

БИЗНЕС- ИНФОРМАТИКА

Научный журнал НИУ ВШЭ

СОДЕРЖАНИЕ

Д.Е. Пальчунов, А.А. Якобсон

Разработка интеллектуального помощника
для подбора товаров в процессе диалога
с пользователем 7

М.Е. Ощепков

Совершенствование процесса целевого
бюджетирования в системе управления
корпоративной результативностью 22

Г.Л. Бекларян

Агентное моделирование и оптимизация
характеристик научно-производственных
кластеров 36

П.А. Колданов, В.А. Колданов

Непараметрическая процедура сравнения
эффективности работы подразделений
сетевой организации 52

K.N. Hardanti, Sutrisno T., E. Saraswati, A. Prastiwi

Determinants of an auditor's continuance intention
with respect to use of the Audit Tools and Linked
Archives System (ATLAS): A model of extended
expectation confirmation 65

T.C. Luong, O.T. Tran

Product information recognition in the retail domain
as an MRC problem 79



Издатель:
Национальный
исследовательский университет
«Высшая школа экономики»

Подписной индекс
Объединенного каталога
«Пресса России» – E79128
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
Е.П. Зараменских

Заместитель главного редактора
Э.А. Бабкин

Компьютерная верстка
О.А. Богданович

Дизайн обложки выполнен
с использованием контента (изображения),
сгенерированного Пользователем
Хрустальной И.И.
(по поручению НИУ ВШЭ),
при помощи Сервиса Kandinsky 3.0
(fusionbrain.ai).

Администратор веб-сайта
И.И. Хрусталева

Адрес редакции:
119049, г. Москва,
ул. Шаболовка, д. 26-28
Тел./факс: +7 (495) 772-9590 *28509
<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 – 100 экз.,
англоязычная версия – 100 экз.,
онлайн-версии на русском и английском –
свободный доступ

Отпечатано в типографии НИУ ВШЭ
г. Москва, Измайловское шоссе, д. 44, стр. 2

© 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 Высшей школой бизнеса НИУ ВШЭ. Журнал выпускается ежеквартально, на русском и английском языках.

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: моделирование социальных и экономических систем, цифровая трансформация бизнеса, управление инновациями, информационные системы и цифровые технологии в бизнесе, анализ данных и системы бизнес-интеллекта, математические методы и алгоритмы бизнес-информатики, моделирование и анализ бизнес-процессов, поддержка принятия управленческих решений.

Журнал «Бизнес-информатика» включен в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней кандидата и доктора наук (Перечень ВАК).

Журнал входит в базы Scopus, Web of Science Emerging Sources Citation Index (WoS ESCI), Russian Science Citation Index на платформе Web of Science (RSCI), EBSCO.

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

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

ГЛАВНЫЙ РЕДАКТОР

Зараменских Евгений Петрович

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА

Бабкин Эдуард Александрович

Национальный исследовательский университет
«Высшая школа экономики», Нижний Новгород, Россия

ЧЛЕНЫ РЕДКОЛЛЕГИИ

Авдошин Сергей Михайлович

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Акопов Андраник Сумбатович

Центральный экономико-математический институт РАН,
Москва, Россия

Алескеров Фуад Тагиевич

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Афанасьев Александр Петрович

Институт проблем передачи информации им. А.А. Харкевича
РАН, Москва, Россия

Афанасьев Антон Александрович

Центральный экономико-математический институт РАН,
Москва, Россия

Баранов Александр Павлович

Главный научно-исследовательский вычислительный центр
Федеральной налоговой службы, Москва, Россия

Барахнин Владимир Борисович

Федеральный исследовательский центр информационных
и вычислительных технологий, Новосибирск, Россия

Беккер Йорг

Университет Мюнстера, Мюнстер, Германия

Вестнер Маркус

Технический университет прикладных наук,
Регенсбург, Германия

Гаврилова Татьяна Альбертовна

Санкт-Петербургский государственный университет,
Санкт-Петербург, Россия

Глотен Эрве

Тулонский университет, Ла-Гард, Франция

Гурвич Владимир Александрович

Ратгерский университет (Университет Нью-Джерси),
Ратгерс, США

Джейкобс Лоренц

Университет Цюриха, Цюрих, Швейцария

Дискин Иосиф Евгеньевич

Всероссийский центр изучения общественного мнения,
Москва, Россия

Зандкуль Курт

Университет Ростока, Росток, Германия

Иванников Александр Дмитриевич

Институт проблем проектирования в микроэлектронике РАН,
Москва, Россия

Исаев Дмитрий Валентинович

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Калягин Валерий Александрович

Национальный исследовательский университет
«Высшая школа экономики», Нижний Новгород, Россия

Кравченко Татьяна Константиновна

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Кузнецов Сергей Олегович

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Лугачев Михаил Иванович

Московский государственный университет
им. М.В. Ломоносова, Москва, Россия

Лин Квей-Жей

Технологический институт Нагои, Нагоя, Япония

Мальцева Светлана Валентиновна

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Мейор Питер

Комиссия ООН по науке и технологиям, Женева,
Швейцария

Миркин Борис Григорьевич

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Назаров Дмитрий Михайлович

Уральский государственный экономический университет,
Екатеринбург, Россия

Пальчунов Дмитрий Евгеньевич

Новосибирский государственный университет, Новосибирск,
Россия

Пардалос Панайот (Панос)

Университет Флориды, Гейнсвилл, США

Пастор Оскар

Политехнический университет Валенсии, Валенсия,
Испания

Посегга Йоахим

Университет Пассау, Пассау, Германия

Самуйлов Константин Евгеньевич

Российский университет дружбы народов, Москва, Россия

Стоянова Ольга Владимировна

Национальный исследовательский университет
«Высшая школа экономики», Санкт-Петербург, Россия

Триболе Жозе

Университет Лиссабона, Лиссабон, Португалия

Ульянов Михаил Васильевич

AVECO, Любляна, Словения

Ускенбаева Раиса Кабиевна

Казахский национальный исследовательский технический
университет им. К.И. Сатпаева, Алматы, Казахстан

Цуканова Ольга Анатольевна

Санкт-Петербургский национальный исследовательский
университет информационных технологий, механики
и оптики, Санкт-Петербург, Россия

Чхартишвили Александр Гедеванович

Институт проблем управления им В.А. Трапезникова РАН,
Москва, Россия

Штраус Кристина

Университет Вены, Вена, Австрия

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

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

BUSINESS INFORMATICS

HSE Scientific Journal

CONTENTS

D.E. Palchunov, A.A. Yakobson

Development of an intelligent assistant for selection
of goods in the process of dialogue with the user 7

M.E. Oshchepkov

Improving target budgeting in a corporate
performance management system 22

G.L. Beklaryan

Agent-based modeling and optimization
of the characteristics for research-and-production
clusters 36

P.A. Koldanov, V.A. Koldanov

Nonparametric procedure for comparing
the performance of divisions of a network
organization 52

K.N. Hardanti, Sutrisno T., E. Saraswati, A. Prastiwi

Determinants of an auditor's continuance intention
with respect to use of the Audit Tools and Linked
Archives System (ATLAS): A model of extended
expectation confirmation 65

T.C. Luong, O.T. Tran

Product information recognition in the retail
domain as an MRC problem 79

Vol. 18 No. 1 – 2024



Publisher:
HSE University

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 Russian Federation

Editor-in-Chief
E. Zaramenskikh

Deputy Editor-in-Chief
E. Babkin

Computer making-up
O. Bogdanovich

The cover design is made
using the content (image)
generated by the User I. Khrustaleva
(on behalf of HSE University),
using the Kandinsky 3.0 Service
(fusionbrain.ai).

Website administration
I. Khrustaleva

Address:
26-28, build. 4, Shablovka Street
Moscow 119049, Russia

Tel./fax: +7 (495) 772-9590 *28509
<http://bijournal.hse.ru>
E-mail: bijournal@hse.ru

Circulation:
English version – 100 copies,
Russian version – 100 copies,
online versions in English and Russian –
open access

Printed in HSE Printing House
44, build. 2, Izmaylovskoye Shosse,
Moscow, Russia

© HSE University

ABOUT THE JOURNAL

Business Informatics is a peer reviewed interdisciplinary academic journal published since 2007 by HSE University, Moscow, Russian Federation. The journal is administered by HSE Graduate School of Business. The journal is issued quarterly, in English and Russian.

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 following areas: modeling of social and economic systems, digital transformation of business, innovation management, information systems and technologies in business, data analysis and business intelligence systems, mathematical methods and algorithms of business informatics, business processes modeling and analysis, decision support in management.

The journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Russian Federation.

The journal is included into Scopus, Web of Science Emerging Sources Citation Index (WoS ESCI), Russian Science Citation Index on the Web of Science platform (RSCI), EBSCO.

The journal is distributed both in printed and electronic forms.

