX \right)^2, \quad (8)$$

где $\sigma^2 = DX = ND\bar{X}$.

Для случая, когда известный квантиль есть медиана, будем иметь:

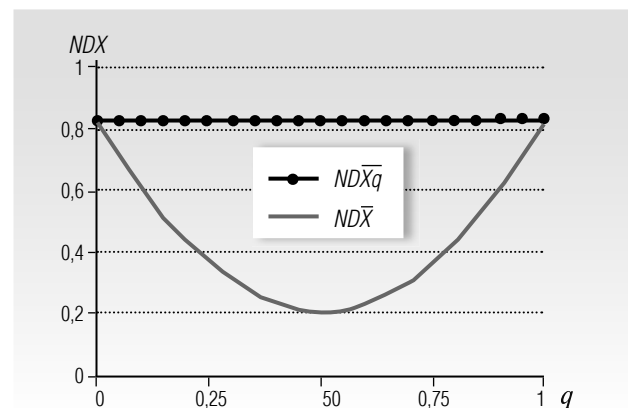
$$\sigma_{0,5}^2 = \sigma^2 - \left(\int_{-\infty}^{x_{0,5}} x dF(x) - \frac{EX}{2} \right)^2.$$

Из формулы (8) очевидно, что $\sigma_q^2 \leq \sigma^2$, т.е. для достаточно больших объемов наблюдений привлечение дополнительной информации о квантиле может привести к снижению среднеквадратической ошибки, а значит, и к улучшению точности оценивания математического ожидания.

Для равномерного в интервале $[0,1]$ распределения формула (8) принимает вид:

$$\sigma_q^2 = \frac{1}{12} - \frac{q(1-q)}{4}. \quad (9)$$

Очевидно, что минимум дисперсии достигается при $q = 0,5$. График зависимости (9) от q приведен на рисунке 2.


Рис. 2. Зависимость σ_q^2 от q для $F(x) = R_{[0,1]}(x)$, $\sigma^2 = \frac{1}{12}$

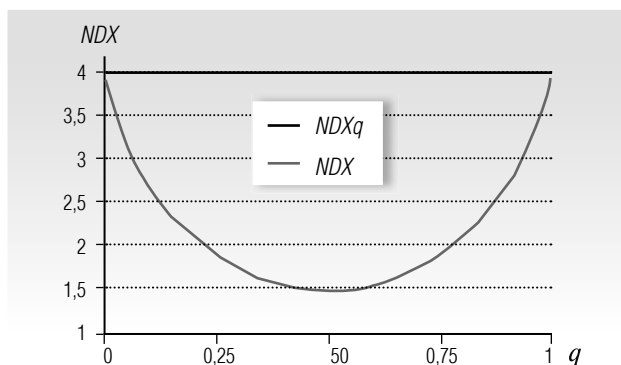


Рис. 3. Зависимость σ_q^2 от q для нормальной функции распределения $N(10,4)$, $\sigma^2 = 4$

На рисунке 3 представлена зависимость σ_q^2 от q для нормальной функции распределения $N(10,4)$.

Для треугольного в $[0,1]$ распределения Симпсона с параметром $c \in (0,1)$ на рисунке 4 приведены графики зависимостей

$$d = -\frac{1}{q(1-q)} \cdot \left(\int_{-\infty}^{x_q} x dF(x) - q \cdot EX \right)^2 \quad (10)$$

второго слагаемого в формуле (8), для разных значений параметра c . Именно d отражает, насколько меньше стала итоговая дисперсия σ_q^2 по сравнению с дисперсией обычной оценки среднего $\sigma^2 = ND\bar{X}$. Заметим, что в сравнении с другими возможными значениями параметра c случай наличия симметрии $c = 0,5$ не дал максимального улучшения. Учет квантиля оказался наиболее значимым для случаев с большей асимметрией ($c = 0,9$ и $c = 0,1$). При этом для случая положительной асимметрии $c < 0,5$ точность повышается, если привлекать информацию о $q > 0,5$, и наоборот (рисунок 5 и таблица 1).

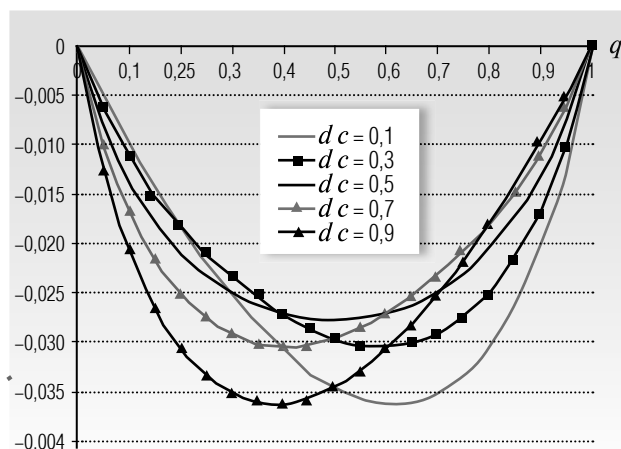


Рис. 4. Зависимость d от q для треугольного в $[0,1]$ распределения с параметром $c \in (0,1)$

На рисунке 5 показано такое сочетание параметра c и квантиля q , которое позволяет получить максимальное улучшение качества оценивания модифицированного среднего в смысле минимума d . В таблице 1 приведены численные значения асимптотических нормированных дисперсий σ_q^2 и минимальных значений d для этих сочетаний c и q .

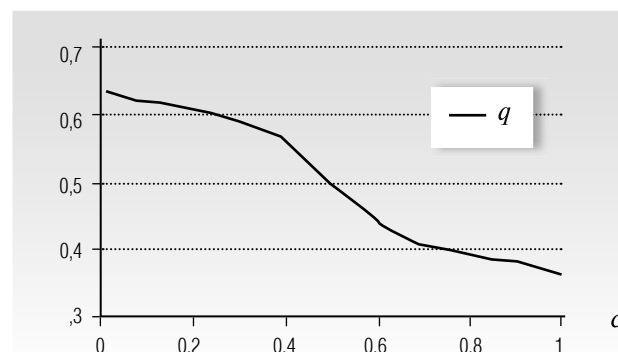


Рис. 5. Сочетание параметра c и квантиля q , максимально улучшающие качество оценивания модифицированного среднего, для треугольного в $[0,1]$ распределения с параметром $c \in (0,1)$

Таблица 1.

Значения асимптотических нормированных дисперсий

$$\sigma_q^2 = \lim_{N \rightarrow \infty} ND\bar{X}^q, \text{ дисперсий } \sigma^2 = ND\bar{X} \text{ и } d = \sigma_q^2 - \sigma^2$$

для оптимальных в смысле минимума d сочетаний c и q

c	q	σ^2	σ_q^2	d
0,01	0,6355	0,055006	0,015375	-0,03963
0,1	0,616	0,050555	0,014224	-0,03633
0,2	0,609	0,046667	0,013541	-0,0331
0,3	0,5935	0,043889	0,013383	-0,03051
0,4	0,564	0,042222	0,013619	-0,0286
0,5	0,500	0,041667	0,013889	-0,02778
0,6	0,436	0,042222	0,013619	-0,0286
0,7	0,4065	0,043889	0,013383	-0,03051
0,8	0,3910	0,046667	0,013541	-0,03313
0,9	0,384	0,050555	0,014224	-0,03633
0,99	0,3645	0,055006	0,015375	-0,03963

Таким образом, для всех рассматриваемых распределений учет дополнительной информации позволил существенно снизить дисперсии оценок. При этом для симметричных распределений минимальное значение дисперсии получается, если квантиль совпадает с медианой — центром сим-

трии. Несимметричные распределения требуют дополнительных исследований, однако на примере треугольного распределения можно предположить, что при положительной асимметрии привлечение информации о квантиле, превышающем медиану, позволяет получить более точную оценку математического ожидания, и наоборот.

Следует отметить, что в работе [7] на примере равномерного распределения показано, что оценка функции распределения с учетом симметрии имеет меньшую дисперсию, чем для случая учета медианы.

В итоге модифицированная оценка чистой нетто-премии (2) с привлечением информации о квантиле может быть рассчитана по формуле

$$p^q = z \cdot \bar{X}^q, \quad (11)$$

при этом она даст более точный результат по сравнению с классическим способом оценивания с использованием выборочного среднего, поскольку среднеквадратическая ошибка $\bar{X}^q$ меньше.

2. Оценка нетто-премии с учетом знания квантиля для добровольного медицинского страхования

Рассматриваемый способ оценивания нетто-премии (11) применялся к реальным данным о выплатах страховой компании по договорам добровольного медицинского страхования. Для сохранения коммерческой тайны исходные значения были масштабированы, название страховой компании не упоминается. Всего за период произошло $N = 239$ страхо-

вых случаев. В таблице 2 приведена выборка $X = \{X_1, X_2, \dots, X_N\}$, при этом показаны неповторяющиеся выплаты Y_i и n_i – количество их повторений в выборке, $N = \sum_{i=1}^k n_i$, k – количество неповторяющихся значений выплат.

При расчете чистой нетто-премии традиционным способом выборочное среднее $\bar{X} = 504,73$ у.е./ед., вероятность наступления страхового случая оценивалась как отношение количества страховых случаев к общему числу страховых договоров, при этом оказалось, что $z = 0,035$. Следовательно, нетто-премия $p = 17,67$ у.е.

Далее выяснилось, что актуарию компании из предыдущего многолетнего опыта работы в области добровольного медицинского страхования известно, что в 90% случаев ущерб при наступлении страхового случая не превышает 750 у.е., т.е. фактически известен квантиль $x_q = x_{0,9}$ уровня $q = 0,9$ исследуемой случайной величины. Воспользовавшись формулой (5), получим $\bar{X}^q = 516,23$ у.е. В итоге модифицированная с учетом знания квантиля нетто-премия $p^q = 18,07$ у.е., что на 2,26% превосходит p .

Таким образом, знание квантиля привело к переоценке чистой нетто-премии. Рассчитанное ранее значение было заниженным, что увеличивало риск банкротства компании. Поскольку модифицированная оценка более точная, если объем выборки большой ($N = 239$), актуарию компании было рекомендовано произвести перерасчет тарифа на добровольное медицинское страхование.

Таблица 2.

Масштабированные данные о выплатах по страховым случаям добровольного медицинского страхования, у.е./ед.

i	Y_i у.е./ед.	n_i ед.	i	Y_i у.е./ед.	n_i ед.	i	Y_i у.е./ед.	n_i ед.	i	Y_i у.е./ед.	n_i ед.	i	Y_i у.е./ед.	n_i ед.	i	Y_i у.е./ед.	n_i ед.
1	16,9	2	11	87,5	1	21	187,5	18	31	437,5	1	41	781,25	1	51	2020,0	1
2	20,6	1	12	93,8	20	22	200	4	32	468,8	5	42	937,5	2	52	2500,0	1
3	25,0	1	13	100,0	5	23	210,0	1	33	487,5	1	43	1000,0	1	53	3125,0	2
4	28,1	1	14	112,5	2	24	218,8	2	34	500,0	3	44	1125,0	2	54	3750,0	3
5	31,3	8	15	125,0	11	25	243,8	2	35	531,3	1	45	1250,0	2	55	4687,5	1
6	37,5	9	16	131,3	2	26	250,0	5	36	562,5	2	46	1343,8	1	56	5000,0	1
7	46,9	1	17	137,5	1	27	281,3	4	37	600,0	1	47	1500,0	1	57	5625,0	1
8	50,0	1	18	156,3	24	28	312,5	19	38	625,0	8	48	1562,5	2	58	8312,5	1
9	56,3	11	19	162,5	1	29	375,0	10	39	739,8	1	49	1662,5	1	59	8437,5	1
10	62,5	19	20	175,0	1	30	406,25	1	40	750,0	2	50	1687,5	1	60	9375,0	1

Заключение

В данной работе реализован новый подход к расчету чистой нетто-премии в страховании ином, чем страхование жизни, за счет привлечения дополнительной информации о квантиле функции распределения при оценивании математического ожидания ущерба. Модифицированная оценка среднего является несмещенной и точнее обычного выборочного среднего, поскольку имеет меньшую асимптотическую среднеквадратическую ошибку, поэтому новая оценка чистой нетто-премии обладает большей точностью для достаточно больших объемов наблюдений.

Также в работе исследовано влияние квантиля на значение асимптотической нормированной дис-

персии модифицированного среднего для случаев равномерного, треугольного и нормального распределений. Показано, что в симметричном случае максимальную точность дает учет медианы; на примере треугольного распределения показано, что для положительной асимметрии лучше привлекать большие значения квантиля, и наоборот.