EDITORIAL BOARD

EDITOR-IN-CHIEF

Evgeny P. Zaramenskikh

HSE University, Moscow, Russia

DEPUTY EDITOR-IN-CHIEF

Eduard A. Babkin

HSE University, Nizhny Novgorod, Russia

EDITORIAL BOARD

Sergey M. Avdoshin

HSE University, Moscow, Russia

Andranik S. Akopov

Central Economics and Mathematics Institute, Russian Academy of Sciences, Moscow, Russia

Fuad T. Aleskerov

HSE University, Moscow, Russia

Alexander P. Afanasyev

Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences, Moscow, Russia

Anton A. Afanasyev

Central Economics and Mathematics Institute, Russian Academy of Sciences, Moscow, Russia

Vladimir B. Barakhnin

Federal Research Center of Information and Computational Technologies, Novosibirsk, Russia

Alexander P. Baranov

Federal Tax Service, Moscow, Russia

Jörg Becker

University of Munster, Munster, Germany

Alexander G. Chkhartishvili

V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

Tatiana A. Gavrilova

Saint-Petersburg University, St. Petersburg, Russia

Hervé Glotin

University of Toulon, La Garde, France

Vladimir A. Gurvich

Rutgers, The State University of New Jersey, Rutgers, USA

Laurence Jacobs

University of Zurich, Zurich, Switzerland

Iosif E. Diskin

Russian Public Opinion Research Center, Moscow, Russia

Dmitry V. Isaev

HSE University, Moscow, Russia

Alexander D. Ivannikov

Institute for Design Problems in Microelectronics, Russian Academy of Sciences, Moscow, Russia

Valery A. Kalyagin

HSE University, Nizhny Novgorod, Russia

Tatiana K. Kravchenko

HSE University, Moscow, Russia

Sergei O. Kuznetsov

HSE University, Moscow, Russia

Kwei-Jay Lin

Nagoya Institute of Technology, Nagoya, Japan

Mikhail I. Lugachev

Lomonosov Moscow State University, Moscow, Russia

Svetlana V. Maltseva

HSE University, Moscow, Russia

Peter Major

UN Commission on Science and Technology for Development, Geneva, Switzerland

Boris G. Mirkin

HSE University, Moscow, Russia

Dmitry M. Nazarov

Ural State University of Economics, Ekaterinburg, Russia

Dmitry E. Palchunov

Novosibirsk State University, Novosibirsk, Russia

Panagote (Panos) M. Pardalos

University of Florida, Gainesville, USA

Óscar Pastor

Polytechnic University of Valencia, Valencia, Spain

Joachim Posegga

University of Passau, Passau, Germany

Konstantin E. Samouylov

Peoples' Friendship University, Moscow, Russia

Kurt Sandkuhl

University of Rostock, Rostock, Germany

Olga Stoyanova

HSE University, St. Petersburg, Russia

Christine Strauss

University of Vienna, Vienna, Austria

José M. Tribolet

Universidade de Lisboa, Lisbon, Portugal

Olga A. Tsukanova

Saint-Petersburg National Research University of Information Technologies, Mechanics and Optics, St. Petersburg, Russia

Mikhail V. Ulyanov

AVECO, Ljubljana, Slovenia

Raissa K. Uskenbayeva

Kazakh National Technical University after K.I. Satpaev, Almaty, Kazakhstan

Markus Westner

Technical University for Applied Sciences (OTH Regensburg), Regensburg, Germany

DOI: 10.17323/2587-814X.2024.1.7.21

Разработка интеллектуального помощника для подбора товаров в процессе диалога с пользователем*

Д.Е. Пальчунов^a 

E-mail: palch@math.nsc.ru

А.А. Якобсон^b

E-mail: a.yakobson@g.nsu.ru

^a Институт математики им. С.Л. Соболева, Сибирское отделение Российской академии наук
Адрес: Россия, 630090, г. Новосибирск, пр-т Академика Коптюга, д. 4

^b Новосибирский национальный исследовательский государственный университет
Адрес: Россия, 630090, г. Новосибирск, ул. Пирогова, д. 1

Аннотация

Статья посвящена разработке методов создания интеллектуальных помощников. Интеллектуальные помощники могут применяться в колл-центрах для решения проблем клиентов, для решения задач техподдержки, для помощи людям с ограниченными возможностями, для помощи в выборе товаров и др. Рассматриваются интеллектуальные помощники, ведущие аргументативный диалог с пользователями, направленный на поиск товаров и услуг, максимально удовлетворяющих желания и потребности пользователей. Разработка интеллектуального помощника ведется на основе четырехуровневой модели предметной области и семантической модели пользователя. Разрабатываемая система автоматизирует процесс поиска и обоснования решения за счет повторного использования прецедентов – накопленных ранее знаний о предыдущих диалогах с пользователями. Это дает разрабатываемой системе преимущество перед имеющимися аналогами, которые неспособны к переиспользованию знаний о предыдущих диалогах. В статье разрабатывается прецедентный подход к созданию интеллектуальной системы, способной аргументировать свои ответы. Для этого строится граф аргументации, разрабатываются методы структурирования прецедентов, используются онтологические гомоморфизмы для преобразования имеющихся прецедентов в готовое решение. Представлено описание теоретико-модельных методов построения интеллектуальных помощников. Прецеденты товаров, пользователей и диалогов интеллектуального

* Статья опубликована при поддержке Программы НИУ ВШЭ «Университетское партнерство»

помощника с пользователями формально описываются в виде частичных моделей. Преобразование прецедентов и сходство прецедентов формализуется при помощи онтологических гомоморфизмов частичных моделей. Цель разрабатываемой диалоговой системы – не только подобрать решение по запросу пользователя, но и выяснить задачи, которые собирается решать пользователь, провести анализ его аргументации, а затем обосновать пользователю предложенное решение, показать, что именно этот товар или услуга смогут удовлетворить его потребности.

Ключевые слова: интеллектуальный помощник, аргументативный диалог, прецедент, частичная модель, онтологический гомоморфизм, онтологическая модель предметной области, семантическая модель пользователя

Цитирование: Пальчунов Д.Е., Якобсон А.А. Разработка интеллектуального помощника для подбора товаров в процессе диалога с пользователем // Бизнес-информатика. 2024. Т. 18. № 1. С. 7–21.

DOI: 10.17323/2587-814X.2024.1.7.21

Введение

Сейчас мир переживает пик популярности технологий искусственного интеллекта (ИИ). Успех ChatGPT как универсальной диалоговой системы показал потребность людей в наличии интеллектуального инструмента для решения различных задач. ChatGPT решает многие задачи: от написания текстов для социальных сетей (с большим успехом) до создания научных работ (с не очень большим успехом).

В большинстве случаев такие системы выступают в роли крайне развитых и мощных компиляторов контента – перебирают существующие данные и по кусочкам собирают из них требуемый результат. При этом сама концепция нейронных сетей (нейросетей) накладывает на них вполне конкретное ограничение: они опираются на собственную обученную модель, на заранее внесенные в обучающую выборку данные. Как итог они не могут использовать недавние результаты собственной работы для улучшения процесса поиска решения, поскольку переобучение нейросети – долгий и ресурсоемкий процесс. Также, при появлении какого-либо принципиально нового объекта в предметной области (например, нового стиля рисования, если говорить о нейросети, создающей изображения), нейросеть при всем желании не сможет самостоятельно получить такой же результат – эти новые данные не были заложены в нее при обучении.

Дополнительно осложняет ситуацию и тот факт, что нейросети являются «черным ящиком», интерпретировать процесс их работы, особенно для

больших промышленных экземпляров, практически невозможно. Одним из возможных решений проблем нейросетевого подхода является построение логической системы, основанной на семантическом анализе и структуризации прецедентов – предыдущих сеансов работы интеллектуальной системы как с данным пользователем, так и со всеми предыдущими пользователями.

1. Задачи интеллектуального помощника

Нашей целью является разработка методов ведения аргументированного диалога интеллектуального помощника с пользователем для того, чтобы помогать пользователю достигать поставленных целей, реализовывать свои намерения, удовлетворять потребности, а также решать возникающие проблемы. В данной работе мы в первую очередь рассматриваем интеллектуальных помощников, помогающих пользователям в выборе подходящих товаров и услуг.

Для достижения этой цели необходимо автоматизировано решать следующие задачи:

- ♦ выявление интересов, потребностей, желаний, целей и намерений пользователя;
- ♦ выяснение, каким образом пользователь достигает своих целей, решает свои задачи, реализует намерения и удовлетворяет свои потребности (например, намерение приобрести желаемый товар);
- ♦ выявление обоснования, объяснения, аргументации пользователя, почему, например, ему нужно именно это устройство; выяснение конкретных

задач, которые пользователь собирается решать при помощи данного устройства (например, просмотр и редактирование фотографий, уборка помещения или управление умным домом);

- ◆ подбор для пользователя товара или услуги, наиболее подходящего для решения задач пользователя и удовлетворения его потребностей;
- ◆ построение аргументации, обоснования того, что данный товар или услуга действительно является самым лучшим для пользователя (при условии выполнения ограничений по цене товара и других нефункциональных требований), либо предложение пользователю набора товаров, которые наилучшим образом подходят для решения его задач;
- ◆ объяснение, в чем состоит отличие этих товаров, их положительные и отрицательные качества (в сравнении друг с другом) с точки зрения решения задач пользователя и удовлетворения его потребностей.

Для разработки методов ведения аргументативного диалога интеллектуального помощника с пользователем мы применяем современную теорию аргументации [1–9].

Данная статья в первую очередь посвящена изложению методов и технологий подбора интеллектуальным помощником товаров, наиболее подходящих пользователю, в процессе диалога с ним. Более подробное описание методов построения обоснования и аргументации того, что данный товар является наилучшим для пользователя, будет предметом следующей статьи.

Необходимую информацию для ведения диалога с пользователем интеллектуальный помощник берет из семантической (онтологической) модели [10–12], структуру которой мы опишем ниже. Семантическая модель наполняется и пополняется за счет извлечения информации как с сайтов производителей товаров и интернет-магазинов, так и за счет анализа отзывов покупателей о приобретенных товарах.

При разработке онтологической модели мы используем ряд онтологий. Это:

- ◆ онтология предметной области в целом;
- ◆ онтология характеристик, свойств, функциональных возможностей различных товаров и устройств;
- ◆ онтология пользователей: пользовательских задач, целей и намерений; каких целей пользователи достигают и какими способами.

Исключительно важным в разрабатываемом подходе является то, что мы сохраняем и анализируем диалоги с пользователем. Это является существенным отличием данного подхода от большинства существующих решений.

Например, Алиса (виртуальный голосовой помощник, созданный компанией Яндекс, ЯндексGPT 2) при ведении диалога с пользователем «не помнит» даже предыдущую реплику или вопрос пользователя. Если пользователь говорит: «Алиса, поставь мне какую-нибудь песню группы X», то Алиса поставит песню этой группы. Но если пользователь говорит: «Алиса, мне нравятся песни группы X. Поставь мне какую-нибудь песню этой группы», то Алиса отвечает «Мне нечего ответить».

Если во время воспроизведения песни спросить: «Алиса, какое название у этой песни?», она ответит. Если попросить: «Алиса, поставь предыдущую песню», она поставит. Но если спросить: «Алиса, какое название у предыдущей песни?», то она не сможет ответить.