Предложенный метод апробирован на реальных данных о выплатах страховой компании по добровольному медицинскому страхованию. Привлечение дополнительной информации позволило получить более точное значение чистой нетто-премии. Актуарию компании рекомендовано осуществить перерасчет тарифной ставки с целью снижения возможного риска банкротства компании. ■

Литература

1. Миронкина Ю.Н., Сорокин А.С. Основы актуарных расчетов. М.: Центр ЕАОИ, 2011.
2. Фалин Г.И., Фалин А.И. Теория риска для актуариев в задачах. М.: Мир, Научный мир, 2004.
3. Abu-Dayyeh W.A., Ahmed M.S., Ahmed R.A., Muttalak H.A. Some estimators of a finite population mean using auxiliary information // Applied Mathematics and Computation. 2003. No. 139. P. 287–298.
4. Haq A., Shabbir J. An improved estimator of finite population mean when using two auxiliary attributes // Applied Mathematics and Computation. 2014. No. 241. P. 14–24.
5. Singh H.P., Tailor R. Estimation of finite population mean with known coefficient of variation of an auxiliary character // Statistica. 2005. Vol. LXV. No. 3. P. 301–313.
6. Tarima S., Pavlov D. Using auxiliary information in statistical function estimation // ESAIM: Probability and Statistics. 2006. No. 10. P. 11–23.
7. Дмитриев Ю.Г. О свойствах оценок функции распределения и функционалов при дополнительной априорной информации // Математическая статистика и ее приложения. 1976. № 4. С. 63–76.
8. Дмитриев Ю.Г., Тарасенко П.Ф. Использование априорной информации в статистической обработке экспериментальных данных // Известия вузов. Физика. 1992. № 9. С. 136–142.
9. Дмитриев Ю.Г., Устинов Ю.К. Статистическое оценивание распределений вероятностей с использованием дополнительной информации. Томск: ТГУ, 1988.
10. Журко Е.С., Зенкова Ж.Н. Влияние априорной информации на результаты метода ценообразования на товар-новинку PSM // Материалы III Международной научно-практической конференции «Актуальные проблемы и перспективы развития государственной статистики в современных условиях», г. Саратов, 5–7 декабря 2016 г. Саратов: Саратовстат, 2017. Т. 2. С. 66–68.
11. Журко Е.С., Зенкова Ж.Н. Модификация метода ценообразования PSM с учетом квантиля заданного уровня // Материалы Международной научно-практической конференции «Информационные технологии Сибири», г. Кемерово, 10 ноября 2016 г. Кемерово: КузГТУ, 2016. С. 134–136.
12. Зенкова Ж.Н., Макеева О.Б. Использование информации о квантиле при анализе оборачиваемости оборотных средств // Материалы III Всероссийской молодежной конференции «Математическое и программное обеспечение информационных, технических и экономических систем», Томск, 22–23 мая 2015 г. Томск: ТГУ, 2015. С. 82–87.
13. Зенкова Ж.Н., Муравлева М.А. Асимптотическая несмещенность оценки функции распределения по однократно интервалом цензурированным данным с привлечением информации о симметрии // Материалы X Всероссийской конференции «Новые информационные технологии в исследовании сложных структур», пос. Катунь, 9–11 июня 2014 г. Томск: ТГУ, 2014. С. 114–115.
14. Зенкова Ж.Н., Краковецкая И.В. Непараметрическая оценка Тёрнбулла для интервально-цензурированных данных в маркетинговом исследовании спроса на биоэнергетические напитки // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2013. № 3(24). С. 64–69.
15. Боровков А.А. Математическая статистика. Новосибирск: Наука; Институт математики, 1997.

Estimating the net premium using additional information about a quantile of the cumulative distribution function

Zhanna N. Zenkova

Associate Professor, Institute of Applied Mathematics and Computer Science
National Research Tomsk State University;
SRO "Association of professional actuaries"
Address: 36, Lenin Avenue, Tomsk, 634050, Russian Federation
E-mail: zhanna.zenkova@mail.tsu.ru

Elizaveta A. Krainova

Doctoral Student, Institute of Applied Mathematics and Computer Science
National Research Tomsk State University;
Address: 36, Lenin Avenue, Tomsk, 634050, Russian Federation
E-mail: lanshakoal@gmail.com

Abstract

In this paper, the task of increasing the accuracy of net premium estimations in non-life insurance is considered. Improvements are achieved by involving additional information about a known quantile of loss cumulative distribution function. The additional information is used by projection the empirical cumulative distribution function onto the class of cumulative distribution functions with a certain quantile, and then the modified empirical cumulative distribution function is substituted into the integral that yields the mean value. This allows us to obtain a modified estimation of mean value using additional information about the quantile which is unbiased and its variance is asymptotically less than the variance of the classical sample mean, so that the mean-square error of the modification is also smaller. Therefore, the modified estimation is more accurate than the classical one for a large sample size.

The influence of a quantile value on the variance of the new estimation is studied for uniform, triangular and normal distributions. It is suggested that the minimum of the variance is reached when a known quantile is equal to the median (symmetry center) for symmetrical distribution. Based on Simpson triangular distribution, it was shown that for cases of skewed distributions involving the quantile allows one to decrease the variance more significantly than for symmetrical ones.

The modified estimation of mean value is applied to a real data set for calculation of a net premium. The data contain information about payments for voluntary health insurance of some insurance company. It is demonstrated that the classical method underestimates the net premium, and so it could lead to the company's bankruptcy. After applying the new modified technique, the net premium becomes higher and the bankruptcy risk is reduced as well.

This paper contains practically significant results which make it possible to give important recommendations to an insurance company.

Key words: net premium, sample mean, additional information, cumulative distribution function quantile, modified estimation of mean value, accuracy of estimation, mean-square error, non-life insurance.

Citation: Zenkova Z.N., Krainova E.A. (2017) Estimating the net premium using additional information about a quantile of the cumulative distribution function. *Business Informatics*, no. 4 (42), pp. 55–63.
DOI: 10.17323/1998-0663.2017.4.55.63.

References

1. Mironkina Y.N., Sorokin A.S. (2011) *Osnovy aktuarnykh raschetov* [Fundamentals of actuarial calculations]. Moscow: Center EAOI (in Russian).
2. Falin G.I., Falin A.I. (2004) *Teoriya riska dlya aktuariyev v zadachakh* [Theory of risk for actuaries in tasks]. Moscow: Mir, Nauchny Mir (in Russian).
3. Abu-Dayyeh W.A., Ahmed M.S., Ahmed R.A., Muttalak H.A. (2003) Some estimators of a finite population mean using auxiliary information. *Applied Mathematics and Computation*, no. 139, pp. 287–298.

4. Haq A., Shabbir J. (2014) An improved estimator of finite population mean when using two auxiliary attributes. *Applied Mathematics and Computation*, no. 241, pp. 14–24.
5. Singh H.P., Tailor R. (2005) Estimation of finite population mean with known coefficient of variation of an auxiliary character. *Statistica*, vol. LXV, no. 3, pp. 301–313.
6. Tarima S., Pavlov D. (2006) Using auxiliary information in statistical function estimation. *ESAIM: Probability and Statistics*, no. 10, pp. 11–23.
7. Dmitriev Y.G. (1976) O svoystvakh otsenok funktsii raspredeleniya i funktsionalov pri dopolnitel'noy apriornoy informatsii [On the features of distribution function and functionals estimations in case of additional a priori information]. *Mathematical Statistics and Its Applications*, no. 4, pp. 63–76 (in Russian).
8. Dmitriev Y.G., Tarasenko P.F. (1992) Ispol'zovanie apriornoy informatsii v statisticheskoy obrabotke eksperimental'nykh dannykh [Using a priori information in statistical processing of experimental data]. *Izvestiya vuzov. Physics*, no. 9, pp. 136–142 (in Russian).
9. Dmitriev Y.G., Ustinov Y.K. (1988) *Statisticheskoe otsenivanie raspredeleniy veroyatnostey s ispol'zovaniem dopolnitel'noy informatsii* [Statistical estimation of probability distributions using additional information]. Tomsk: TSU (in Russian).
10. Zhurko E.S., Zenkova Z.N. (2017) Vliyaniye apriornoy informatsii na rezul'taty metoda tsenoobrazovaniya na tovar-novinku PSM [The influence of a priori information on the results of the PSM new product pricing method]. Proceedings of the *III International Scientific and Practical Conference "Actual Problems and Perspectives of State Statistics Development in Modern Conditions"*. Saratov, 5–7 December 2016. Saratov: Saratovstat, vol. 2, pp. 66–68 (in Russian).
11. Zhurko E.S., Zenkova Z.N. (2016) Modifikatsiya metoda tsenoobrazovaniya PSM s uchedom kvantilya zadannogo urovnya [Modification of the PSM pricing method considering quantile of the specified level]. Proceedings of the *International Scientific and Practical Conference "Information Technologies of Siberia"*. Kemerovo, 10 November 2016. Kemerovo: KuzSTU, pp. 134–136 (in Russian).
12. Zenkova Z.N., Makeeva O.B. (2015) Ispol'zovanie informatsii o kvantile pri analize oborachivaemosti oborotnykh sredstv [Using quantile information in current assets turnover analysis]. Proceedings of the *III All-Russian Youth Conference "Mathematical Support and Software for Informational, Technical and Economical Systems"*. Tomsk, 22–23 May 2015. Tomsk: TSU, pp. 82–87 (in Russian).
13. Zenkova Z.N., Muravleva M.A. (2014) Asimptoticheskaya nesmeshchennost' otsenki funktsii raspredeleniya po odnokratno intervalom tsenzurirovannym dannym s privilecheniem informatsii o simmetrii [Asymptotic unbiasedness of distribution function estimation based on single interval-censored data using information about symmetry]. Proceedings of the *X All-Russian Conference "New Information Technologies in Complex Structures Research"*. Katun', 9–11 June 2014. Tomsk: TSU, pp. 114–115 (in Russian).
14. Zenkova Z.N., Krakovetskaya I.V. (2013) Neparametricheskaya otsenka Ternbulla dlya interval'no-tsenzurirovannykh dannykh v marketingovom issledovanii sprosa na bioenergeticheskie napitki [Nonparametric Turnbull estimator for interval-censored data in the marketing research of the demand of bio-energy driks]. *Tomsk State University Journal of Control and Computer Science*, no. 3(24), pp. 64–69 (in Russian).
15. Borovkov A.A. (1997) *Matematicheskaya statistika* [Mathematical Statistics]. Novosibirsk: Nauka; Institute of Mathematics (in Russian).

Комбинированный алгоритм выделения сообществ в графах взаимодействующих объектов¹

С.Ю. Лобанова

студент бакалавриата

Московский институт электроники и математики им. А.Н. Тихонова

Национальный исследовательский университет «Высшая школа экономики»

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

E-mail: s.lobanova@usabilitylab.net

А.А. Чеповский

кандидат физико-математических наук, доцент департамента прикладной математики

Московский институт электроники и математики им. А.Н. Тихонова

Национальный исследовательский университет «Высшая школа экономики»

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

E-mail: aachepovsky@hse.ru

Аннотация

В статье предложен и реализован алгоритм выделения пересекающихся и вложенных сообществ в графах взаимодействующих объектов различной природы. Для этого рассмотрены два классических алгоритма: иерархический агломеративный и основанный на поиске k -клик. Представленный комбинированный алгоритм основан на последовательном их применении. Кроме того, разработаны параметрические опции, отвечающие за действия с сообществами, размеры которых меньше заданного k , а также с единичными вершинами. Варьирование этих параметров позволяет учитывать различия в топологии исходного графа и корректировать тем самым алгоритм.

Проведено тестирование на реальных данных, в том числе на группе графов социальной сети, исследовано качественное содержание полученного разбиения. Для оценки различий, получаемых интегрированным методом и классическими алгоритмами выделения сообществ, была использована общепринятая мера подобия. В результате явно показано, что результирующие разбиения значительно отличаются. Выявлено, что для предложенного в статье подхода показатель числовой характеристики корректности разбиений, модулярности, может быть ниже соответствующего значения при применении других подходов. При этом содержательно результат интегрированного метода зачастую более информативен в силу пересечений и вложенной структуры сообществ.

Представлена визуализация разбиения, получаемого для одного из примеров интегрированным методом на первом и последнем шагах. Наряду с успешно найденным набором параметров интегрированного метода для малых сообществ и отсеченных вершин в случае социальных сетей отмечены некоторые недостатки предложенной модели. Сделаны предложения по развитию данного подхода в виде использования набора параметрических алгоритмов.

Ключевые слова: граф, анализ графа, выделение сообществ, структура графа, анализ социальной сети, большие данные.

Цитирование: Лобанова С.Ю., Чеповский А.А. Комбинированный алгоритм выделения сообществ в графах взаимодействующих объектов // Бизнес-информатика. 2017. № 4 (42). С. 64–73.
DOI: 10.17323/1998-0663.2017.4.64.73.

¹ Работа выполнена при поддержке Российского фонда фундаментальных исследований (проекты №16-29-09546 и №16-07-00641)

Введение

Многие системы в различных областях могут быть описаны как сложные сети, то есть наборы вершин и ребер с нетривиальными топологическими свойствами. Примеры варьируются довольно широко, к ним относятся технологические [1], биологические [2] и социальные [3] системы.

Понятие сообщества — неоднозначное и связано с классификацией объектов по категориям с целью запоминания и поиска информации. В зависимости от контекста оно может быть эквивалентно модулю, классу, группе, кластеру и т.д. Данное понятие играет важную роль в задаче выделения сообществ в сети Интернет [4]. Проблема актуальна при реализации поисковых систем нового поколения, фильтрации данных и их автоматической классификации.

Типичными примерами анализируемых сетей являются популярные сервисы социальных сетей. Тогда в качестве узлов выступают аккаунты пользователей, а связи между ними определяются формальными для конкретного сервиса понятиями «дружбы», «подписки» и т.д. В результате получается граф взаимодействующих объектов. Подобные сети обладают структурой неявных сообществ — групп людей, объединенных по какому-либо дополнительному признаку, например, однокурсники, подписчики ряда публичных персон или тематических пабликов, криминальные элементы.

Выделение сообществ может помочь при анализе графов взаимодействующих объектов. В общем случае эта задача является NP-полной, но, поскольку в этом контексте рассматриваются различные области [5, 6], предлагается использовать алгоритм с корректируемыми опциями, значения которых можно выбирать в зависимости от топологии графа, а значит — от исходной предметной области, из которой получен граф взаимодействующих объектов.

1. Методы разбиения на сообщества

В 2002 году М. Гирван и М. Ньюман [7, 8] представили алгоритм, ставший одним из первых дивизивных методов выделения сообществ. Изначально все вершины ассоциированы с единственным сообществом. Каждый шаг алгоритма Гирван–Ньюмана начинается с вычисления значения характеристики центральности по посредничеству [9] для каждого ребра в графе [8], а затем ребро с наибольшим значением этой характеристики удаляется. Так сеть

разбивается на несвязные компоненты, каждая из которых, в свою очередь, подвергается той же процедуре. Разбиение может проводиться до тех пор, пока в графе не останется ребер, либо пока модулярность [10] результирующего разбиения не достигнет максимума.

Модулярность — числовая характеристика корректности разбиения. На текущий момент это самая популярная мера качества разбиения, прямо или косвенно используемая при исследовании сложных сетей. Как правило, более близкое к реальному разбиению имеет большие значения модулярности.

Модулярность Q определена для существующей структуры сообществ графа следующим образом:

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j), \quad (1)$$

где m — количество ребер,

A_{ij} — элемент матрицы смежности;

k_i — степень вершины i ;

c_i — номер класса, к которому принадлежит вершина;

$$\delta(c_i, c_j) = \begin{cases} 0, & c_i \neq c_j; \\ 1, & c_i = c_j. \end{cases}$$

Центральность по посредничеству для заданного ребра — это доля кратчайших путей, проходящих это ребро, среди всех кратчайших путей:

$$c_B(e) = \sum_{\substack{s \neq t \\ s, t \in V}} \frac{\sigma_{st}(e)}{\sigma_{st}}, \quad (2)$$

где e — заданное ребро;

V — множество вершин графа;

σ_{st} — количество кратчайших путей между вершинами s и t ;

$\sigma_{st}(e)$ количество кратчайших путей между вершинами s и t , проходящих через ребро e .

Расчет $c_B(e)$ является вычислительно затратным [1, 8]. Поэтому в случае крупных сетей целесообразно использовать другие алгоритмы [1].

Быстрый алгоритм иерархического разбиения для общего случая взвешенных графов был предложен Блонделем и другими авторами в 2008 году [11]. Алгоритм агломеративный, на каждом шаге он объединяет текущие вершины графа в супервершины (сообщества) так, чтобы при этом прирост модулярности этого объединения был максимальным. В результате каждой итерации получается иерархическое разбиение следующего уровня. Если достичь

увеличения модулярности не удастся, алгоритм завершает работу. Применение метода ограничено скорее объемом выделенной памяти, чем временем вычисления.

Так как этот алгоритм выделяет сообщества, основываясь на непосредственных окрестностях вершин, на практике полученное разбиение может не соответствовать реальному. Поэтому не всегда очевидно, соответствует ли очередное промежуточное разбиение какому-либо релевантному иерархическому уровню графа. Кроме того, результаты алгоритма зависят от порядка рассмотрения вершин на первом шаге [12].

Описанные выше подходы подразумевают, что сообщества не имеют пересечений друг с другом. При этом принадлежность вершины в реальных сетях редко ограничена единственным сообществом [13]. Как следствие, полученное разбиение может недостаточно точно описывать реальную сеть.

Далее под СРМ будем иметь в виду один из известных алгоритмов, представленный коллективом авторов в работах [14, 15] и позволяющий получать в результате пересекающиеся сообщества. В данном методе нахождение пересекающейся структуры сообществ связано с задачей нахождения k -клик (полных подграфов размера k) в графе. Для фиксированного k алгоритм сперва находит все k -клики сети. Затем, если две k -клики имеют $k - 1$ общую вершину, то они считаются смежными друг с другом (рисунк 1). Тогда сообщество определяется максимальным набором смежных k -клик [15].

Для оценки качества подобных разбиений существует множество модификаций классической модулярности. В данной статье будет рассмотрен вариант, предложенный в работе [13]:

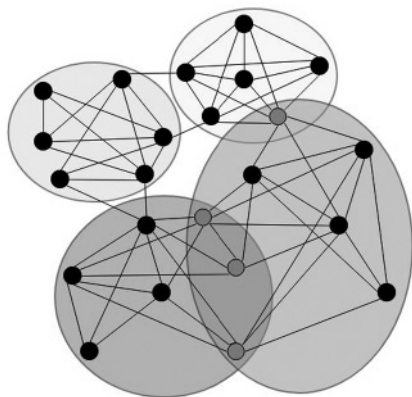


Рис. 1. Выделение пересекающихся сообществ алгоритмом СРМ для 4-клик

$$Q_{ov} = \frac{1}{2m} \sum_{c \in C} \sum_{i, j \in c} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \frac{1}{O_i O_j}, \quad (3)$$

где $C = \{c_1, c_2, \dots, c_{|C|}\}$ – множество пересекающихся сообществ;

Q_i – количество сообществ, которым принадлежит вершина i .

2. Описание комбинированного алгоритма

В статье рассматривается объединение иерархического алгоритма Блонделя и метода разбиения сети на пересекающиеся сообщества СРМ. Комбинирование заключается в их последовательном применении. Идея состоит в том, что сначала сеть разделяется на крупные непересекающиеся сообщества, а затем в каждом из них находятся пересекающиеся сообщества меньшего размера. Далее будем называть данный алгоритм «комбинированным алгоритмом» (КА).

Однако, если методом Блонделя было выявлено какое-то конечное сообщество, не дробящееся на части с точки зрения реальной сети, то запускать СРМ на нем уже нецелесообразно. Поэтому к каждому отдельному выделенному методом Блонделя сообществу с применяется алгоритм СРМ, только если для него верно условие, что доля вершин с высокой центральностью по посредничеству не превосходит заранее заданного значения, зависящего от размера сообщества c :

$$\frac{\sum_{v \in c} \alpha(v)}{n_c} \leq \beta(n_c), \quad \alpha(v) = \begin{cases} c_B(v), & c_B(v) \geq \tilde{\alpha} \\ 0, & c_B(v) < \tilde{\alpha} \end{cases}, \quad (4)$$

где n_c – количество вершин в сообществе c ;

$c_B(v)$ – центральность по посредничеству вершины v ;

$\tilde{\alpha}$ и $\beta(n_c)$ – параметры алгоритма, $0 \leq \tilde{\alpha} < 1$; $0 < \beta \leq 1$.

Относительно большое значение центральности по посредничеству характеризует вершину как возможно связывающую разные сообщества, то есть если ее убрать, граф может разделиться на несколько компонент связности. Схожее соображение используется в алгоритме Гирван–Ньюмана [1, 8]. Наличие достаточного количества таких вершин в сообществе может говорить о том, что оно оформлено не окончательно, и требуется дальнейшее дробление алгоритмом Блонделя. Поэтому условие перехода к СРМ зависит как от самого значения центральности по посредничеству, так и от доли вершин, у которых оно выше некоторого порогового $\tilde{\alpha}$. Пороговые значения для $c_B(v)$ и доли вершин определяются вход-

ными параметрами алгоритма $\tilde{\alpha}$ и β , чтобы иметь возможность влиять на итоговое разбиение.

В результате применения СРМ отсекаются вершины, не входящие в клики, поэтому встает вопрос определения их в сообщества. В качестве решения выступает система из двух новых параметров i_{opt} и m_{opt} комбинированного алгоритма. В зависимости от их значений перераспределяются не только отсеченные вершины, но и маленькие сообщества. Это позволяет частично влиять на результирующую структуру, а также, в отдельных случаях, — избежать объединения листов графа или маленьких сообществ в большие бессмысленные сообщества. Такая особенность разбиения, которую можно назвать «сбором мусора», свойственна многим методам оптимизации модулярности, в том числе алгоритму Блонделя [11, 16, 17].

Для более точного определения и наглядной иллюстрации принципов работы параметров i_{opt} и m_{opt} рассмотрим $C_R = \{C_{R_1}, C_{R_2}, \dots, C_{R_{rc}}\}$ множество сообществ, выделенных алгоритмом Блонделя, где rc их количество. Для удобства записи будем считать, что если отдельное сообщество $C_{R_i} \in C_R$ не удовлетворяет условию запуска СРМ, то его потом дублирует множество $C_{CPM_i} = C_{R_i}$. В противном случае C_{CPM_i} определяется как множество сообществ, полученных в результате работы СРМ внутри C_{R_i} :

$$C_{CPM_i} = \{C_{CPM_{i_1}}, C_{CPM_{i_2}}, \dots, C_{CPM_{i_{mc_i}}}\}, \quad (5)$$

где mc_i количество выделенных методом СРМ сообществ в $C_{R_i} \in C_R$.

Зафиксируем полученное после применения СРМ разбиение как C_f и обозначим структуру сообществ после перераспределения в соответствии с параметром i_{opt} как C_f^i . Параметры i_{opt} и m_{opt} являются обязательными при запуске алгоритма наряду с размером клики k и ограничениями $\tilde{\alpha}$ и β . Структура сообществ C_f преобразуется сначала в соответствии со значением параметра i_{opt} , а затем — со значением параметра m_{opt} .

Первый параметр i_{opt} отвечает за распределение отсеченных на этапе применения СРМ вершин и может иметь следующие значения (таблица 1).

Второй параметр m_{opt} работает с маленькими сообществами. Сообщество считается маленьким, если число вершин в нем не превышает размера клики k , также являющегося входным параметром алгоритма. Такое ограничение выбрано по причине того, что k — тот минимальный размер сообщества, который может быть выделен с помощью метода СРМ. Параметр m_{opt} может иметь следующие значения (таблица 2).

Таблица 1.

Описание принципа работы параметра i_{opt}

None	Отсеченные вершины не принадлежат ни одному сообществу
i2c	Каждая отсеченная вершина выделяется в индивидуальное (единичное) сообщество внутри того сообщества, в котором оно находилось до применения СРМ
i2r	Каждая отсеченная вершина выделяется в индивидуальное (единичное) сообщество, не имеющее пересечений с другими
ia2c	Группа из всех отсеченных вершин, ранее находящихся в одном сообществе до применения СРМ, образует внутри него сообщество
ia2r	Все отсеченные вершины образуют единое сообщество, не имеющее пересечений с другими

Таблица 2.

Описание принципа работы параметра m_{opt}

None	Структура сообществ не изменяется
mdel	Маленькие сообщества расформируются, при этом, если вершина принадлежит другим сообществам, то она в них остается. В противном случае она не будет принадлежать ни одному сообществу
m2c	Все маленькие сообщества, находящиеся в общем для них сообществе (т.е. вложенные в него), объединяются в нем в одно общее сообщество, не имеющее вложенных сообществ
m2r	Все маленькие сообщества образуют одно общее сообщество, не имеющее пересечений с другими, даже если вершины принадлежали и другим сообществам

Каждое из преобразований производится однократно, поэтому в случае появления новых маленьких сообществ (возможно для $m2r$), они не будут объединены с новым общим сообществом. Итоговая структура сообществ C_f^* сети формируется из C_f в соответствии с принципами, продемонстрированными в таблицах 1 и 2.

Отдельно следует рассмотреть комбинацию параметров со значениями «i2c/None». В этом случае при выделении небольшого сообщества часто выходит так, что все его члены также выделяются и в собственные индивидуальные сообщества, а большие сообщества разделяются на одну большую часть и несколько индивидуальных. Во избежание нагромождения сообществ в итоговой структуре авторами применялась модификация: в описанном для маленьких сообществ случае индивидуальные сообщества расформируются и, таким образом, остается только само маленькое сообщество. Для большого сообщества остается оно само и выделенные внутри него индивидуальные сообщества (то есть убирается та часть внутреннего разбиения, в которой находят-

ся все вершины исходного сообщества, за исключением выделенных в единичные сообщества). На рисунке 2 приведена возможная картина результата работы КА внутри одного из сообществ, полученных на первом шаге. Подобный результат не дают примененные по отдельности алгоритмы Блонделя и СРМ. Между тем, в ряде случаев при анализе графов такое разбиение будет крайне информативно.

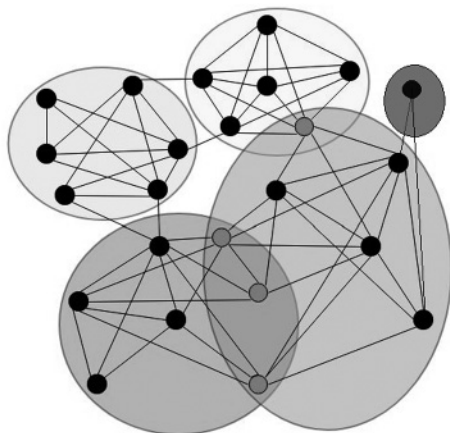


Рис. 2. Возможное итоговое разбиение на сообщества внутри одного из полученных на первом этапе работы КА

3. Тестирование

В ходе тестирования были оценены результаты работы КА на разных типах графов и произведены сравнения с результатами алгоритмов Блонделя и СРМ по отдельности. Также сравнение было произведено с результатами некоторых классических алгоритмов: GN, быстрой жадной оптимизации модулярности FG [18] и дающего сообщества с пересечениями алгоритма CONGA (Cluster Overlap Newman-Girvan Optimized Algorithm) [19]. Рассмотрена зависимость результатов КА от параметров k , $\tilde{\alpha}$, β , i_{opt} , m_{opt} . Размер клики k рассматривался в диапазоне $k \in [3, 10]$. Для каждого разбиения вычислено значение Q_{ov} , а для разбиений, полученных КА, рассчитана их мера подобия ω другим разбиениям.