Разрабатываемый нами интеллектуальный помощник при подборе товаров для пользователя и при порождении аргументации может обращаться:

- ◆ ко всему текущему диалогу с этим пользователем;
- ◆ к предыдущим диалогам с данным пользователем;
- ◆ к предыдущим диалогам с другими пользователями.

Таким образом интеллектуальный помощник работает с прецедентами ранее проведенных диалогов. Они находятся на третьем уровне четырехуровневой онтологической модели, которая будет подробно описана ниже.

В настоящее время различными организациями, такими, как интернет-магазины, банки и др., используются виртуальные ассистенты, предназначенные для помощи пользователю в поиске подходящего товара или услуги. Однако, как правило, эти системы работают по заранее заданному сценарию и при возникновении новой, не предусмотренной заранее ситуации, они не могут оказать помощь клиенту и перенаправляют его на взаимодействие с консультантом – человеком.

Эта ситуация, в которой виртуальный ассистент не способен найти решение или выдать нужную рекомендацию, приводит к снижению мотивации пользователя работать с ним в дальнейшем. Кроме того, подобные системы практически всегда ведут диалог «с нуля», не запоминая пользователя и кон-

текст диалога. Если пользователь уже обращался с подобной проблемой, он должен пройти весь путь поиска решения заново.

Разрабатываемая нами система автоматизирует процесс поиска решения за счет повторного использования накопленных ранее случаев (ситуаций, прецедентов). Путем сравнения цели текущего пользователя и информации из предыдущих диалогов возможно не только найти похожее решение в прошлом, но и дополнительно аргументировать предложенное решение на основе совпадения намерений. Это дает разрабатываемой системе преимущество перед существующими аналогами, которые неспособны к накоплению прецедентов и аргументации своих решений.

В данной работе мы разрабатываем прецедентный подход к созданию интеллектуальной системы, способной аргументировать свои ответы. Для этого строится граф аргументации, разрабатываются методы структурирования прецедентов, используются онтологические гомоморфизмы для преобразования имеющихся прецедентов в готовое решение. Целью интеллектуальной диалоговой системы является помощь человеку в поиске ответа на тот или иной вопрос. Наша задача — обойти ограничение нейросетевого подхода, состоящее в невозможности учитывать результаты недавних сессий работы с пользователем. Для этого мы реализуем прецедентный подход к построению решения с использованием онтологических гомоморфизмов. На его основе происходит построение и обоснование решения.

2. Существующие подходы и решения

В настоящее время существует множество диалоговых систем, созданных под разные задачи: систем, поддерживающих диалог с помощью простых фраз, систем — голосовых помощников (Алиса, Siri и прочие), способных к более сложному общению, шуткам или исполнению простых заданий (найти информацию в интернете, включить электрический прибор) и пр. На данный момент времени вершиной развития подобных диалоговых систем являются сложные языковые универсальные модели (LLM), разработанные под решение произвольных задач по типу ChatGPT.

Отдельно стоит выделить подкласс рекомендательных систем — для них важно не только найти и выдать корректную информацию, но также и обо-

сновать, почему система вывела определенное решение. Мы рассмотрим системы разных типов с точки зрения подходов к получению решения и его обоснованию.

2.1. ChatGPT

Диалоговая система (языковая модель LLM) от OpenAI, ориентированная на решение произвольных задач [13]. В ее основе лежит сильная предобученная языковая модель InstructGPT, используемая для формализации ввода пользователя, сама же модель обучена с применением подхода Reinforcement Learning with Human Feedback (RLHF). При помощи экспертов была создана модель вознаграждения, присваивавшая оценку корректности решения основной модели, после чего было запущено автоматизированное обучение с подкреплением.

На данный момент имеется исключительно мало исследовательских статей об архитектуре ChatGPT, поскольку OpenAI не разглашала подобную информацию, кроме той, что есть в блоге компании [13].

Модель имеет ряд недостатков.

1. При ведении длинного диалога, ответы становятся нечеткими, система начинает выдавать некорректные ответы. Причина этого состоит в том, что модель не обучается на длинных диалогах, фокус смещен на более подробные и детальные ответы на небольшое количество вопросов в одном сеансе.
2. Исходная модель не использует данные из сети интернет, а ограничена данными, которые попали в нее во время обучения. Как итог, она не может использовать сведения, полученные из диалогов с пользователями (например, новый факт об окружающем мире), что повышает зависимость модели от качества обучающих данных и создает временной лаг между появлением новых знаний и внесением их в модель.
3. Модель не верифицирует данные, полученные в ходе работы, что приводит к парадоксальной ситуации, когда система очень подробно рассуждает о бессмысленных вещах, вводя пользователя в заблуждение (этот феномен метафорически называется «галлюцинации»).

Кроме того, важно отметить, что ChatGPT может хранить контекст текущего диалога, но при следующем сеансе общения с этим же пользователем эта информация использоваться не будет. В результате, мы получаем, что модель может генерировать неактуальные для пользователя ответы,

которые ранее уже были получены и использованы в диалоге с ним же, «забывать» то, что пользователь сообщал модели ранее.

В то же время, задача дообучения нейронной сети такого масштаба, в зависимости от результатов диалогов, для решения проблемы динамического обновления набора прецедентов требует значительных затрат вычислительных ресурсов и может приводить к проблеме катастрофического забывания [14], поскольку поступающие данные могут быть какими угодно. В результате, разработчики предпочитают сначала набрать новые данные, обработать их и запустить разовый, но длительный процесс обучения.

2.2. BlenderBot

Модель BlenderBot от ParlAI, ориентированная на диалоги [15]. Благодаря наличию долговременной памяти, система лучше, чем ChatGPT поддерживает длинный диалог, она способна использовать информацию, ранее полученную от пользователя, однако данные, полученные от одного пользователя, не используются в диалоге с другими пользователями.

Архитектура системы BlenderBot построена по принципу конвейера [16], система формирует ответ, последовательно используя серию модулей, каждый из которых выполняет свою задачу, затем передает свой вывод в следующий модуль. Модель существует в трех видах, в зависимости от количества параметров (3, 30, 175 миллиардов параметров).

Порядок вызова модулей [16] зависит от контекста, в котором происходит диалог. Система формирует решение в зависимости от контекста диалога, обращаясь как к собственной долговременной памяти, так и формируя интернет-запросы. В случае, если поиск в памяти и в сети не требуется, данные будут извлекаться из текущего диалога. Ответ формируется из этих данных, кроме того, на этом этапе задействуются модули, ответственные за симуляцию эмпатии и личности в диалоге. Полный список модулей и их подробное описание присутствует в техническом отчете разработчиков BlenderBot [16]. Используя множество модулей и регулируя порядок их вызова, система может, поддерживая длинный диалог, обновлять свои данные о пользователе, вести диалог на несколько разных тем, переключаясь между ними, в зависимости от контекста последнего сообщения пользователя.

Рассматривая BlenderBot в контексте нашей задачи, можно отметить успешную реализацию в модели системы долговременной памяти и системы принятия решения о необходимости поиска в базе данных предыдущих диалогов или интернете. Однако, данные о предыдущих диалогах хранятся в памяти только в виде набора фактов (например, «Пользователь 1 любит собак», «Пользователь 2 живет в стране А»), не имеют семантической связи друг с другом, привязаны к конкретному пользователю. Таким образом, теряется основная часть контекста предыдущих диалогов. Также система не может воспроизвести собственные шаги в решении конкретной проблемы.

Таким образом, мы рассмотрели две популярные языковые модели, одна из которых предназначена для решения произвольных задач, а другая — для поддержания длительного и сложного по структуре диалога. Отметим, что обе эти модели не показывают пользователю в явном виде, как было получено решение и никак это решение не обосновывают, не аргументируют.

Далее мы рассмотрим примеры систем, разработанных для более узких областей применения, но с более структурированным подходом к построению аргументативного диалога.

2.3. Система ведения аргументативного диалога, основанная на аргументативных структурах

Система, представленная в [17], была разработана для ведения дискуссии между пользователем и системой на различные темы. Это текстовая система (хотя поддерживается и голосовой интерфейс), анализирует сообщения пользователя, извлекает из них предпосылки аргументации [17], на их основе генерирует аргументы. Программно система реализована как набор модулей, объединенных с помощью Apache ActiveMQ.

Фраза пользователя преобразуется в «диалоговое действие». Авторами рассматриваются четыре типа диалоговых действий: утверждение, вопрос, уступка и отступление. Для распознавания этих действий используется классификатор на основе логистической регрессии. Далее идет поиск подходящего узла аргументации в графе аргументации на основе косинусного сходства между векторами предложений (причем сходство рассматривается

между усреднением значения узла и фразы пользователя). Извлеченная аргументация обрабатывается с учетом предметной области и генерируется следующий аргумент.

Отметим важную особенность этой системы — она способна оценивать действия пользователя, в частности, продолжает ли он свою мысль, собирается начать говорить или собирается закончить — это служит дополнительной информацией при порождении ответа и его обосновании.

Наполнение базы аргументов происходит при помощи методов автоматического извлечения аргументов, разработанных в [18]. На момент публикации статьи [17], система была способна понимать пять тем для обсуждения и поддерживать до 2000 узлов аргументации для каждой из них. При этом отсутствует возможность динамического добавления в структуру аргументации новых данных по ходу диалога, так что возникает естественный вопрос: как отреагирует система на новую информацию, которой нет в ее данных.

2.4. Системы аргументации, основанные на деревьях коммуникационного дискурса

Согласно теории речевых структур, любой связный дискурс может быть описан одним деревом дискурса, описанным в виде структуры дерева с использованием теории речевых действий [30]. Каждый абзац текста (или весь текст целиком) преобразуется в дерево через связывание предложений с помощью речевых действий (например, «Обоснование»), в листьях дерева находятся сами предложения. Таким образом, строится дерево аргументации, по которому можно определить наличие аргументации в абзаце/тексте. Кроме того, такой подход позволяет сохранить контекст аргументации, без которого даже эксперт-человек не сможет провести анализ наличия аргументации.

Важно отметить, что в работе [30] рассматривается именно факт наличия аргументации, а не ее смысловая часть и/или убедительность. Тем не менее, подобный подход можно и использовать для семантического анализа аргументации.

Подход, основанный на деревьях коммуникационного дискурса, также рассматривается в работах [31, 32].

2.5. Система ведения аргументативного диалога о вакцинации против COVID-19

Диалоговая система [19] предназначена для консультации пользователя на тему вакцинации, с максимальным обоснованием рекомендаций системы. Система основана на построении графа аргументации, согласно подходу Чалагина и Хантера [20] — выяснение сходства предложений, чтобы получить ответ из базы знаний. Метод не учитывает предыдущие действия пользователя и, как итог, теряет контекст диалога. Система стремится учитывать аргументы пользователя и строить ответ, не противоречащий им и, в то же время, согласующийся с базой знаний.

Модуль аргументации системы [20] состоит из графа аргументации, составленного экспертом. Узлы этого графа представляют либо аргументы состояния, либо аргументы ответа. С каждым узлом связан набор предложений на естественном языке, отражающий возможные аргументы пользователя для данного узла. Поиск подходящего узла в графе осуществляется при помощи меры сходства предложений. Порождение решения происходит на основе предоставленной пользователем информации, причем найденный узел должен согласоваться с данными пользователя и предотвращать «опасности», обозначенные пользователем неприемлемые моменты решения (в примере статьи речь идет о противопоказаниях к вакцинации). При этом, если система не может найти «безопасное» решение, она все равно выдаст ответ пользователю, но с запросом дополнительной информации для корректировки решения. Таким образом, каждый новый аргумент пользователя «включает» соответствующий узел в графе аргументации, а идущие от этого узла связи либо подкрепляют соответствующие варианты решений, либо выключают их из графа.

Таким образом, известные универсальные диалоговые системы хорошо справляются со многими задачами, но при этом они не имеют механизмов пояснения хода построения решения для пользователя. Кроме того, их архитектура не позволяет им быстро интегрировать в свои базы знаний решения, полученные от успешных сеансов работы с пользователями.

С другой стороны, специализированные системы, в которых аргументация закладывалась как изначальное требование, в основном полагаются на заранее подготовленную базу знаний, составлен-

ную экспертом, и строят свои решения и аргументы на ее основе, выход за пределы этой базы знаний приводит к построению ненадежного решения. За счет этого проявляется невозможность работы с результатами предыдущих сеансов, поскольку у них отсутствует механизм вставки такой информации в базу знаний.

3. Четырехуровневая онтологическая модель предметной области

Как было указано выше, целью работы является создание методов ведения интеллектуальным помощником (цифровым ассистентом) аргументированного диалога с пользователем [21]. В рамках нашего исследования разработка интеллектуального помощника [22] основана на семантической модели – четырехуровневой онтологической модели предметной области [10, 11]. Опишем эту модель более подробно.

Первый уровень онтологической модели – это онтологии:

1. Онтология предметной области рассматриваемых товаров (устройств) – набор понятий, описывающих: виды устройств; строение и характеристики устройств; функциональность устройств.
2. Онтология пользователей – набор понятий, описывающих: цели и намерения; интересы, желания, потребности; виды (классы) решаемых задач.
3. Онтология диалогов – набор понятий, описывающих: аргументацию (аргументы, контраргументы и др.); эмоциональные оценки пользователей, их удовлетворенность или неудовлетворенность; успешность данного диалога (приобретение товара пользователем, продолжение общения и других товарах или прекращение диалога пользователем, нежелание далее его продолжать).

Под онтологией мы понимаем знания только о смысле понятий, то есть аналитические утверждения [23–25], не содержащие информации о состоянии реального мира.

Второй уровень онтологической модели – это общие (универсальные) знания. Это синтетические утверждения [23, 26], знания о реальном мире:

1. Теория предметной области – свойства конкретных товаров, их характеристики, функциональность и т.д.
2. Знания о видах пользователей, их классификации (по уровню доходов, социальному положению,

образовательному уровню), классах задач, решаемых пользователями, иерархии задач, методам сведения задач к подзадачам и возможности решения одних задач при помощи разных устройств.

3. Знания о методах ведения диалогов с пользователями – о методах аргументации, обоснования конкретных предложений пользователю; о методах выявления целей и потребностей пользователей, решаемых ими задачах; о методах определения эмоционального состояния и эмоциональных оценок пользователей.

Третий уровень онтологической модели, наиболее важный в рамках данной работы – уровень прецедентов. Это:

1. Прецеденты товаров и устройств – конкретные устройства, комплектующие, аксессуары, наборы устройств, цена и наличие товаров в магазинах и пр.
2. Прецеденты пользователей – те пользователи, с которыми интеллектуальный помощник уже вел диалоги, вместе с их свойствами и характеристиками; знания о пользователях, их целях, намерениях, интересах, потребностях, решаемых ими задачах.
3. Прецеденты диалогов с пользователями, иерархически структурированные: диалог с одним пользователем; все диалоги с данным пользователем; диалоги с классами пользователей.

Четвертый уровень онтологической модели – это оценочные и вероятностные знания. Они порождаются за счет анализа прецедентов, содержащихся на третьем уровне онтологической модели. Это, в частности:

- ♦ вероятности того, что пользователь с определенными характеристиками, имеющий определенные потребности захочет приобрести данное устройство;
- ♦ вероятности того, что пользователь, имеющий (купивший) устройство **A** захочет приобрести устройство **B**;
- ♦ оценки сходства прецедентов: устройств, параметров устройств, пользователей и диалогов с ними.

На основе данной четырехуровневой семантической модели мы разрабатываем прецедентный подход к построению аргументированного диалога интеллектуального помощника с пользователем.

Активно используемые сейчас нейросетевые алгоритмы, в силу своей структуры, ограничены в использовании недавних прецедентов в своей модели –

лишь при обучении следующей версии эти данные могут быть включены в обучающую выборку.

Применение прецедентного подхода решает эту проблему: мы можем добавлять новые прецеденты «на лету», в процессе работы системы. Также использование прецедентного подхода делает диалоговую систему способной к аргументации собственных выводов, обоснованию выбора предлагаемых пользователю товаров.

Прецедентный подход опирается на набор примеров-прецедентов из прошлых сеансов работы с пользователями. Для определенных предметных областей это позволяет строить решения из уже имеющихся данных, применяя к ним некоторые преобразования, меняющие структуру прецедента-решения в соответствии с поставленной пользователем задачей (такие преобразования, в частности, реализуются при помощи онтологических гомоморфизмов частичных моделей, которые будут подробно описаны ниже).

При реализации прецедентного подхода возникает ряд проблем. Во-первых, прецеденты должны быть структурированы, причем не в виде реляционной таблицы с набором некоторых колонок – такой способ организации сделает поиск подходящего прецедента слабо связанным со смысловым содержанием прецедента. Во-вторых, необходимо оценивать степень сходства прецедентов – как для поиска подходящей «стартовой точки», так и для преобразования прецедента в итоговое решение. В-третьих, система должна иметь возможность преобразовывать прецеденты в соответствии с требованиями пользователя.

Решением указанных проблем является организация прецедентов в семантический граф, где связи между прецедентами будут отражать их схожие стороны, показывая степень подобия в том или ином свойстве прецедента. Это решает проблему семантического поиска – нужно пройти по графу прецедентов, следуя нужным семантическим связям. При этом удаленность двух узлов друг от друга будет явным образом отражать степень сходства соответствующих прецедентов. Процесс преобразования прецедентов может быть рассмотрен как преобразование частичных моделей, формально описывающих данные прецеденты, при помощи онтологических гомоморфизмов, расширений и сужений частичных моделей. В данном случае свойства и параметры исходного прецедента будут преобразовываться не обязательно в те же самые понятия, но в

онтологически близкие. Возможно и изоморфное вложение – расширение исходного прецедента, а также сужение прецедента, удаление лишних элементов модели.

Таким образом, мы реализуем программную систему, осуществляющую поиск решения на основе семантического сходства прецедентов с объяснимым механизмом поиска решения.

4. Теоретико-модельный подход к разработке интеллектуального помощника

Важно отметить, что большинство рассматриваемых нами прецедентов – как прецеденты товаров, так и прецеденты пользователей, содержат только часть всей информации о пользователе или устройстве. Поэтому в рамках теоретико-модельного подхода прецеденты должны формально описываться не обычными моделями (алгебраическими системами), а частичными моделями.

Определение. Рассмотрим сигнатуру $\sigma = \langle P_1, \dots, P_m, c_1, \dots, c_l \rangle$, в которой P_1, \dots, P_m – символы предикатов, а c_1, \dots, c_l – символы констант. Рассмотрим кортеж $\mathcal{A}^p = \langle A, P_1, \dots, P_m, c_1, \dots, c_l \rangle$, пусть для каждого $n \leq m$ значение n -местного предиката P_i на \mathcal{A}^p определено как пара $P_i^{\mathcal{A}^p} = (P_i^+, P_i^-)$, где $P_i^+, P_i^- \subseteq |\mathcal{A}^p|^n$ и $P_i^+ \cap P_i^- = \emptyset$. Назовем \mathcal{A}^p частичной моделью в сигнатуре σ . Будем считать, что для элементов $a_1, \dots, a_n \in |\mathcal{A}^p|^n$ если $(a_1, \dots, a_n) \in P_i^+$, то выполнено $\mathcal{A}^p \models P_i(a_1, \dots, a_n)$, если $(a_1, \dots, a_n) \in P_i^-$, то выполнено $\mathcal{A}^p \models \neg P_i(a_1, \dots, a_n)$, а если $(a_1, \dots, a_n) \notin (P_i^+ \cup P_i^-)$, то значение предиката $P_i(a_1, \dots, a_n)$ на частичной модели \mathcal{A}^p не определено.

Класс частичных моделей сигнатуры σ обозначим через $K^p(\sigma)$.