В работах [20, 21] предложен индекс ω (Omega Index) для оценки подобия двух разбиений, исходя из количества пар вершин, согласованных в одном и том же количестве сообществ. Пусть C_1 и C_2 различные разбиения, K_1 , K_2 соответствующие им количества сообществ. Пусть $t_j(C)$ — множество пар вершин, встречающихся вместе в разбиении C ровно j раз. Тогда ω определяется следующим образом:

$$\omega(C', C'') = \frac{\omega_u(C', C'') - \omega_e(C', C'')}{1 - \omega_e(C', C'')}, \quad (6)$$

где $\omega_u(C', C'')$ — наблюдаемый показатель, а $\omega_e(C', C'')$ — ожидаемый:

$$\omega_u(C', C'') = \frac{1}{m} \sum_{j=0}^{\min(K_1, K_2)} |t_j(C') \cap t_j(C'')|,$$

$$\omega_e(C', C'') = \frac{1}{m^2} \sum_{j=0}^{\min(K_1, K_2)} |t_j(C')| \cdot |t_j(C'')|.$$

Значение $\omega(C', C'') = 1$ в случае абсолютного совпадения двух разбиений.

Авторами были изучены результаты применения КА и выбранных классических алгоритмов на более чем 70 графах взаимодействующих объектов различной природы, таких как социальные сети, сети цитирования ученых, белок-белковые взаимодействия, сети инфраструктуры. В ряде случаев для графов из одной группы и со схожими топологическими параметрами выявлены наиболее оптимальные значения корректирующих параметров для КА. Программная реализация выполнена на Python 3.5, а для работы с графами были использованы средства стандартных и открытых библиотек, содержащих, в том числе, и классические алгоритмы. Реализованы полуавтоматические блоки тестирования для поиска лучших по Q_{ov} значений параметров. Визуализация полученных разбиений на графах обеспечивается сторонним программным обеспечением.

Более подробно покажем результаты на эго-графах, полученных из социальной сети ВКонтакте (таблица 3). Эго-графами в данном случае были графы, сформированные поиском в ширину от одной вершины с глубиной 1 и добавлением всех ребер между полученными таким образом вершинами. При этом в качестве стартовой брались вершины со степенью в диапазоне примерно от 100 до 500. Это позволило получить не только графы с разными топологическими свойствами, но и удобных для качественного анализа размеров. Именно изучение содержательной стороны в получаемых разбиениях и выявление наиболее подходящих значений параметров КА было одной из основных целей данного исследования.

На данных графах КА в большинстве случаев ожидаемо выделял большее число сообществ (таблица 4). Для данной группы графов, представляющих эго-графы одной и той же социальной сети, наиболее эффективным по содержанию разбиений показал себя вариант параметров $i2c/m2r$ и $k = 4$. Качественный экспертный анализ состава выделенных рассмотренными алгоритмами групп, особенно маленьких сообществ из нескольких вершин, показал

Таблица 3.

Графы социальной сети для тестирования

	n	c	avk	dst	$avntr$	$avclu$	$maxcl$	bet	clo
G_1	105	792	15,1	0,1451	38,04	0,712	15	44,5	0,5434
G_2	120	1226	20,4	0,1717	52,39	0,668	14	49,3	0,5507
G_3	158	1352	17,1	0,1090	44,41	0,767	17	69,9	0,5309
G_4	158	1596	20,2	0,1287	63,80	0,638	21	68,4	0,5374
G_5	187	1873	20,0	0,1077	51,11	0,674	16	83,0	0,5308
G_6	229	1726	15,1	0,0661	26,67	0,633	14	106,5	0,5184
G_7	439	4143	18,9	0,0431	41,47	0,602	16	209,6	0,5117
G_8	461	4299	18,7	0,0405	49,36	0,665	17	220,7	0,5110

Обозначения: n – число вершин; c – число ребер; avk – средняя степень вершин; dst – плотность связей; $avntr$ – среднее число треугольников; $avclu$ – средний кластерный коэффициент; $maxcl$ – размер максимальной клики; bet – усредненная центральность по посредничеству; clo – усредненная центральность по близости.

Таблица 4.

Результаты тестов

	$s(KA)$	Q_{ov}^{KA}	$s(BL)$	Q_{ov}^{BL}	$s(CP)$	Q_{ov}^{CP}	$s(GN)$	Q_{ov}^{GN}	$s(FG)$	Q_{ov}^{FG}	$s(CN)$	Q_{ov}^{CN}
G_1	38	0,213	4	0,320	4	0,167	17	0,250	3	0,315	3	0,245
G_2	21	0,367	4	0,385	4	-0,013	20	0,329	4	0,371	4	0,136
G_3	28	0,650	4	0,614	5	0,390	15	0,581	5	0,581	5	0,494
G_4	39	0,327	4	0,342	3	0,016	47	0,175	3	0,313	4	0,100
G_5	36	0,048	5	0,369	3	-0,008	27	0,250	5	0,325	110	0,011
G_6	51	0,224	5	0,465	4	-0,029	22	0,408	6	0,393	4	0,273
G_7	90	0,501	7	0,572	8	-0,149	54	0,536	8	0,509	5	0,304
G_8	99	0,349	7	0,569	11	-0,272	53	0,514	6	0,552	8	0,195

Обозначения: $s(KA)$ – число выделенных КА сообществ; $s(BL)$ – число выделенных алгоритмом Блонделя сообществ; $s(CP)$ – число выделенных CPM сообществ; $s(GN)$ – число выделенных GN сообществ; $s(FG)$ – число выделенных FG сообществ; $s(CN)$ – число выделенных CONGA сообществ; каждое из Q_{ov}^A значение Q_{ov} для разбиения, полученного алгоритмом A.

Таблица 5.

Значения ω между результатами разбиения графов КА и другими алгоритмами

	$\omega(BL)$	$\omega(CPM)$	$\omega(GN)$	$\omega(FG)$	$\omega(CONGA)$
G_1	0,0811	-0,0478	0,0209	0,0729	0,0226
G_2	-0,0125	0,0139	0,0152	-0,0087	0,0655
G_3	0,0160	-0,0166	-0,0017	0,0043	0,0149
G_4	0,0459	-0,0077	-0,0098	0,0047	-0,0147
G_5	0,0651	-0,0058	0,0232	0,0133	-0,0055
G_6	0,0915	0,0003	0,0004	0,0159	0,0088
G_7	0,0452	0,0009	0,0061	0,0020	0,0066
G_8	0,0152	0,0137	-0,0070	0,0096	0,0079

Обозначения:

$\omega(BL)$ – для разбиения графа с помощью КА и алгоритма Блонделя;
 $\omega(CPM)$ – для разбиения графа с помощью КА и алгоритма CPM;
 $\omega(GN)$ – для разбиения графа с помощью КА и алгоритма GN;
 $\omega(FG)$ – для разбиения графа с помощью КА и алгоритма FG;
 $\omega(CONGA)$ – для разбиения графа с помощью КА и алгоритма CONGA.

наибольшую эффективность именно КА. Так, например, алгоритмы Блонделя, CPM и FG ожидаемо не оставляли большие сообщества, а среди выделяемых алгоритмом GN наиболее маленьких сообществ (в том числе из 1 элемента) для каждого из графов не все были содержательно обоснованы и в ряде случаев с учетом имеющихся в графе ребер должны были бы быть объединены с другими.

При этом необходимо отметить, что значение Q_{ov}^{KA} не всегда было больше аналогичного для иных алгоритмов в силу нахождения КА иного локального экстремума данной меры качества, что частично подтверждается рассчитанными значениями меры подобия ω данных разбиений (таблица 5). Как видно из таблиц, сильно различается не только число полученных КА и остальными алгоритмами сообществ, но и сами наборы объединенных в группы вершин, ведь в силу особенностей рассмотренных классических алгоритмов, некоторые виды разбиений не

могут быть получены по результатам их применения даже теоретически (например, ситуация из *рисунка 2* для СРМ нереализуема). Все это совокупности показывает, что полученное при КА разбиение, как правило, существенно отличалось от остальных.

Получаемое в рамках КА разбиение является иерархическим, с пересекающимися и вложенными сообществами, что дает возможность более полного понимания структуры исследуемого графа и лучше соответствует реальной системе взаимодействия объектов некоторой природы, в частности, пользователей социальных сетей. В качестве иллюстрации рассмотрим пример для G_3 подробнее.

На первом шаге КА выделяет в графе G_3 четыре сообщества: C_{R_1} , C_{R_2} , C_{R_3} , C_{R_4} размеров 44, 41, 37, 36 соответственно (*рисунок 3*). Данный граф является очень показательным, поскольку выделенные алгоритмом Блонделя сообщества близки по размерам. Тут C_{R_1} является типичным «мусорным» сообществом, содержащим исходную вершину эго-графа, листовые вершины и мелкие сообщества. Остальные три выделенных сообщества содержательно более обоснованные, они представляют из себя группы пользователей, большинство из которых не имеют связей с представителями других двух сообществ.

Далее после применения следующих шагов КА на графе G_3 эти четыре сообщества преобразуются в 28 следующим образом. Внутри трех содержательных плотных сообществ происходят минимальные изменения. Так, внутри C_{R_3} алгоритм нашел одну вершину, менее связанную с остальными, для которой согласно опциям *i2c/None* было выделено отдельное сообщество внутри исходного, остальные 40 вершин попали во второе внутреннее сообщество

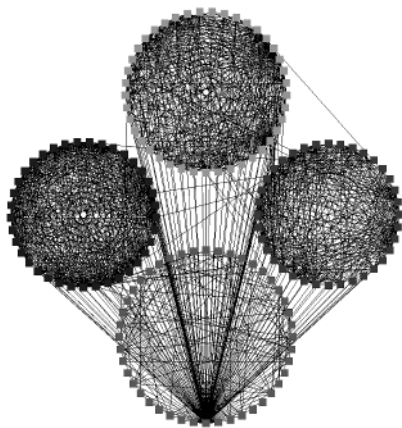


Рис. 3. Выделенные алгоритмом Блонделя в графе G_3 четыре сообщества

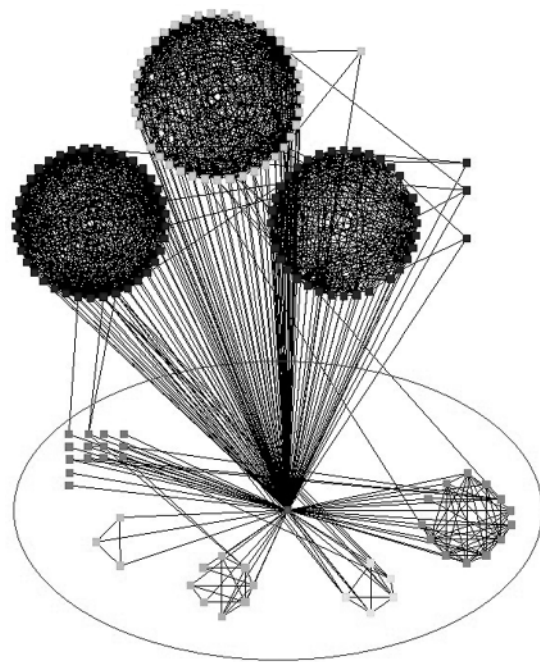


Рис. 4. Выделенные КА сообщества

для C_{R_2} . Сообщество C_{R_3} осталось совсем без изменений, а для C_{R_4} были найдены свои три вершины, которые выделились в свои индивидуальные внутренние сообщества, аналогично ситуации с C_{R_2} , создав при этом еще и большой остаток — сообщество из 33 вершин.

Основные изменения произошли на втором шаге КА внутри C_{R_1} . Так, в этом сообществе были выделены вложенные еще 4 других, пересекающихся по исходной вершине эго-графа. Кроме того, аналогично ситуации в других двух сообществах верхнего уровня были выявлены 14 индивидуальных сообществ внутри исходного (*рисунок 4*).

В другом эго-графе G_4 с тем же числом вершин, но с иной структурой (*таблица 3*) после применения к нему КА было выявлено 39 сообществ. При этом изначально на первом шаге так же были выделены 4 больших группы (здесь $|C_{R_1}| = 50$, $|C_{R_2}| = 44$, $|C_{R_3}| = 44$, $|C_{R_4}| = 20$), внутри которых уже на втором шаге произведен поиск пересекающихся вложенных сообществ. При этом C_{R_2} и C_{R_3} не претерпели изменений, но внутри C_{R_1} и C_{R_4} выделено 28 и 7 вложенных сообществ с пересечениями различных размеров соответственно.

Необходимо отметить, что по оценке самих пользователей социальных сетей, чьи эго-графы были проанализированы в процессе работы, иные разбиения, полученные сравниваемыми методами, несли в себе меньшую содержательную информацию,

хотя в ряде случаев показатель Q_{ov} был выше, чем для КА (таблица 4). В том числе, менее информативными были признаны результаты алгоритмов, взятых за составляющие КА, а, например, для GN разбиение оказалось частично излишним.