Для преобразования частичных моделей, формализующих прецеденты, мы используем онтологические гомоморфизмы. В данной статье мы рассмотрим три вида онтологических гомоморфизмов, наиболее важных для данного изложения; в качестве иллюстрации возьмем пример устройства – ноутбук. Это гомоморфизмы **обобщения** (в частичной модели наличие в ноутбуке разъема *USB A* заменяем на просто наличие разъема *USB*), гомоморфизм **уточнения** (наличие в ноутбуке разъема *USB* заменяем на наличие *USB A*) и гомоморфизм **сходства** (наличие разъема *USB A* заменяем на наличие разъема *USB C*).

Дадим строгие определения онтологических гомоморфизмов. Для этого рассмотрим онтологиче-

ские отношения: на множестве ключевых понятий онтологии предметной области рассматриваемых товаров (устройств) – предикатах сигнатуры σ , введем два двуместных отношения: отношение *общее-частное* $Hyp(Q, P)$ и отношение *сходства* $Sim(P, Q)$. Отношение Hyp является частичным порядком, а отношение Sim является рефлексивным и симметричным (но не обязательно транзитивным).

Определение. Рассмотрим частичные модели $\mathfrak{A}^p \in K^p(\sigma_1)$ и $\mathfrak{B}^p \in K^p(\sigma_2)$, пусть $P^n \in \sigma_1$, $Q^n \in \sigma_2$, $\sigma_1 \setminus \{P^n\} \subseteq \sigma_2$ и выполнено $Sim(P, Q)$. Отображение $h: |\mathfrak{A}^p| \rightarrow |\mathfrak{B}^p|$ назовем онтологическим гомоморфизмом *сходства* частичной модели \mathfrak{A}^p в частичную модель \mathfrak{B}^p ($h: \mathfrak{A}^p \rightarrow \mathfrak{B}^p$) если для любых $c \in \sigma_1$ и $a_1, \dots, a_n \in |\mathfrak{A}^p|$ выполнено:

- а) если $\mathfrak{A}^p \models P(a_1, \dots, a_n)$, то $\mathfrak{B}^p \models Q(h(a_1), \dots, h(a_n))$;
- б) если $\mathfrak{A}^p \models \neg P(a_1, \dots, a_n)$, то $\mathfrak{B}^p \models \neg Q(h(a_1), \dots, h(a_n))$;
- в) $h(c^{\mathfrak{A}^p}) = c^{\mathfrak{B}^p}$.

Истинность и ложность остальных предикатов из σ_1 сохраняются.

Определение. Рассмотрим частичные модели $\mathfrak{A}^p \in K^p(\sigma_1)$ и $\mathfrak{B}^p \in K^p(\sigma_2)$, $\sigma_1 \setminus \{P^n\} \subseteq \sigma_2$ и выполнено $Hyp(Q, P)$. Отображение $h: |\mathfrak{A}^p| \rightarrow |\mathfrak{B}^p|$ назовем онтологическим гомоморфизмом *обобщения* частичной модели \mathfrak{A}^p в частичную модель \mathfrak{B}^p ($h: \mathfrak{A}^p \rightarrow \mathfrak{B}^p$) если для любых $c \in \sigma_1$ и $a_1, \dots, a_n \in |\mathfrak{A}^p|$ выполнено:

- а) если $\mathfrak{A}^p \models P(a_1, \dots, a_n)$, то $\mathfrak{B}^p \models Q(h(a_1), \dots, h(a_n))$;
- б) $h(c^{\mathfrak{A}^p}) = c^{\mathfrak{B}^p}$.

Истинность и ложность остальных предикатов из σ_1 сохраняются.

Определение. Рассмотрим частичные модели $\mathfrak{A}^p \in K^p(\sigma_1)$ и $\mathfrak{B}^p \in K^p(\sigma_2)$, пусть $P^n \in \sigma_1$, $Q^n \in \sigma_2$, $\sigma_1 \setminus \{P^n\} \subseteq \sigma_2$ и выполнено $Hyp(Q, P)$. Отображение $h: |\mathfrak{A}^p| \rightarrow |\mathfrak{B}^p|$ назовем онтологическим гомоморфизмом *уточнения* частичной модели \mathfrak{A}^p в частичную модель \mathfrak{B}^p ($h: \mathfrak{A}^p \rightarrow \mathfrak{B}^p$) если для любых $c \in \sigma_1$ и $a_1, \dots, a_n \in |\mathfrak{A}^p|$ выполнено:

- а) если $\mathfrak{A}^p \models \neg P(a_1, \dots, a_n)$, то $\mathfrak{B}^p \models \neg Q(h(a_1), \dots, h(a_n))$;
- б) $h(c^{\mathfrak{A}^p}) = c^{\mathfrak{B}^p}$.

Истинность и ложность остальных предикатов из σ_1 сохраняются.

Однократное или многократное применение онтологических гомоморфизмов позволяет интеллектуальному помощнику автоматически переходить от описаний одних устройств к описаниям других, в определенной степени подобных.

Например, пользователь хочет купить некоторое устройство с определенными характеристиками, но требуемого устройства нет в продаже (или его цена не устраивает пользователя). Тогда интеллектуальный помощник автоматически находит другое устройство, частичная модель которого онтологически гомоморфна модели исходного, но которое имеется в продаже, и предлагает это устройство пользователю. Автоматически ищется товар (или несколько товаров), наиболее близкий к требуемому.

Предлагая пользователю устройства, интеллектуальный помощник также предоставляет объяснение, почему их отличие от желаемого пользователем не являются существенными с точки зрения решаемых пользователем задач. Такие объяснения на естественном языке либо заранее задаются в семантической модели, при определении онтологического сходства понятий, либо извлекаются из текстов естественного языка в процессе диалога (в частности, из описаний товаров на сайтах производителей и магазинов, из отзывов покупателей и т.п. [27, 28]).

Данный процесс продолжается итеративно, пока не будет найдено устройство, устраивающее пользователя: пользователь указывает, что ему не нравится, интеллектуальный помощник подбирает новый вариант. Таким образом строится граф аргументации, в вершинах которого находятся частичные модели, соответствующие устройствам, а переходы происходят при помощи онтологических гомоморфизмов.

Здесь важно отметить, что под прецедентами, мы понимаем всевозможные объекты, субъекты и ситуации, с которыми работает интеллектуальный помощник. Прецедентами являются как объекты предметной области, сами пользователи, так и диалоги с пользователями, результаты работы предыдущего сеанса программной системы – артефакты взаимодействия интеллектуального помощника с пользователем. В качестве примера рассмотрим иерархию структурированных прецедентов – объектов предметной области, связанных с удовлетворением потребностей пользователя:

1. Объект предметной области, товар, который нужен пользователю (например, компьютер или смартфон).
2. Объект предметной области + потребности пользователя (которые выявил интеллектуальный помощник в процессе диалога).

3. Объект предметной области + потребности пользователя + класс решаемых пользователем задач. Мы расширяем прецедент, добавив к нему задачи, которые необходимо решать пользователю. Важно отметить, что свойства объекта предметной области (например, функциональность устройства) четко определены в базе знаний и независимы от восприятия пользователя. С другой стороны, задачи, которые пользователь намерен решить при помощи данного устройства, зависят от конечных целей пользователя, его желаний и потребностей. Класс решаемых задач определяет сам пользователь. Имея данные о потребностях пользователя и решаемых им задачах, мы строим прецедент как тройку: <частичная модель, описывающая устройство; потребности пользователя; набор (класс) решаемых пользователем задач>.

Такой способ представления прецедентов позволяет объединить объективную информацию о товарах с субъективной информацией о пользователе, полученной интеллектуальным помощником в процессе диалога. Напомним, что прецеденты представлены на третьем уровне онтологической модели предметной области.

Построение решения – подбор товара, нужного пользователю, предполагает сравнение как объектов предметной области (товаров, устройств), так и структурированных прецедентов, описанных выше. Для этого используется аппарат метрик, которые позволяют проводить сравнение объектов и прецедентов в ходе работы интеллектуального помощника. В частности, используется семантический граф прецедентов с заранее рассчитанной (или установленной экспертом) мерой сходства (подобия) прецедентов. Знания о мерах сходства прецедентов относятся к четвертому уровню онтологической модели предметной области. Исходя из этих знаний, в частности, задается отношение сходства $Sim(P, Q)$, рассмотренное выше при определении онтологических гомоморфизмов.

При подборе товаров приоритеты пользователя рассчитываются исходя из двух параметров: во-первых, свойств и функциональности устройств и, во-вторых, потребностей и желаний пользователя, класса решаемых им задач – то есть, объективных и субъективных параметров. Ориентируясь на эти два вида параметров, мы вычисляем сходство различных прецедентов, в том числе, объектов данной предметной области.

5. Программная реализация диалоговой системы

Разработанная программная система [22] представляет собой набор из пяти блоков (модулей), обеспечивающих различные этапы работы системы. Техническая реализация представляет собой MVC приложение на Java Spring, с REST интерфейсом.

Блок 1 отвечает за осуществление ввода пользователя и его формализацию через механизм поиска речевых действий [29], коррекцию модели пользователя и выявление пользовательских намерений.

Блок 2 отвечает за анализ полученного ввода, а также порождение сообщений системы о необходимости получения дополнительной информации у пользователя.

Блок 3 отвечает за поиск требуемого товара – прецедента предметной области. Осуществляется проверка – является ли частичная модель требуемого товара, построенная в результате диалога с пользователем, подмоделью модели некоторого прецедента предметной области (т.е., товара, имеющегося в наличии). На вход блока подаются сформулированные и проверенные на непротиворечивость потребности пользователя. Далее при помощи онтологических гомоморфизмов, реализованных на основе функции сходства двух частичных моделей, происходит поиск прецедента, наиболее сходного с желаемым пользователем.

Если прецедент найден, то данное решение будет предложено пользователю, описание найденного прецедента на естественном языке будет использовано в качестве обоснования (аргумента).

Блок 4 отвечает за анализ реакции пользователя. Основная функция блока – уточнить требования пользователя. Если в ответе пользователя есть новые данные, они формализуются с помощью механизмов блоков 1 и 2 и снова запускается блок 3. Таким образом процесс поиска решения происходит итеративно.