Заключение

Таким образом, разработан и представлен новый алгоритм поиска эффективного выделения неявных сообществ в графах взаимодействующих объектов различной природы, названный авторами «комбинированным алгоритмом» (КА). На наборе графов из разных отраслей, взятых из открытых источников, применены КА и широко известные классические алгоритмы. Изучены полученные разбиения графов на сообщества. Сравнение значения модулярности Q_{ov} на результирующих разбиениях и меры подобия ω между результатом КА и остальных алгоритмов показали, что в большинстве случаев имеются значительные различия в получаемых наборах сообществ. Помимо этого, для набора графов из социальной сети проведен содержательный анализ полученных разбиений, который показал, что результат КА при выбранных на данном наборе опциях более точно отражает группы близких связей между аккаунтами.

Предложенный алгоритм запускался на графах при разных параметрах. В ходе тестирования было обнаружено, что использование параметров ограничений $\tilde{\alpha}$ и β на данных графах не повышает качества результирующего разбиения, так как результат, отличный от результатов применения алгоритма Блонделя и КА (при отсутствии ограничений) достигался лишь в случае $\tilde{\alpha} > 0$ и $\beta = 0$. Дальнейшее исследование проводилось при значениях параметров $\tilde{\alpha} = 0$ и $\beta = 1$. Таким образом, метод СРМ запускался на всех полученных алгоритмом Блонделя сообществах.

Проведенное исследование показало, что наиболее подходящие значения дополнительных параметров для социальных сетей — это *i2c/None* (в рамках обозначений КА). Вместе с тем нельзя утверждать, что для графов взаимодействующих объектов иной природы эта комбинация будет эффективнее остальных. Поэтому авторы считают целесообразными дальнейшие исследования на тему конструирования и реализации параметрических алгоритмов выделения сообществ, а также возможного использования автоматизированных средств по подбору подходящих опций и значений параметров, исходя из топологических особенностей и исходной природы графов. ■

Литература

1. Guo Y., Xu K. An improved algorithm for finding community structure in networks with an application to IPv6 backbone network // *Frontiers of Computer Science in China*. 2007. Vol. 1. P. 459–467.
2. Wu F., Chen L., Wang J., Alhajj R. Biomolecular networks and human diseases // *BioMed Research International*. 2014. Vol. 2014. P. 1–2.
3. White H.C., Boorman S.A., Breiger R.L. Social structure from multiple networks. I. Blockmodels of roles and positions // *American Journal of Sociology*. 1976. Vol. 81. No. 4. P. 730–780.
4. Graph structure in the Web / A.Z. Broder [et al.] // *Computer Networks*. 2000. Vol. 33. No. 1–6. P. 309–320.
5. Ahuja M.S., Singh J., Neha. Practical applications of community detection // *International Journal of Advanced Research in Computer Science and Software Engineering*. 2016. Vol. 6. No. 4. P. 412–415.
6. Pennacchioli D., Coscia M., Pedreschi D. Overlap versus partition: Marketing classification and customer profiling in complex networks of products // *Proceedings of 2014 IEEE 30th International Conference on Data Engineering Workshops (ICDEW 2014)*. Chicago, 31 March – 4 April 2014. P. 103–110.
7. Newman M., Girvan M. Finding and evaluating community structure in networks // *Physical Review E*. 2004. Vol. 69. No. 2. P. 1–15.
8. Girvan M., Newman M. Community structure in social and biological networks // *Proceedings of the National Academy of Sciences*. 2002. Vol. 99. No. 12. P. 7821–7826.
9. Newman M. *Networks: An introduction*. Oxford: Oxford University Press, 2010.
10. Newman M. Modularity and community structure in networks // *Proceedings of the National Academy of Sciences*. 2006. Vol. 103. No. 23. P. 8577–8582.
11. Blondel V.D., Guillaume J.L., Lambiotte R., Lefebvre E. Fast unfolding of communities in large networks // *Journal of Statistical Mechanics: Theory and Experiment*. 2008. No. 10. P10008.
12. Fortunato S. Community detection in graphs // *Physics Reports*. 2010. Vol. 486. No. 3–5. P. 75–174.
13. Shen H.W., Cheng X.Q., Cai K., Hu M.B. Detect overlapping and hierarchical community structure in networks // *Physica A: Statistical Mechanics and its Applications*. 2009. Vol. 388. No. 8. P. 1706–1712.
14. Palla G., Derényi I., Farkas I., Vicsek T. Uncovering the overlapping community structure of complex networks in nature and society // *Nature*. 2005. Vol. 435. No. 7043. P. 814–818.
15. Palla G., Ábel D., Farkas I.J., Pollner P., Derényi I., Vicsek T. (2009) *k*-clique percolation and clustering / G. Palla [et al.] // *Handbook of Large-scale Random Networks*. Springer. 2009. Ch. 9. P. 369–408.

16. Fortunato S., Barthelemy M. Resolution limit in community detection // Proceedings of the National Academy of Sciences. 2006. Vol. 104. No. 1. P. 36–41.
17. Орлов А.О., Чеповский А.А. Особенности алгоритма Блонделя при выявлении сообществ в графе социальной сети / Труды Международной научной конференции Московского физико-технического института (государственного университета) и Института физико-технической информатики (SCVRT1516). М., Протвино: Институт физико-технической информатики. 2016. С. 124–129.
18. Clauset A., Newman M., Moore C. Finding community structure in very large networks // Physical Review. 2004. E 70. 066111.
19. Gregory S. An algorithm to find overlapping community structure in networks // Proceedings of Knowledge Discovery in Databases: PKDD 2007: 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland, 17–21 September 2007. Berlin, Heidelberg: Springer, 2007, Vol. 4702. P. 91–102.
20. Collins L.M., Dent C.W. Omega: A general formulation of the Rand index of cluster recovery suitable for non-disjoint solutions // Multivariate Behavioral Research. 1988. Vol. 23. No. 2. P. 231–242.
21. Gregory S. Fuzzy overlapping communities in networks // Journal of Statistical Mechanics: Theory and Experiment. 2011. Vol. 2011. No. 02. P. 1–18.

Combined method to detect communities in graphs of interacting objects¹

Sofiya Y. Lobanova

BSc Program Student

Tikhonov Moscow Institute of Electronics and Mathematics

National Research University Higher School of Economics

Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation

E-mail: s.lobanova@usabilitylab.net

Alexander A. Chepovskiy

Associate Professor, School of Applied Mathematics

Tikhonov Moscow Institute of Electronics and Mathematics

National Research University Higher School of Economics

Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation

E-mail: aachepovskiy@hse.ru

Abstract

In this paper, we propose and implement a method for detecting intersecting and nested communities in graphs of interacting objects of different natures. For this, two classical algorithms are taken: a hierarchical agglomerate and one based on the search for k -cliques. The combined algorithm presented is based on their consistent application. In addition, parametric options are developed that are responsible for actions with communities whose sizes are smaller than the given k , and also with single vertices. Varying these parameters allows us to take into account differences in the topology of the original graph and thus to correct the algorithm.

The testing was carried out on real data, including on a group of graphs of a social network, and the qualitative content of the resulting partition was investigated. To assess the differences between the integrated method and the classical algorithms of community detections, a common measure of similarity was used. As a result, it is clearly shown that the resulting partitions are significantly different. We found that for the approach proposed in the article the index of the numerical characteristic of the partitioning accuracy, modularity, can be lower than the corresponding value for other approaches. At the same time, the result of an integrated method is often more informative due to intersections and nested community structure.

A visualization of the partition obtained for one of the examples by an integrated method at the first and last steps is presented. Along with the successfully found set of parameters of the integrated method for small communities and cut off vertices in the case of social networks, some shortcomings of the proposed model are noted. Proposals are made to develop this approach by using a set of parametric algorithms.

¹ This work was supported by the Russian Foundation for Basic Research (projects No. 16-29-09546 and No. 16-07-00641)

Key words: graph, graph analysis, community detection, graph structure, social network analysis, big data.

Citation: Lobanova S.Y., Chepovskiy A.A. (2017) Combined method to detect communities in graphs of interacting objects. *Business Informatics*, no. 4 (42), pp. 64–73. DOI: 10.17323/1998-0663.2017.4.64.73.

References

1. Guo Y., Xu K. (2007) An improved algorithm for finding community structure in networks with an application to IPv6 backbone network. *Frontiers of Computer Science in China*, vol. 1, pp. 459–467.
2. Wu F., Chen L., Wang J., Alhaji R. (2014) Biomolecular networks and human diseases. *BioMed Research International*, vol. 2014, pp. 1–2.
3. White H.C., Boorman S.A., Breiger R.L. (1976) Social structure from multiple networks. I. Blockmodels of roles and positions. *American Journal of Sociology*, vol. 81, no. 4, pp. 730–780.
4. Broder A.Z., Kumar S.R., Maghoul F., Raghavan P., Rajagopalan S., Stata R., Tomkins A., Wiener J. (2000) Graph structure in the Web. *Computer Networks*, vol. 33, no. 1–6, pp. 309–320.
5. Ahuja M.S., Singh J., Neha (2016) Practical applications of community detection. *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 6, no. 4, pp. 412–415.
6. Pennacchioli D., Coscia M., Pedreschi D. (2014) Overlap versus partition: Marketing classification and customer profiling in complex networks of products. *Proceedings of 2014 IEEE 30th International Conference on Data Engineering Workshops (ICDEW 2014). Chicago, 31 March – 4 April 2014*, pp. 103–110.
7. Newman M., Girvan M. (2004) Finding and evaluating community structure in networks. *Physical Review E*, vol. 69, no. 2, pp. 1–15.
8. Girvan M., Newman M. (2002) Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826.
9. Newman M. (2010) *Networks: An introduction*. Oxford: Oxford University Press.
10. Newman M. (2006) Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, pp. 8577–8582.
11. Blondel V.D., Guillaume J.L., Lambiotte R., Lefebvre E. (2008) Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, no. 10, P10008.
12. Fortunato S. (2010) Community detection in graphs. *Physics Reports*, vol. 486, no. 3–5, pp. 75–174.
13. Shen H.W., Cheng X.Q., Cai K., Hu M.B. (2009) Detect overlapping and hierarchical community structure in networks. *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 8, pp. 1706–1712.
14. Palla G., Derényi I., Farkas I., Vicsek T. (2005) Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, vol. 435, no. 7043, pp. 814–818.
15. Palla G., Ábel D., Farkas I.J., Pollner P., Derényi I., Vicsek T. (2009) *k*-clique percolation and clustering. *Handbook of Large-scale Random Networks*, Springer, ch. 9, pp. 369–408.
16. Fortunato S., Barthelemy M. (2006) Resolution limit in community detection // *Proceedings of the National Academy of Sciences*, vol. 104, no. 1, pp. 36–41.
17. Orlov A.O., Chepovskiy A.A. (2016) Osobennosti algoritma Blondelya pri vyyavlenii soobshchestv v grafe sotsial'noy seti [Features of Blondel algorithm to reveal communities in social networks]. *Proceedings of International Scientific Conference of Moscow Institute of Physics and Techniques and Institute of Physical and Technical Information (SCVRT1516)*. Moscow, Protvino: Institute of Physical and Technical Information, pp. 124–129 (in Russian).
18. Clauset A., Newman M., Moore C. (2004) Finding community structure in very large networks. *Physical Review E*, 70. 066111.
19. Gregory S. (2007) An algorithm to find overlapping community structure in networks. *Proceedings of Knowledge Discovery in Databases: PKDD 2007: 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland, 17–21 September 2007. Berlin, Heidelberg: Springer, 2007*, Vol. 4702. P. 91–102.
20. Collins L.M., Dent C.W. (1988) Omega: A general formulation of the Rand index of cluster recovery suitable for non-disjoint solutions. *Multivariate Behavioral Research*, vol. 23, no. 2, pp. 231–242.
21. Gregory S. (2011) Fuzzy overlapping communities in networks. *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2011, no. 02, pp. 1–18.

Моделирование и оптимизация бизнес-процессов и процессных систем в условиях неопределенности

А.Г. Мадера

*доктор технических наук, профессор департамента математики факультета экономики
Национальный исследовательский университет «Высшая школа экономики»*

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

E-mail: amadera@hse.ru

Аннотация

Данная работа посвящена математическому моделированию и оптимизации бизнес-процессов и процессных систем в условиях неопределенности. В настоящее время моделирование бизнес-процессов носит в основном дескриптивный описательный характер, что не позволяет осуществлять количественное моделирование и оптимизацию при проектировании процессов и процессных систем. Кроме того, существующие методы принятия решений в бизнес-процессах исходят из допущения о детерминированности определяющих бизнес-процессы факторов. Несмотря на неопределенность условий развития реальных процессов, обусловленных неопределенностью будущих цен на ресурсы, конъюнктуры рынка, экономики, финансов, и пр., факторы неопределенного будущего либо не принимаются во внимание, либо полагаются теми же, что наблюдаются и в настоящее время.

В статье разработана оптимизационная интервально стохастическая математическая модель, позволяющая в количественном виде моделировать бизнес-процессы и процессные системы, в которых они протекают, в условиях неопределенности будущих состояний экономики, финансов, конъюнктуры рынка, цен на ресурсы, а также актуализации шансов и рисков, имеющих место при осуществлении производственных, обеспечивающих и сервисных процессов. Критерием оптимальности модели является максимизация минимального отклонения прогнозируемых шансов и рисков, что позволяет принимать наилучшее решение при наступлении в будущем наиболее неблагоприятных для бизнес-процесса условий. Принятый в математической модели критерий оптимальности учитывает не только неопределенность будущих состояний экономики, финансов и рыночной конъюнктуры, но и психологию принятия решений и вынесения субъективных суждений и оценок. Приведены концепция и метод оценивания индуктивных (логических, субъективных) вероятностей наступления неопределенных прогнозируемых факторов бизнес-процесса.