Блок 5 отвечает за окончательную генерацию решения и его обоснование (аргументацию). Пользователю предлагается товар, полностью соответствующий его требованиям, выявленным в процессе диалога. Также может быть представлен набор товаров, отвечающих требованиям, но отличающихся по цене или несущественным для пользователя характеристикам.

Заключение

В статье разрабатываются методы создания интеллектуальных помощников. Интеллектуальные помощники могут применяться для помощи выбора товаров — как рекомендательные системы, в колл-центрах для решения различных проблем клиентов, для решения задач техподдержки, для помощи людям с ограниченными возможностями. В данной работе в первую очередь рассматриваются интеллектуальные помощники, предназначенные для помощи выбора товаров пользователем.

Для создания интеллектуальных помощников мы разрабатываем методы ведения аргументированного диалога с пользователем. Для этого мы разрабатываем методы автоматизированного построения обоснований и аргументации. Формализация рассуждений и аргументации производится при помощи частичных моделей, гомоморфизмов и онтологических гомоморфизмов частичных моделей. Онтологические гомоморфизмы сходства частичных моделей формально описывают подобие прецедентов, что служит математическим базисом построения аргументации на основе прецедентов.

Предложенная архитектура программной системы реализует методы ведения диалога с использованием прецедентов, пополнением базы прецедентов после каждого сеанса работы, их организацию в виде семантической сети. Подобный подход позволяет добиться прозрачности работы системы, повысить гибкость подбора решения за счет анализа смыслового содержания вводимых пользователем фраз (с помощью атомарных моделей), что отличает систему от имеющихся аналогов.

Дальнейшее развитие системы возможно в направлении улучшения алгоритма поиска прецедента в семантической сети, с введением большего количества связей между прецедентами для повышения детализации поиска, а также разработки методов определения сходства и подобия прецедентов.

Благодарности

Работа выполнена в рамках государственного задания Института математики им. С.Л. Соболева Сибирского отделения Российской академии наук, проект FWNF-2022-0011. ■

Литература

1. Toulmin S.E. The uses of argument. Cambridge University Press, 2003.
2. Reed C., Rowe G. Araucaria: Software for argument analysis, diagramming and representation // International Journal on Artificial Intelligence Tools. 2004. Vol. 13. No. 4. P. 961–979.
3. Walton D., Reed C., Macagno F. Argumentation schemes. Cambridge University Press, 2008.
4. Macagno F., Walton D. Classifying the patterns of natural arguments // Philosophy & Rhetoric. 2015. Vol. 48. No. 1. P. 26–53. <https://doi.org/10.5325/philrhet.48.1.0026>
5. Wagemans J. Constructing a periodic table of arguments // Argumentation, objectivity, and bias: Proceedings of the 11th international conference of the Ontario Society for the Study of Argumentation, Windsor, Ontario, Canada, 2016. P. 1–12. <https://doi.org/10.2139/ssrn.2769833>
6. Walton D., Gordon T.F. Argument invention with the Carneades argumentation system // SCRIPTed. 2017. Vol. 14. No. 2. P. 168–207. <https://doi.org/10.2966/scrip.140217.168>
7. Wachsmuth H., Potthast M., Al-Khatib K., Ajjour Y. Building an argument search engine for the Web // Proceedings of the 4th Workshop on Argument Mining, 2017. P. 49–59.
8. Slonim N. et al. An autonomous debating system // Nature. 2021. Т. 591. No. 7850. P. 379–384. <https://doi.org/10.1038/s41586-021-03215-w>
9. Stab C., Miller T., Shiller B., Rai P., Gurevych I. Cross-topic argument mining from heterogeneous sources // Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018. P. 3664–3674.
10. Naydanov Ch., Palchunov D., Sazonova P. Development of automated methods for the critical condition risk prevention, based on the analysis of the knowledge obtained from patient medical records // Proceedings of the International Conference on Biomedical Engineering and Computational Technologies, Novosibirsk, 28–30 October 2015. P. 33–38. <https://doi.org/10.1109/SIBIRCON.2015.7361845>
11. Naydanov Ch., Palchunov D., Sazonova P. Development of automated methods for the prevention of risks of critical conditions, based on the analysis of the knowledge extracted from the medical histories // The Siberian Scientific Medical Journal. 2016. Vol. 36. No. 1. P. 105–113.
12. Найданов Ч.А. Разработка ядра онтологической модели, настраиваемой под предметную область // Вестник Новосибирского государственного университета. Серия: Информационные технологии. 2019. Т. 17. №. 1. С. 72–81.

13. Introducing ChatGPT // OpenAI, 2022. [Электронный ресурс]: <https://openai.com/blog/chatgpt> (дата обращения 26.01.2024).
14. McCloskey M., Cohen N.J. Catastrophic interference in connectionist networks: The sequential learning problem // *Psychology of Learning and Motivation*. 1989. Vol. 24. P. 109–165.
15. BlenderBot 3: A Conversational AI Prototype // Blenderbot AI. [Электронный ресурс]: <https://blenderbot.ai> (дата обращения 26.01.2024).
16. Shuster K. et al. Blenderbot 3: A deployed conversational agent that continually learns to responsibly engage // *arXiv:2208.03188*. 2022. <https://doi.org/10.48550/arXiv.2208.03188>
17. Higashinaka R. et al. Argumentative dialogue system based on argumentation structures // *Proceedings of the 21st Workshop on the Semantics and Pragmatics of Dialogue*. 2017. P. 154–155.
18. Lippi M., Torroni P. Argumentation mining: State of the art and emerging trends // *ACM Transactions on Internet Technology*. 2016. Vol. 16. No. 2. P. 1–25. <https://doi.org/10.1145/2850417>
19. Fazzinga B., Galassi A., Torroni P. An argumentative dialogue system for COVID-19 vaccine information // *International Conference on Logic and Argumentation*. Cham: Springer International Publishing, 2021. P. 477–485.
20. Chalaguine L.A., Hunter A. A persuasive chatbot using a crowd-sourced argument graph and concerns // *IOS Press Ebooks Series Frontiers in Artificial Intelligence and Applications*. 2020. Vol. 326: Computational Models of Argument. P. 9–20. <https://doi.org/10.3233/FAIA200487>
21. Palchunov D., Yakobson A. Automated methods for conducting an argued dialogue with the user // *2022 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*, 2022. P. 880–885. <https://doi.org/10.1109/SIBIRCON56155.2022.10016975>
22. Пальчунов Д.Е., Якобсон А.А. Программная система «АргументДиалог». Свидетельство о государственной регистрации программы для ЭВМ № 2022664041, зарегистрировано 22.07.2022.
23. Carnap R. *Meaning and necessity. A study in semantics and modal logic*. Chicago, 1956.
24. Пальчунов Д.Е. Моделирование мышления и формализация рефлексии I: Теоретико-модельная формализация онтологии и рефлексии // *Философия науки*. 2006. № 4(31). С. 86–114.
25. Palchunov D.E. Methodological aspects of the application of model theory in knowledge engineering and artificial intelligence // *2022 Ural-Siberian Conference on Computational Technologies in Cognitive Science, Genomics and Biomedicine (CSGB)*, 2022. P. 210–215. <https://doi.org/10.1109/CSGB56354.2022.9865602>
26. Пальчунов Д.Е. Моделирование мышления и формализация рефлексии. Ч. 2: Онтологии и формализация понятий // *Философия науки*. 2008. № 2(37). С. 62–99.
27. Пальчунов Д.Е. Решение задачи поиска информации на основе онтологий // *Бизнес-информатика*. 2008. № 1(3). С. 3–13.
28. Orlovsky A., Palchunov D. Development of automated methods for the domain ontology population with the help of a virtual assistant // *2021 IEEE 22nd International Conference of Young Professionals in Electron Devices and Materials (EDM)*, 2021. P. 537–541. <https://doi.org/10.1109/EDM52169.2021.9507641>
29. Махина Е.Д., Пальчунов Д.Е. Программная система для определения речевых действий в текстах естественного языка // *Вестник Новосибирского государственного университета. Серия: Информационные технологии*. 2018. Т. 16. № 4. С. 95–106. <https://doi.org/10.25205/1818-7900-2018-16-4-95-106>
30. Galitsky B., Ilvovsky D., Kuznetsov S. Detecting logical argumentation in text via communicative discourse tree // *Journal of Experimental and Theoretical Artificial Intelligence*. 2018. Vol. 30. No. 5. P. 637–663. <https://doi.org/10.1080/0952813X.2018.1467492>
31. Galitsky B., Ilvovsky D., Kuznetsov S.O. Text integrity assessment: Sentiment profile vs rhetoric structure // *Computational Linguistics and Intelligent Text Processing* (eds. A. Gelbukh). CILing 2015. Lecture Notes in Computer Science. Vol. 9042. Springer, Cham, 2015. P. 126–139. https://doi.org/10.1007/978-3-319-18117-2_10
32. Galitsky B., Ilvovsky D., Kuznetsov S.O. Rhetoric map of an answer to compound queries // *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Beijing, China, 2015*. Vol. 2: Short Papers. P. 681–686. <https://doi.org/10.3115/v1/P15-2112>

Об авторах

Пальчунов Дмитрий Евгеньевич

д.ф.-м.н., академик Российской инженерной академии;

в.н.с., Институт математики им. С.Л. Соболева, Сибирское отделение Российской академии наук, Россия, 630090, г. Новосибирск, пр-т Академика Коптюга, д. 4;

E-mail: palch@math.nsc.ru

ORCID: 0000-0001-9487-3256

Якобсон Александр Алексеевич

ассистент, кафедра общей информатики, Новосибирский национальный исследовательский государственный университет, Россия, 630090, г. Новосибирск, ул. Пирогова, д. 1;

E-mail: a.yakobson@g.nsu.ru

Development of an intelligent assistant for selection of goods in the process of dialogue with the user

Dmitry E. Palchunov^a

E-mail: palch@math.nsc.ru

Alexander A. Yakobson^b

E-mail: a.yakobson@g.nsu.ru

^a S.L. Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences
Address: 4, Akademika Koptyuga Ave., Novosibirsk 630090, Russia

^b Novosibirsk National Research State University
Address: 1, Pirogova St., Novosibirsk 630090, Russia

Abstract

This article is devoted to the development of methods for creating intelligent assistants. Intelligent assistants can be used in call centers to solve customer problems, to solve technical support tasks, to help people with disabilities, to help in choosing goods, etc. We consider intelligent assistants that engage in argumentative dialogue with users, aimed at finding goods and services that maximally satisfy users' wants and needs. The development of the intelligent assistant is based on a four-level model of the subject domain and a semantic model of the user. The system under development automates the process of search and decision justification through the reuse of domain cases: accumulated knowledge about previous dialogues with users. This gives the system we developed an advantage over existing analogues, which are incapable of reusing knowledge about previous dialogues. The paper develops a case-based approach to building an intelligent system capable of reasoning about its responses. For this purpose, an argumentation graph is constructed, methods for structuring domain cases are developed, and ontological homomorphisms are used to transform the available domain cases into a finished solution. A description of model-theoretical methods for constructing intelligent assistants is presented. The cases of goods, users and dialogues of an intelligent assistant with users are formally described in the form of partial models. The transformation of domain cases and similarity of cases are formalized using ontological homomorphisms of partial models. The purpose of the developed dialogue system is not only to select a solution according to the user's request, but also to find out the tasks that the user is going to solve, to analyze his argumentation, and then to justify the proposed solution to the user, to show that this particular product or service will be able to meet his needs.

Keywords: intelligent assistant, argumentative dialogue, domain case, partial model, ontological homomorphism, ontological model of the subject domain, semantic model of the user

Citation: Palchunov D.E., Yakobson A.A. (2024) Development of an intelligent assistant for selection of goods in the process of dialogue with the user. *Business Informatics*, vol. 18, no. 1, pp. 7–21. DOI: 10.17323/2587-814X.2024.1.7.21

References

1. Toulmin S.E. (2003) *The uses of argument*. Cambridge University Press.
2. Reed C., Rowe G. (2004) Araucaria: Software for argument analysis, diagramming and representation. *International Journal on Artificial Intelligence Tools*, vol. 13, no. 4, pp. 961–979.
3. Walton D., Reed C., Macagno F. (2008) *Argumentation schemes*. Cambridge University Press.
4. Macagno F., Walton D. (2015) Classifying the patterns of natural arguments. *Philosophy & Rhetoric*, vol. 48, no. 1, pp. 26–53. <https://doi.org/10.5325/philrhet.48.1.0026>
5. Wagemans J. (2016) Constructing a periodic table of arguments. *Argumentation, objectivity, and bias: Proceedings of the 11th international conference of the Ontario Society for the Study of Argumentation, Windsor, Ontario, Canada*, pp. 1–12. <https://doi.org/10.2139/ssrn.2769833>
6. Walton D., Gordon T.F. (2017) Argument invention with the Carneades argumentation system. *SCRIPTed*, vol. 14, no. 2, pp. 168–207. <https://doi.org/10.2966/scrip.140217.168>
7. Wachsmuth H., Potthast M., Al-Khatib K., Ajjour Y. (2017) Building an argument search engine for the Web. *Proceedings of the 4th Workshop on Argument Mining*, pp. 49–59.
8. Slonim N. et al. (2021) An autonomous debating system. *Nature*, vol. 591, no. 7850, pp. 379–384. <https://doi.org/10.1038/s41586-021-03215-w>
9. Stab C., Miller T., Shiller B., Rai P., Gurevych I. (2018) Cross-topic argument mining from heterogeneous sources. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 3664–3674.
10. Naydanov Ch., Palchunov D., Sazonova P. (2015) Development of automated methods for the critical condition risk prevention, based on the analysis of the knowledge obtained from patient medical records. *Proceedings of the International Conference on Biomedical Engineering and Computational Technologies, Novosibirsk, 28–30 October 2015*, pp. 33–38. <https://doi.org/10.1109/SIBIRCON.2015.7361845>
11. Naydanov Ch., Palchunov D., Sazonova P. (2016) Development of automated methods for the prevention of risks of critical conditions, based on the analysis of the knowledge extracted from the medical histories. *The Siberian Scientific Medical Journal*, vol. 36, no. 1, pp. 105–113.
12. Naidanov Ch.A. (2019) Development of ontological model kernel, customisable for the subject domain. *Vestnik of Novosibirsk State University. Series: Information technologies*, vol. 17, no. 1, pp. 72–81.
13. OpenAI (2022) *Introducing ChatGPT*. Available at: <https://openai.com/blog/chatgpt> (accessed 26 January 2024).
14. McCloskey M., Cohen N.J. (1989) Catastrophic interference in connectionist networks: The sequential learning problem. *Psychology of Learning and Motivation*, vol. 24, pp. 109–165.
15. Blenderbot AI (2022) *BlenderBot 3: A Conversational AI Prototype*. Available at: <https://blenderbot.ai> (accessed 26 January 2024).
16. Shuster K. et al. (2022) Blenderbot 3: A deployed conversational agent that continually learns to responsibly engage. *arXiv:2208.03188*. <https://doi.org/10.48550/arXiv.2208.03188>
17. Higashinaka R. et al. (2017) Argumentative dialogue system based on argumentation structures. *Proceedings of the 21st Workshop on the Semantics and Pragmatics of Dialogue*, pp. 154–155.
18. Lippi M., Torroni P. (2016) Argumentation mining: State of the art and emerging trends. *ACM Transactions on Internet Technology*, vol. 16, no. 2, pp. 1–25. <https://doi.org/10.1145/2850417>
19. Fazzinga B., Galassi A., Torroni P. (2021) An argumentative dialogue system for COVID-19 vaccine information. *International Conference on Logic and Argumentation*. Cham: Springer International Publishing, pp. 477–485.
20. Chalaguine L.A., Hunter A. (2020) A persuasive chatbot using a crowd-sourced argument graph and concerns. *IOS Press E-books Series Frontiers in Artificial Intelligence and Applications*, vol. 326: Computational Models of Argument, pp. 9–20. <https://doi.org/10.3233/FAIA200487>
21. Palchunov D., Yakobson A. (2022) Automated methods for conducting an argued dialogue with the user. *2022 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*, pp. 880–885. <https://doi.org/10.1109/SIBIRCON56155.2022.10016975>

22. Palchunov D.E., Yakobson A.A. (2022) *Software system "ArgumentDialogue"*. Certificate of state registration of computer programme No. 2022664041, registered 22.07.2022.
23. Carnap R. (1956) *Meaning and necessity*. A study in semantics and modal logic. Chicago.
24. Palchunov D.E. (2006) Modelling of thinking and formalization of reflexion I: Theoretical-model formalization of ontology and reflexion. *Philosophy of Science*, vol. 4(31), pp. 86–114.
25. Palchunov D.E. (2022) Methodological aspects of the application of model theory in knowledge engineering and artificial intelligence. *2022 Ural-Siberian Conference on Computational Technologies in Cognitive Science, Genomics and Biomedicine (CSGB)*, pp. 210–215. <https://doi.org/10.1109/CSGB56354.2022.9865602>
26. Palchunov D.E. (2008) Modelling of thinking and formalization of reflexion. Part 2: Ontologies and formalisation of concepts. *Philosophy of Science*, vol. 2(37), pp. 62–99.
27. Palchunov D.E. (2008) Solution of the information search task on the basis of ontologies. *Business-informatics*, vol. 1(3), pp. 3–13 (in Russian).
28. Orlovsky A., Palchunov D. (2021) Development of automated methods for the domain ontology population with the help of a virtual assistant. *2021 IEEE 22nd International Conference of Young Professionals in Electron Devices and Materials (EDM)*, pp. 537–541. <https://doi.org/10.1109/EDM52169.2021.9507641>
29. Makhina E.D., Palchunov D.E. (2018) Software system to determine speech actions in natural language texts. *Vestnik of Novosibirsk State University. Series: Information technologies*, vol. 16, no. 4, pp. 95–106. <https://doi.org/10.25205/1818-7900-2018-16-4-95-106>
30. Galitsky B., Ilvovsky D., Kuznetsov S. (2018) Detecting logical argumentation in text via communicative discourse tree. *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 30, no. 5, pp. 637–663. <https://doi.org/10.1080/0952813X.2018.1467492>
31. Galitsky B., Ilvovsky D., Kuznetsov S.O. (2015) Text integrity assessment: Sentiment profile vs rhetoric structure. *Computational Linguistics and Intelligent Text Processing* (eds. A. Gelbukh). CICLing 2015. Lecture Notes in Computer Science, vol. 9042. Springer, Cham, pp. 126–139. https://doi.org/10.1007/978-3-319-18117-2_10
32. Galitsky B., Ilvovsky D., Kuznetsov S.O. (2015) Rhetoric map of an answer to compound queries. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Beijing, China*, vol. 2: Short Papers, pp. 681–686. <https://doi.org/10.3115/v1/P15-2112>

About the authors

Dmitry E. Palchunov

Dr. Sci. (Phys.-Math.), Academician of the Russian Academy of Engineering;

Leading Researcher, S.L. Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, 4, Akademika Koptuyuga Ave., Novosibirsk 630090, Russia;

E-mail: palch@math.nsc.ru

ORCID: 0000-0001-9487-3256

Alexander A. Yakobson

Assistant, Department of General Informatics, Novosibirsk National Research State University, 1, Pirogova St., Novosibirsk 630090, Russia;

E-mail: a.yakobson@g.nsu.ru

DOI: 10.17323/2587-814X.2024.1.22.35

Совершенствование процесса целевого бюджетирования в системе управления корпоративной результативностью

М.Е. Ощепков 

E-mail: moshchepkov@hse.ru

Высшая школа бизнеса, Национальный исследовательский университет «Высшая школа экономики»
Адрес: Россия, 119049, г. Москва, ул. Шаболовка, д. 26–28