Разработанные в статье модели и методы позволяют осуществлять математическое моделирование и оптимизацию бизнес-процессов при разнообразных видах деятельности без ограничений на сложность структурной модели бизнес-процесса, а также качественный и количественный состав звеньев в процессных системах. На их основе могут быть разработаны программные комплексы количественного проектирования бизнес-процессов и процессных систем в условиях неопределенности.

Ключевые слова: бизнес-процесс, процессная система, условия неопределенности, математическая модель, оптимизация, интервально стохастическая модель, вероятность, критерий оптимизации, шансы, риски.

Цитирование: Мадера А.Г. Моделирование и оптимизация бизнес-процессов и процессных систем в условиях неопределенности // Бизнес-информатика. 2017. № 4 (42). С. 74–82.
DOI: 10.17323/1998-0663.2017.4.74.82.

Введение

Количественное проектирование бизнес-процессов и процессных систем, в которых они совершаются, требует наличия количественных методов моделирования и оптимизации. Между тем моделирование процессов и процессных систем, к какой бы сфере деятельности они ни относились (производству, сервису, обслуживанию, обеспечению, снабжению, логистике, цепям поставок), на сегодняшний день носит в основном качественный дескриптивный описательный характер и осуществляется в виде вербального, текстового, табличного, графического и других описаний и нотаций потоков работ, ресурсов и информации. При этом в существующей литературе по бизнес-процессам под их моделированием понимается так называемая регламентация, документирование и сопровождающий документооборот, а под оптимизацией — реинжиниринг, заключающийся в выполнении ряда мероприятий по согласованию и частичному улучшению процессов [1, 2]. Однако выполнение каких бы то ни было организационных предписаний и мероприятий оптимизацией, на самом деле, не является, поскольку основывается не на математическом анализе достижения экстремального значения некоторого количественного критерия оптимальности, а на качественном, при котором обеспечить выбор оптимального решения в принципе невозможно.

Несмотря на актуальность разработки математических методов моделирования и оптимизации бизнес-процессов и процессных систем, их проектирование продолжает оставаться, по преимуществу, качественным. Вышесказанное подтверждается как явно недостаточным количеством работ по математическим методам моделирования и оптимизации процессов (на что указывается, например, в работе [3]), так и отмечаемым в работах [4, 5] фактическим отсутствием сведений о конкретных примерах, относящихся к реальным компаниям, которые на практике осуществляют количественное проектирование, моделирование и оптимизацию своих бизнес-процессов. Отметим, что работы [3, 6–14], хотя и содержат математические модели бизнес-процессов, применимы лишь *ad hoc* — к частным конкретным случаям. При этом практически во всех работах бизнес-процессы рассматриваются исключительно в детерминированных условиях, когда все факторы, их определяющие, полностью и однозначно известны.

Между тем, реальные бизнес-процессы и процессные системы на практике функционируют в условиях неопределенности всех определяющих их факторов: будущих состояний экономики, финансов, конъюнктуры рынка, будущих цен на ресурсы и энергоносители, объемов инвестиций, спроса на новый продукт, шансов, рисков, а также будущей финансовой устойчивости организационных единиц, входящих в процессную систему. Неопределенность как настоящего, так и будущего, является неотъемлемым атрибутом реальной действительности, поэтому для адекватного описания реальности необходимо, чтобы условия неопределенности включались в создаваемые методы моделирования и оптимизации бизнес-процессов.

В настоящей статье развивается оптимизационная математическая модель бизнес-процессов и процессных систем в условиях неопределенности, с использованием интервально стохастических факторов, подчиняющихся равномерному закону распределения вероятностей. Разработанные в статье модели и методы моделирования и оптимизации могут быть положены в основу количественного проектирования процессов и процессных систем и разработки соответствующего программного комплекса, для применения на практике в сферах производства, обеспечения, снабжения, сервиса, оказания услуг, логистики, цепях поставок, и др.

1. Структурные и математические модели процессных звеньев

Структурная модель процесса и процессной системы представляет собой совокупность взаимодействующих между собой организационных единиц, моделируемых процессными звеньями, и связей между ними. В каждом процессном звене осуществляются различные виды деятельности, подчиненные достижению единой общей цели, направленной на производство конечного выходного продукта, имеющего ценность для потребителя. На входы процессных звеньев поступают потоки ресурсов (факторов производства), которые преобразуются в продукты (конечный или промежуточный) на выходе звеньев.

Следуя работе [2], выделяют следующие виды бизнес-процессов и соответствующих им звеньев (рисунк 1):

♦ производственные, или основные, процессы и звенья (рисунк 1а) — процессы, добавляющие как ценность, так и стоимость к конечному или проме-

жуточному продукту, в количестве, равном затратам на производство продукта. В производственном звене происходит преобразование поступающих на его вход материальных факторов производства, энергоносителей и живого труда, в продукт на выходе (промежуточный продукт, полуфабрикат, незавершенное производство, окончательный продукт всего процесса);

♦ обеспечивающие процессы и звенья (рисунк 1b) – процессы, в которых продукт не производится, качественно и количественно не изменяется, ценность не добавляется, в отличие от добавленной стоимости, обусловленной затратами энергоносителей и живого труда по обеспечению производства. В обеспечивающем звене объемы факторов производства и продуктов на входе и выходе звена равны между собой. К обеспечивающим видам работ относятся погрузка, разгрузка, доставка факторов производства и продуктов производства, хранение, оформление документации и др.;

♦ сервисные процессы и звенья, или услуги (рисунк 1c) – процессы, добавляющие как стоимость, так и ценность для потребителя, за которую последний готов платить. В сервисном звене, как и в производственном, добавление стоимости к продукту обуславливается затратами факторов производства, энергоносителей и живого труда при выполнении сервисных работ, но в то же время, как и в обеспечивающем звене, объемы продукта на входе и выходе сервисного звена равны между собой. К сервисным работам могут быть отнесены все работы над конечным продуктом, выполняемые по заказам потребителя, за которые он готов платить: заказная упаковка, сортировка, маркировка, придание товарного вида, доставка непосредственно потребителю, и др.

Входные потоки в каждом звене включают в себя как экзогенные потоки ресурсов, закупаемых во внешней среде (множество X), так и эндогенные

потоки продуктов, произведенных в том же процессе в предыдущих звеньях (множество Y), и поступающих для производства в звенья того же процесса. Нумерация экзогенных и эндогенных потоков в множествах X и Y является независимой и сквозной по всей структуре процессной системы.

Рассмотрим структурную модель процессной системы конкретного бизнес-процесса (рисунк 2). Потоки экзогенных и эндогенных ресурсов обозначаются кружками, с указанными внутри них объемами ресурсов (x – экзогенных, y – эндогенных); поступающие ресурсы и/или продукты – стрелками; производственные звенья – прямоугольниками (рисунк 1a), обеспечивающие – ромбами (рисунк 1b), сервисные – треугольниками (рисунк 1c). Структурная модель на рисунке 2 содержит два обеспечивающих звена (1 и 6), четыре производственных звена (2, 3, 4, 5 и 7) и одно сервисное звено 8. Экзогенные ресурсы в объемах x_1, x_2, x_3, x_4, x_5 закупаются вне процесса, эндогенные продукты в объемах y_1, y_2, y_3, y_4 производятся в процессных звеньях 2, 3, 4 и 5. Конечный продукт всего бизнес-процесса в объеме y_5 с выхода производственного звена 7 поступает на вход сервисного звена 8, над которым совершаются сервисные работы, причем $y_6 = y_5$.

Работы по созданию конечного выходного продукта объемом y_5 осуществляются следующим образом. Ресурс в объеме x_1 и часть ресурса x_2 в объеме x'_2 ($x_2 = x'_2 + x''_2$) поступают в складское звено 1, а из него – на входы производственных звеньев 3 и 4. Часть x''_2 ресурса x_2 , а также ресурс объемом x_3 подаются непосредственно в производственное звено 2, где преобразуются в продукт объемом y_1 , одна часть которого в объеме y'_1 ($y_1 = y'_1 + y''_1$) подается в производственное звено 4, а другая, в объеме y''_1 – в производственное звено 5. С выхода производственных звеньев 3 и 4 продукты в объемах y_1 и y_3 , совместно с ресурсом объемом x_5 , поступают

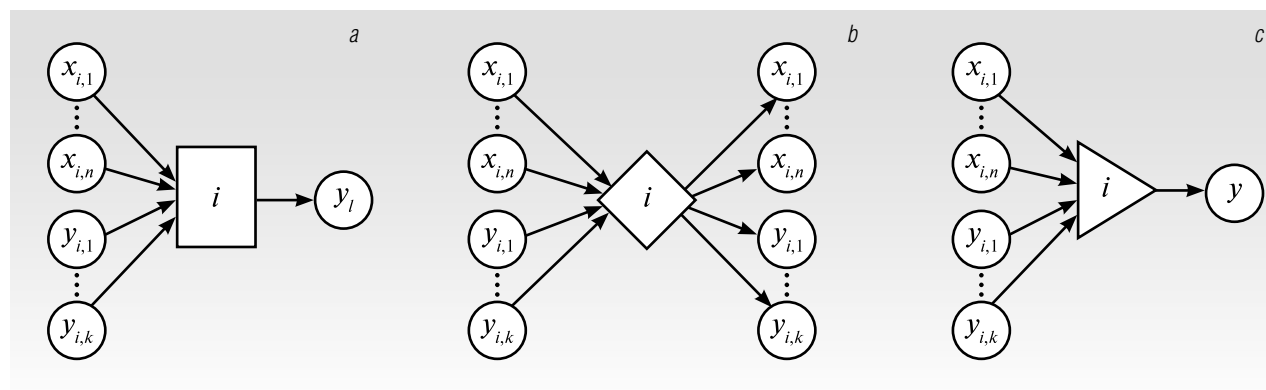


Рис. 1. Структурные модели производственного (a), обеспечивающего (b) и сервисного (c) звеньев

на преобразование в производственное звено 7. Продукт объемом y_4 с выхода производственного звена 5 и ресурс объемом x_4 , поступают на склад, а после него — на вход производственного звена 7, где совместно с продуктами y_2 , y_3 и ресурсом x_5 , перерабатываются в конечный выходной продукт бизнес-процесса объемом y_5 . Последний подвергается в сервисном звене 8 сервисным операциям, приобретая свойства, удовлетворяющие запросы заказчика и оплаченные им.

Математическая модель производственного звена i (рисунки 1а), на вход которого подаются упорядоченные множества экзогенных X_i и эндогенных Y_i потоков ресурсов $\{x_{i,1}, x_{i,2}, \dots, x_{i,n}\} \in X_i \subset X$ и $\{y_{i,1}, y_{i,2}, \dots, y_{i,k}\} \in Y_i \subset Y$, представляет собой многофакторную производственную функцию [15, 16], в натуральном выражении имеющую вид

$$y_i = f_i(x_{i,1}, x_{i,2}, \dots, x_{i,n}; y_{i,1}, y_{i,2}, \dots, y_{i,k}),$$

где $y_i \in Y$ — объем продукта, произведенного в звене i .

В отличие от производственных звеньев, в которых факторы производства на входе преобразуются в качественно новый продукт на выходе, в обеспечивающих звеньях (рисунки 1б) совершаются работы по обеспечению производственного процесса, а в сервисных (рисунки 1с) — по сервисной обработке уже произведенного конечного продукта. При этом как в обеспечивающих, так и в сервисных звеньях при совершении различных видов деятельности факторы производства не затрачиваются, в отличие от энергоносителей, живого труда и средств труда, которые затрачиваются на выполнение соответствующих работ над готовым продуктом, но без изменения его количества. В то время как обеспечивающие процессы не добавляют ценности к производимым продуктам, сервисные — напротив, добавляют ценность к готовому продукту,

как путем придания ему требуемых заказчиком свойств, так и оказания оплаченных клиентами услуг. Таким образом, математические модели обеспечивающих звеньев представляют собой тождественные производственные функции преобразования факторов производства самих в себя

$$f_i(x_{i,1}, x_{i,2}, \dots, x_{i,n}; y_{i,1}, y_{i,2}, \dots, y_{i,k}) = x_{ij}, j = 1, 2, \dots, n,$$

$$f_i(x_{i,1}, x_{i,2}, \dots, x_{i,n}; y_{i,1}, y_{i,2}, \dots, y_{i,k}) = y_{ij}, j = 1, 2, \dots, k,$$

а сервисных звеньев производственные функции, преобразующие факторы производства и готовый продукт объемом y в продукт того же объема $f(x_{i,1}, x_{i,2}, \dots, x_{i,n}; y) = y$ [15, 16].

2. Математическая модель бизнес-процесса и процессной системы в условиях неопределенности

Математические модели, адекватно описывающие реальные бизнес-процессы, должны учитывать неопределенность будущих состояний экономики, финансов, конъюнктуры рынка, цен на факторы производства и энергоносители, объемов инвестиций, шансы, риски, и прочие факторы неопределенности (см. раздел 1) и, кроме того, — психологию принятия решений и вынесения субъектом суждений и оценок [17–21]. Неопределенность моделируется интервально стохастическими факторами, равномерно распределенными внутри интервалов со значениями границ, оцениваемыми индуктивной (логической, субъективной) вероятностью [16, 20–25]. Для формулирования оптимизационной модели необходимо определить количественный критерий оптимизации и ограничения области допустимых решений D [26].

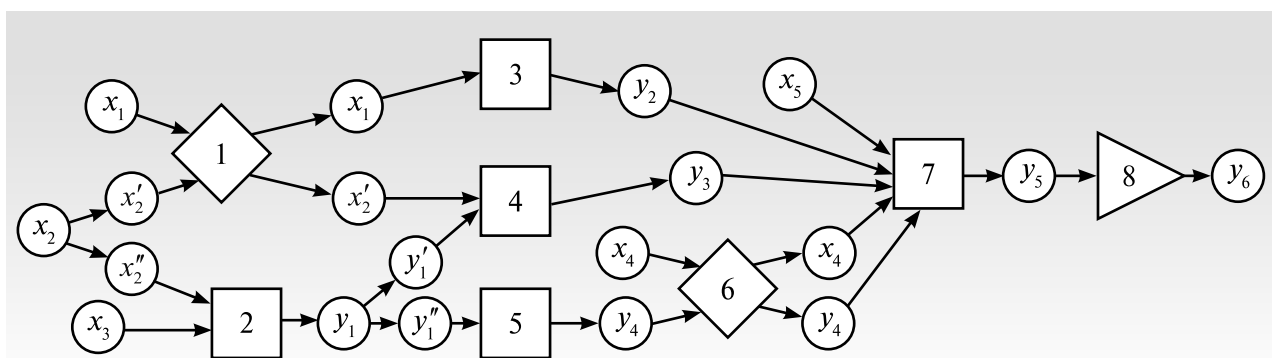


Рис. 2. Структурная модель бизнес-процесса и процессной системы
(1, 6 — обеспечивающие звенья, 2, 3, 4, 5, 7 — производственные звенья, 8 — сервисное звено)

Критерий оптимизации математической модели. Неопределенные события, актуализирующиеся в будущем, сводятся, по существу, к двум категориям: событиям, наступление которых благоприятно для бизнес-процесса (шансы) — высокие прибыли, достижение цели, получение запланированных результатов, и событиям, которые будут неблагоприятными для бизнес-процесса (риски) — убытки, недостаточная прибыль, провалы, банкротства и др. Как показано в работах [16, 24], для принятия наилучших решений критерий оптимизации бизнес-процесса должен максимизировать шансы и минимизировать риски. Этим условиям удовлетворяет комплексный критерий «шансы — риски» $Ch\&R$ [16, 24], определяемый как

$$Ch\&R = \beta_{Ch} Ch - \beta_R |R|,$$

где Ch и R — полные прогнозируемые шансы и риски, релевантные рассматриваемому процессу;

$\beta_{Ch}, \beta_R \in [0, 1]$ — коэффициенты относительной важности шансов и рисков.

Величины полных шансов $Ch \in M$ и рисков $R \in M$ (множество M содержит L шансов и K рисков) определяются [16, 24] как суммы произведений количественных мер шансов (доходы, прибыль, выручка) $M_{Ch,k}, (k=1, 2, \dots, L)$ и рисков (потери, убытки, ущербы) $M_{R,k}, (k=1, 2, \dots, K)$ на меры их возможной актуализации, или вероятности $P_{Ch,k}$ и $P_{R,k}$, то есть

$$Ch = \sum_{k=1}^L M_{Ch,k} P_{Ch,k}, \quad R = \sum_{k=1}^K M_{R,k} P_{R,k}.$$

При этом количественные меры каждого k -го шанса и риска являются интервально стохастическими и изменяющимися в интервалах $M_{Ch,k} \in [\underline{M}_{Ch,k}, \bar{M}_{Ch,k}]$, $M_{R,k} \in [\underline{M}_{R,k}, \bar{M}_{R,k}]$, где $\underline{M}_{Ch,k}, \bar{M}_{Ch,k}$ и $\underline{M}_{R,k}, \bar{M}_{R,k}$ — нижние и верхние значения границ интервалов количественных мер шансов и рисков, вычисляемые с использованием правил интервальной арифметики [27].

Для получения наилучшего, релевантного бизнес-процессу решения, обеспечивающего максимизацию шансов при одновременной минимизации рисков, критерий оптимальности «шансы — риски» $Ch\&R$ необходимо максимизировать. Это равносильно определению такого решения в условиях неопределенности, при котором гарантируется достижение максимальной разности между нижней границей шансов $\underline{Ch}$ и верхней границей рисков $\bar{R}$ [16, 24] для любого $\omega \in \Omega$ (ω — элементарные события, Ω — пространство элементарных со-

бытий). Получим следующий критерий оптимизации бизнес-процессов:

$$\max \{Ch\&R\} = \max \{ \beta_{Ch} \underline{Ch} - \beta_R \bar{R} \}, \text{ где} \quad (1)$$

$$\underline{Ch} = \sum_{k=1}^L \underline{M}_{Ch,k} P_{Ch,k}, \quad \bar{R} = \sum_{k=1}^K \bar{M}_{R,k} P_{R,k}.$$

Ограничения оптимизационной математической модели. Ограничения математической модели являются бюджетными и определяют условия на величину полных затрат по осуществлению процессной деятельности, величина которых не должна превышать вложенные в бизнес-процесс финансовые средства.

Обозначим объемы закупаемых экзогенных факторов производства как $\{x\} = \{x_1, x_2, \dots, x_m\} \in D$, $x_i \geq 0$, а их цены как $c_{Ch,i}$ и $c_{R,i}$, ($i = 1, 2, \dots, m$). Учитывая, что цены на факторы производства являются неопределенными и интервально стохастическими, изменяющимися в интервалах $c_{Ch,i}(\omega) \in [\underline{c}_{Ch,i}, \bar{c}_{Ch,i}]$ и $c_{R,i}(\omega) \in [\underline{c}_{R,i}, \bar{c}_{R,i}]$, $\bar{c}_{Ch,i} \leq \underline{c}_{R,i}$, с нижними и верхними границами $\underline{c}_{Ch,i}, \bar{c}_{Ch,i}$ и $\underline{c}_{R,i}, \bar{c}_{R,i}$, получим выражения для полных затрат по закупке факторов производства для каждого $\omega \in \Omega$:

$$C_{Ch}(\omega) = \sum_{i=1}^m c_{Ch,i}(\omega) x_i, \quad C_R(\omega) = \sum_{i=1}^m c_{R,i}(\omega) x_i.$$

В силу интервально стохастической неопределенности как затрат, так и объемов финансирования бизнес-процесса, бюджетные ограничения должны пониматься в вероятностном смысле, а именно — как вероятности выполнения бюджетных ограничений-неравенств, которые не должны быть ниже некоторых пороговых значений p_{Ch} и p_R , задаваемых владельцем бизнес-процесса [28] и бизнес-аналитиками. Таким образом, ограничения модели для шансов и рисков будут записаны в виде:

$$P\{C_{Ch}(\omega) \leq I_{Ch}(\omega)\} \geq p_{Ch}, \quad P\{C_R(\omega) \leq I_R(\omega)\} \geq p_R \quad (2)$$

где $I_{Ch}(\omega) \in [\underline{I}_{Ch}, \bar{I}_{Ch}]$ и $I_R(\omega) \in [\underline{I}_R, \bar{I}_R]$, $\bar{I}_{Ch} \leq \underline{I}_R$ — интервально стохастические объемы финансирования бизнес-процесса;

$P\{\cdot\}$ — индуктивные (субъективные, логические) вероятности.

Интервально стохастическая оптимизационная математическая модель. Интервально стохастическая оптимизационная математическая модель бизнес-процессов при неопределенности определяющих факторов, формулируется следующим образом: найти объемы факторов производства $\{x\} = \{x_1, x_2, \dots, x_m\} \in D$, $x_i \geq 0$, $i = 1, 2, \dots, m$, доставляющие

максимальное значение критерию «шансы – риски» (1) и удовлетворяющие вероятностным ограничениям для шансов и рисков (2), то есть:

$$\max_x \{Ch \& R\} = \max_x \left\{ \beta_{Ch} \underline{Ch} - \beta_R \bar{R} \right\},$$

$$P\{C_{Ch}(\omega) \leq I_{Ch}(\omega)\} \geq p_{Ch}, \quad P\{C_R(\omega) \leq I_R(\omega)\} \geq p_R.$$

Для определения оптимального решения данной оптимизационной модели в условиях интервально стохастической неопределенности необходимо преобразовать вероятностные ограничения (2) в соответствующие им детерминированные эквиваленты [29].

Эквивалентная детерминированная оптимизационная математическая модель. Заметим, что интервально стохастические цены факторов производства $c_{Ch,i}(\omega)$, $c_{R,i}(\omega)$ и объемы финансирования $I_{Ch}(\omega)$, $I_R(\omega)$, $\omega \in \Omega$ являются интервальными и подчиняются равномерным законам распределения вероятностей, а их суммы, равные полным затратам, $c_{Ch}(\omega)$ и $c_R(\omega)$, согласно центральной предельной теореме Ляпунова [30], – нормальному распределению с математическими ожиданиями

$$m(C_{Ch}) = \sum_{i=1}^m m(c_{Ch,i})x_i, \quad m(C_R) = \sum_{i=1}^m m(c_{R,i})x_i$$

и дисперсиями

$$\sigma^2(C_{Ch}) = \sum_{i=1}^m \sigma^2(c_{Ch,i})x_i^2, \quad \sigma^2(C_R) = \sum_{i=1}^m \sigma^2(c_{R,i})x_i^2, \text{ где}$$

$$m(c_{Ch,i}) = (\underline{c}_{Ch,i} + \bar{c}_{Ch,i})/2, \quad m(c_{R,i}) = (\underline{c}_{R,i} + \bar{c}_{R,i})/2,$$

$$\sigma^2(c_{Ch,i}) = (\underline{c}_{Ch,i} - \bar{c}_{Ch,i})^2/12, \quad \sigma^2(c_{R,i}) = (\underline{c}_{R,i} - \bar{c}_{R,i})^2/12.$$

Приводя, согласно теореме [29], вероятностные неравенства в ограничениях (2) к детерминированному виду, получим окончательную эквивалентную детерминированную оптимизационную математическую модель:

$$\max_{\{x\}} \{Ch \& R\} = \max_{\{x\}} \left\{ \beta_{Ch} \underline{Ch} - \beta_R \bar{R} \right\},$$

$$m(C_{Ch}) + k_{Ch} \sqrt{\sigma^2(C_{Ch}) + \sigma^2(I_{Ch})} \leq m(I_{Ch}),$$

$$m(C_R) + k_R \sqrt{\sigma^2(C_R) + \sigma^2(I_R)} \leq m(I_R),$$

$$x_i \geq 0, \quad i = 1, 2, \dots, m,$$

где k_{Ch} , k_R – квантили стандартных нормальных распределений;

$m(I_{Ch})$, $m(I_R)$ и $\sigma^2(I_{Ch})$, $\sigma^2(I_R)$ – математические ожидания и дисперсии интервально стохастических объемов финансирования, равные

$$m(I_{Ch}) = (\underline{I}_{Ch} + \bar{I}_{Ch})/2, \quad m(I_R) = (\underline{I}_R + \bar{I}_R)/2,$$

$$\sigma^2(I_{Ch}) = (\underline{I}_{Ch} - \bar{I}_{Ch})^2/12, \quad \sigma^2(I_R) = (\underline{I}_R - \bar{I}_R)^2/12.$$

Полученная детерминированная оптимизационная математическая модель относится к классу задач нелинейного программирования общего вида. Ее оптимальное решение может быть получено с помощью существующих пакетов компьютерных программ, например, NEOS, AMPL, MATLAB, MathOptimzier, GALAHAD, MS Excel. По найденному оптимальному решению $\{x^*\} = \{x_1^*, x_2^*, \dots, x_m^*\}$ определяются оптимальные значения всех факторов бизнес-процесса и процессной системы: объем и стоимость конечного продукта, объемы и стоимости промежуточных продуктов на выходах всех звеньев, доходы и прибыль при различных прогнозируемых условиях в будущем и т.д.

3. Метод оценивания индуктивных вероятностей прогнозируемых событий

Вероятности наступления будущих событий (состояний экономики, финансов, рисков, шансов и др.) необходимо оценивать согласно концепции индуктивной вероятности, поскольку прогнозируемые события являются уникальными и единичными, к которым понятия классической или статистической вероятностей неприменимы. Это обусловлено тем, что единичные уникальные события не являются случайными объектами, рассматриваемыми в классической теории вероятностей, поскольку они не удовлетворяют условиям массовости, однородности, неограниченной воспроизводимости в идентичных условиях, и устойчивости частот [20–24].

Концепция и метод оценивания индуктивных вероятностей $P(A)$ наступления событий $\{A\} = \{A_1, A_2, \dots, A_n\}$ разработаны в работе [21] и позволяют существенно увеличить адекватность прогнозирования. Метод основан как на статистических данных по прогнозированию релевантных событий за прошлые периоды, так и на новой имеющейся информации о наблюдаемых в настоящий момент времени тенденциях. Данные за прошлые периоды характеризуют погрешности, присущие субъекту при прогнозировании событий и оценивании их вероятностей, а данные, получаемые из информации о развитии событий и наблюдаемых в настоящем тенденциях, позволяют провести их коррекцию. По этим двум видам данных формируются две матрицы: матрица погрешностей прогнозирования

субъекта $\mathcal{L}$ и матрица уточненных прогнозов $\mathcal{M}$, сделанных на основании современной информации. Как показано в работе [21], искомый вектор вероятностей наступления прогнозируемых событий $P(A)$ является собственным вектором полной матрицы погрешностей прогнозирования $\mathcal{K} = \mathcal{M} \cdot \mathcal{L}$ отвечающим ее единичному собственному значению, а именно, $P(A) = \mathcal{K} \cdot P(A)$.

Заключение

Предложенная в работе оптимизационная интервально стохастическая математическая модель позволяет моделировать оптимальные бизнес-процессы и оптимальные структуры и параметры процессных систем в условиях неопределенности будущих состояний экономики, финансов, конъюнктуры рынка, спроса на новый продукт, цен на факторы производства, цены продажи конечного

продукта, объемов инвестиций, наступления в будущем шансов и рисков, и др. Структурная модель процессной системы может включать производственные, обеспечивающие и сервисные звенья, отражающие весь диапазон видов деятельности, совершаемых в различных бизнес-процессах при переработке входов в продукты на выходах. Разработанные в статье математические модели и методы позволяют осуществлять математическое моделирование и оптимизацию разнообразных бизнес-процессов и процессных систем без ограничений на сложность, качественный и количественный состав процессных звеньев и сложность структуры процессной системы. Они могут быть положены в основу разработки программных комплексов количественного проектирования бизнес-процессов и сложных процессных систем в условиях неопределенности. ■

Литература

1. Robson M., Ullah P. A practical guide to business process re-engineering. Hampshire: Gower, 1996.
2. Harrington H.J., Esseling K., van Nimwegen H. Business process improvement workbook: Documentation, analysis, design, and management of business process improvement. N.Y.: McGraw-Hill, 1997.
3. Vergidis K., Tiwari A., Majeed B. Business process analysis and optimization: Beyond reengineering // IEEE Transactions on Systems, Man, and Cybernetics. Part C (Applications and Reviews). 2008. Vol. 38. No. 1. P. 69–82.
4. Stock J.R., Lambert D.M. Strategic logistics management. Boston: McGraw-Hill, 2001.
5. Shapiro J.F. Modeling the supply chain. Duxbury, CA: Thomson Learning, 2001.
6. van der Aalst W.M.P. The application of Petri nets to workflow management // Journal of Circuits, Systems and Computers. 1998. Vol. 8. No. 1. P. 21–66.
7. Hofacker I., Vetschera R. Algorithmical approaches to business process design // Computers & Operations Research. 2001. No. 28. P. 1253–1275.
8. Koubarakis M., Plexousakis D. A formal framework for business process modelling and design // Information Systems. 2002. Vol. 27. No. 5. P. 299–319.
9. Valiris G., Glykas M. Business analysis metrics for business process redesign // Business Process Management Journal. 2004. Vol. 10. No. 4. P. 445–480.
10. Seuring S. A review of modeling approaches for sustainable supply chain management // Decision Support Systems. 2013. Vol. 54. No. 4. P. 1513–1520.
11. Wibig M. Dynamic programming and genetic algorithm for business processes optimization // International Journal of Intelligent Systems and Applications. 2013. Vol. 5. No. 1. P. 44–51.
12. Sawicki P., Sawicka H. Logistics process improvement using simulation and stochastic multiple criteria decision aiding // Procedia – Social and Behavioral Sciences. 2014. No. 111. P. 213–223.
13. Sustainable supply chain optimisation: An industrial case study / Q. Zhang [et al.] // Computers & Industrial Engineering. 2014. No. 74. P. 68–83.
14. Ahmadikatoouli A., Motameni H. Enrichment of object oriented Petri net and object Z aiming at business process optimization // International Journal of Advanced Computer Science and Applications. 2015. Vol. 6. No. 7. P. 13–19.
15. Клейнер Г.Б. Производственные функции: Теория, методы, применение. М.: Финансы и статистика, 1986.
16. Мадера А.Г. Математическое моделирование и оптимизация бизнес-процессов на основе комплексного критерия «шансы – риски» // Российский журнал менеджмента. 2015. Т. 13, № 4. С. 51–68.
17. Козелецкий Ю. Психологическая теория решений / Пер. с польск. М.: Прогресс, 1979.
18. Plous S. The psychology of judgment and decision making. N.Y.: McGraw-Hill, Wesleyan University, 1993.
19. Werth L. Psychologie für die wirtschaft. Grundlagen und anwendungen. Berlin: Spectrum Akademischer Verlag, 2004.
20. Мадера А.Г. Интервально стохастическая неопределенность оценок в многокритериальных задачах принятия решений // Искусственный интеллект и принятие решений. 2014. № 3. С. 105–115.
21. Мадера А.Г. Метод определения вероятностей прогнозируемых событий при принятии решений // Искусственный интеллект и принятие решений. 2016. № 2. С. 38–45.
22. Carnap R. Logical foundations of probability. Chicago: The University of Chicago Press. 1971.
23. de Finetti B. Bayesinism: Its unifying role for both the foundations and applications of statistics // International Statistical Review. 1974. No. 42. P. 117–130.

24. Мадера А.Г. Риски и шансы: неопределенность, прогнозирование и оценка. М.: Красанд, 2014.
25. Madera A.G. Estimating the probability of forecasted events // International Journal of Accounting and Economics Studies. 2016. Vol. 4. No. 1. P. 76–80.
26. Мадера А.Г. Моделирование и принятие решений в менеджменте. М.: ЛКИ, 2009.
27. Alefeld G., Herzberger J. Introduction to interval computations. N.Y.: Academic Press, 1983.
28. ISO/IEC 15288:2002. System engineering — System life cycle processes. 2002.
29. Liu B. Theory and practice of uncertain programming. Heidelberg: Physica-Verlag, 2002.
30. Feller W. An introduction to probability theory and its applications. N.Y.: John Wiley, 1970.

Modeling and optimization of business processes and process systems under conditions of uncertainty

Alexander G. Madera

*Professor, Department of Mathematics of the Faculty of Economics
National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow, 101000, Russian Federation
E-mail: amadera@hse.ru*

Abstract

This paper is devoted to mathematical modeling and optimization of business processes and process systems under conditions of uncertainty. At present, modeling of business processes is mainly descriptive, which does not allow quantitative modeling and optimization in the design of processes and process systems. In addition, the existing methods of decision-making in business processes are based on the assumption that the decisive factors are deterministic. Despite uncertainty of the real processes caused by the uncertainty of future costs of resources, the market environment, economy, finances, etc., the factors of an uncertain future are either not taken into account, or are believed to be the same as those observed currently.

In this paper, a stochastic interval mathematical optimization model is developed. This model allows us to simulate in a quantitative way the business processes and process systems in which they take place, taking into account the uncertainties of the future state of the economy, finances, market environment, costs of resources, as well as future realization of chances and risks related to the productive, supporting, and service processes. The criterion for optimality of the model is the maximization of the smallest deviation of the projected chances and risks, which makes it possible to make the best decision in the case that the most unfavorable conditions for the business process occur in the future. The criterion of optimality adopted in the mathematical model takes into account not only the uncertainty of the future state of the economy, finance, and market environment, but also the psychology of decision-making and the subjective nature of judgments and estimates. We present a concept and method for estimating the inductive (logical, subjective) probabilities of the occurrence of uncertain projected business process factors.

The models and methods developed in the paper make it possible to carry out mathematical modeling and optimization of business processes in a variety of activities without restrictions on the complexity of the structural model of the business process, the qualitative and quantitative composition of the connections in the process systems. On their basis, a software package for the quantitative design of business processes and process systems under conditions of uncertainty can be developed.

Key words: business process, process system, conditions of uncertainty, mathematical model, optimization, interval stochastic model, probability, optimization criterion, chances, risks.

Citation: Madera A.G. (2017) Modeling and optimization of business processes and process systems under conditions of uncertainty. *Business Informatics*, no. 4 (42), pp. 74–82. DOI: 10.17323/1998-0663.2017.4.74.82.

References

- 1 Robson M., Ullah P. (1996) *A practical guide to business process re-engineering*. Hampshire: Gower.
- 2 Harrington H.J., Esseling K., van Nimwegen H. (1997) *Business process improvement workbook: Documentation, analysis, design, and management of business process improvement*. N.Y.: McGraw-Hill.

3. Vergidis K., Tiwari A., Majeed B. (2008) Business process analysis and optimization: Beyond reengineering. *IEEE Transactions on Systems, Man, and Cybernetics. Part C (Applications and Reviews)*, vol. 38, no. 1, pp. 69–82.
4. Stock J.R., Lambert D.M. (2001) *Strategic logistics management*. Boston: McGraw-Hill.
5. Shapiro J.F. (2001) *Modeling the supply chain*. Duxbury, CA: Thomson Learning.
6. van der Aalst W.M.P. (1998) The application of Petri nets to workflow management. *Journal of Circuits, Systems and Computers*, vol. 8, no. 1, pp. 21–66.
7. Hofacker I., Vetschera R. (2001) Algorithmical approaches to business process design. *Computers & Operations Research*, no. 28, pp. 1253–1275.
8. Koubarakis M., Plexousakis D. (2002) A formal framework for business process modelling and design. *Information Systems*, vol. 27, no. 5, pp. 299–319.
9. Valiris G., Glykas M. (2004) Business analysis metrics for business process redesign. *Business Process Management Journal*, vol. 10, no. 4, pp. 445–480.
10. Seuring S. (2013) A review of modeling approaches for sustainable supply chain management. *Decision Support Systems*, vol. 54, no. 4, pp. 1513–1520.
11. Wibig M. (2013) Dynamic programming and genetic algorithm for business processes optimization. *International Journal of Intelligent Systems and Applications*, vol. 5, no. 1, pp. 44–51.
12. Sawicki P., Sawicka H. (2014) Logistics process improvement using simulation and stochastic multiple criteria decision aiding. *Procedia – Social and Behavioral Sciences*, no. 111, pp. 213–223.
13. Zhang Q., Shah N., Wassick J., Helling R., van Egerschot P. (2014) Sustainable supply chain optimisation: An industrial case study. *Computers & Industrial Engineering*, no. 74, pp. 68–83.
14. Ahmadikatoouli A., Motameni H. (2015) Enrichment of object oriented Petri net and object Z aiming at business process optimization. *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 7, pp. 13–19.
15. Kleiner G.B. (1986) *Proizvodstvennye funktsii: Teoriya, metody, primeneniye* [Production functions: Theory, methods, application]. Moscow: Finance and Statistics (in Russian).
16. Madera A.G. (2015) Matematicheskoe modelirovanie i optimizatsiya biznes-protsessov na osnove kompleksnogo kriteriya “shansy – riski” [Mathematical modeling and optimization of business processes based on an integrated criterion of chances and risks]. *Russian Management Journal*, vol. 13, no. 4, pp. 51–68 (in Russian).
17. Kozielski J. (1979) *Psikhologicheskaya teoriya resheniy* [Psychological decision theory]. Moscow: Progress (in Russian).
18. Plous S. (1993) *The psychology of judgment and decision making*. N.Y.: McGraw-Hill, Wesleyan University.
19. Werth L. (2004) *Psychologie für die wirtschaft. Grundlagen und anwendungen*. Berlin: Spectrum Akademischer Verlag.
20. Madera A.G. (2014) Interval’no stokhasticheskaya neopredelennost’ otsenok v mnogokriterial’nykh zadachakh prinyatiya resheniy [Interval stochastic uncertainty of estimates in multiple criteria decision making problems]. *Artificial Intelligence and Decision Making*, no. 3, pp. 105–115 (in Russian).
21. Madera A.G. (2016) Metod opredeleniya veroyatnostey prognozirovemykh sobyitiy pri prinyatii resheniy [Method for determination of probabilities of the projected events when making decisions]. *Artificial Intelligence and Decision Making*, no. 2, pp. 38–45 (in Russian).
22. Carnap R. (1971) *Logical foundations of probability*. Chicago: The University of Chicago Press.
23. de Finetti B. (1974) Bayesianism: Its unifying role for both the foundations and applications of statistics. *International Statistical Review*, no. 42, pp. 117–130.
24. Madera A.G. (2014) *Riski i shansy: neopredelennost’, prognozirovanie i otsenka* [Risks and chances: Uncertainty, forecasting and estimation]. Moscow: Krasand (in Russian).
25. Madera A.G. (2016) Estimating the probability of forecasted events. *International Journal of Accounting and Economics Studies*, vol. 4, no. 1, pp. 76–80.
26. Madera A.G. (2009) *Modelirovanie i prinyatie resheniy v menedzhmente* [Modeling and decision making in management]. Moscow: LKI (in Russian).
27. Alefeld G., Herzberger J. (1983) *Introduction to interval computations*. N.Y.: Academic Press.
28. ISO/IEC 15288:2002. *System engineering – System life cycle processes*.
29. Liu B. (2002) *Theory and practice of uncertain programming*. Heidelberg: Physica-Verlag.
30. Feller W. (1970) *An introduction to probability theory and its applications*. N.Y.: John Wiley.

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:
biijournal@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 Math Type or MS Equation tool.

FIGURES should be of high quality, black and white, legible and numbered consecutively with Arabic numerals. All figures (charts, diagrams, etc.) should be submitted in electronic form (photo images – in TIF, PSD 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.