Аннотация

В настоящее время представление процедуры выработки управленческих воздействий при осуществлении процесса планирования в научном и профессиональном сообществе зачастую не соответствует практике системного и последовательного формирования планов при поддержке информационно-аналитической системы. Вместо этого процесс планирования опирается на неформализованную деятельность лица, принимающего решение, которой характерен точечный ситуативный экспертный подход. Данное исследование направлено на разработку аналитического подхода к осуществлению процедуры согласования планов в процессе целевого бюджетирования, который позволит расширить использование возможностей системы управления корпоративной результативностью и обеспечить математическую формализацию задачи выработки управленческих воздействий при корректировке значений плановых ключевых показателей. С этой целью типовой процесс планирования в рамках системы управления корпоративной результативностью дополняется блоками аналитической поддержки, а именно алгоритмом обратных вычислений частных ключевых показателей и расширенным модулем сценарного моделирования. Представленная усовершенствованная модель процесса целевого бюджетирования обеспечивает автоматизированное формирование управленческих воздействий отдела бюджетирования и руководства подразделений в направлении достижения стратегических целей. Применение обратных вычислений обеспечивает математическую постановку задачи вычисления индикаторов плановых ключевых значений, а система Sense and Respond (SaR) позволяет дополнить математическую постановку весовыми коэффициентами первичных показателей эффективности, рассчитанными алгоритмически на основе управленческих решений менеджера вместо экспертной оценки. Реализация разработанного подхода обеспечивает повышение качества планирования по наиболее приоритетным критериям качества: оперативности, точности и адаптивности – за счет системности и методичности составления бюджетов с привлечением современных информационных технологий.

Ключевые слова: управление корпоративной результативностью, бюджетирование, информационно-логическая модель, обратные вычисления, математическое программирование

Цитирование: Ошепков М.Е. Совершенствование процесса целевого бюджетирования в системе управления корпоративной результативностью // Бизнес-информатика. 2024. Т. 18. № 1. С. 22–35. DOI: 10.17323/2587-814X.2024.1.22.35

Введение

Современная экономика характеризуется изменением экономических и социальных парадигм в ответ на различные кризисные явления, что вызвало повышенное внимание к измерению результативности не только в целом в экономике, но также и в отдельных компаниях и их структурных подразделениях [1]. Данный фактор стимулировал активное развитие и постепенное внедрение систем измерения и управления корпоративной результативностью (corporate performance management, СРМ), которые основаны на системном подходе и мышлении для обеспечения правдивости в измерении экономических явлений.

На пути развития и применения систем управления корпоративной результативностью возникает множество проблем, которые требуют решения для обеспечения устойчивого развития и получения конкурентных преимуществ. В современных реалиях конкурентоспособность компаний обеспечивается за счет высокого качества планирования экономической деятельности, определяемое исходя их различных критериев, основными из которых являются точность, оперативность и адаптивность. Достижимость высоких значений обозначенных критериев зависит от развития инструментария (системы и правила) и технологии управления (модели и методы). Методы управления на корпоративном уровне определяют успех организации, поскольку они определяют стратегию организации и обеспечивают ее выполнение [2].

Практический опыт компаний в процессе создания и внедрения СРМ-систем демонстрирует наличие существенных трудностей при выстраивании и развитии инструментария и технологии управления в процессе целереализации. Выработка управленческих воздействий, как правило, представляет собой процесс неформализованной деятельности менеджера, основанный на экспертном подходе. При этом методы и процедуры, применяемые в практике компаний, не позволяют обеспечить системность и методичность установления плановых ключевых показателей, соответствующих требова-

ниям единой стратегической цели для всех структурных единиц компании. Одна из причин возникновения обозначенных проблем выражается в отсутствии четкого понимания того, что невозможно воплотить корпоративную стратегию в бюджетное планирование, не имея конкретной математической базы, на которой основывается управление результативностью деятельности компании.

Данная проблематика актуальна на протяжении последних десятилетий, что находит отражение в исследованиях как зарубежных, так и отечественных авторов [1, 3–9].

В работе предлагается аналитический подход к совершенствованию типового процесса бюджетирования за счет разработки и внедрения алгоритма обратных вычислений, позволяющего формализовать и автоматизировать процедуру формирования управленческих воздействий на частные плановые показатели, и модуля «Факт–Прогноз», обеспечивающего расширение возможностей сценарного моделирования.

Предлагаемый алгоритм обратных вычислений обеспечивает расчет индикаторов плановых ключевых показателей с учетом их приоритетности на основе управленческих решений менеджера. Методическая основа алгоритма строится на объединении теории обратных вычислений и методологии SaR. Математическая постановка задачи вычисления индикаторов плановых ключевых значений обеспечивается применением метода обратных вычислений, а система SaR дополняет математическую постановку весовыми коэффициентами ключевых первичных показателей эффективности, рассчитанными алгоритмически на основе управленческих решений менеджера вместо экспертной оценки.

Модуль сценарного моделирования позволяет планомерно-экономическому управлению параллельно основному процессу бюджетирования осуществлять выработку управленческих воздействий на основе реальных данных от центров финансовой ответственности (ЦФО) и вычисленных с применением алгоритма обратных вычислений и SaR

плановых ключевых показателей, формируя в качестве результата сценарии бюджетирования.

Преимущества предлагаемого подхода заключаются в повышении качества процесса планирования по всем приоритетным критериям качества: оперативность, точность, адаптивность. Оперативность достигается за счет сокращения времени и трудозатрат плано-экономического управления (ПЭУ) на осуществление процедуры выработки управленческих воздействий. Кроме того, сокращается количество корректировок плановых расчетов бюджета со стороны подразделений компании. Повышение точности планирования обеспечивается методичностью и системностью реализации процесса, которые предполагают взаимодействие разработанных аналитических блоков, способствующее эффективно согласованию стратегических целей и действий по их достижению. Адаптивность планирования повышается за счет сокращения времени на корректировки данных для новых шагов планирования и увеличения скорости адаптации системы к наблюдаемым изменениям внешней среды, поскольку в алгоритм заложено условие аппроксимации поведения лица, принимающего решение (ЛПР).

1. Понятие «управление корпоративной результативностью»

Понятие «управление корпоративной результативностью» включает информационные технологии и инструменты, методы и процессы управления, а также человеческие ресурсы и предполагает периодические измерение и анализ ключевых показателей, ориентированных на достижение определенных целей [6]. Система измерения результативности, согласно авторам основополагающих работ в данной области – это набор показателей, используемых для количественной оценки как внутренней эффективности (экономичность), так и внешней эффективности (эффективность) [10, 11].

В литературе по измерению корпоративной результативности описывается широкий спектр подходов к проектированию систем измерения результативности, которые классифицируются по различным аспектам. Согласно Bourne [3] можно использовать два измерения: лежащая в основе измерения процедура (индикаторы потребностей [12, 13], индикаторы аудита [14], индикаторы модели [15]) и основополагающий подход с точки зрения роли руководителя процесса (руководство консультанта [16], руководство фасилитатора [11]).

Исходя из концепции целевого управления, все множество ключевых показателей результативности (key performance indicators, KPI) можно классифицировать следующим образом [8]: первичные показатели, предназначенные для измерения расходов и потребностей в ресурсах, необходимых для достижения поставленных целей – хранятся в базе данных бухгалтерского учета; интегрированные показатели-индикаторы, предназначенные для оценки эффективности по отдельным направлениям деятельности компании – являются расчетными показателями; показатели-индикаторы, предназначенные для измерения результативности деятельности компании в целом и уровня достижения стратегических целей – являются производными показателями.

Целевое управление имеет три уровня: стратегический, тактический и оперативный. Стратегическому уровню соответствует стратегическое планирование, осуществляемое руководителями, которое определяет направление развития компании в долгосрочной перспективе (более 3 лет). В фокусе данной работы находится плано-экономическое управление (ПЭУ) или отдел бюджетирования, которые осуществляют целевое бюджетирование. Целевое бюджетирование предполагает разработку бюджета, ориентированного на достижение целей оперативного планирования, определенных на основе значений ключевых показателей результативности из стратегической карты компании. В этом случае бюджет представляет собой тактический план с горизонтом в один год, отражающий результаты операционной, инвестиционной и иной деятельности компании [17].

Технология управления результативностью бизнеса включает, как минимум, четыре основных блока: моделирование, планирование, мониторинг, анализ – и является последовательностью с замкнутым циклом [18].

Блок моделирования находится за пределами общего цикла управления, так как выполняется поэтапно в зависимости от темпов развития компании и степени ее адаптации к внешней среде. Данный блок включает в себя формулирование стратегической карты, построение деревьев целей, установление ключевых показателей, моделирование бизнес-процессов.

Блок планирования является укрупненным и в интерпретациях различных авторов зачастую раз-

бывается на несколько блоков [18]. Так, согласно Б.Е. Одинцову [8], двумя этапами, следующими за этапом моделирования, являются прогнозирование ключевых показателей результативности и расчет плановых значений ключевых показателей результативности.

Блок непосредственного планирования ключевых показателей результативности является наиболее важной составляющей технологии управления результативностью бизнеса, поскольку обеспечивает систематическое представление целевых значений ключевых показателей результативности ПЭУ и другим структурным подразделениям для согласования прогнозов и формирования окончательно утвержденного плана на будущий период. Данный блок включает факт–прогноз анализ и расчет прогнозных значений показателей в рамках скользящего отчетного периода. Именно от значений показателей, полученных на данном этапе, зависит достижение оперативной цели, указанной руководством на последующий период.

Два других блока концепции управления результативностью бизнеса, мониторинг и анализ, предполагают оперативный контроль за финансово-экономическим состоянием компании и комплексное диагностирование показателей компании соответственно.

2. Проблема исследования

В качестве основы формирования предположений по совершенствованию методической базы управления корпоративной результативностью было рассмотрено типовое решение системы управления корпоративной результативностью замкнутого цикла, характерное для подавляющего большинства отечественных, а также международных компаний. На его примере продемонстрированы узкие места алгоритма, на разрешение которых направлено данное исследование (рис. 1).

Схема отражает этапы типового процесса финансового планирования и прогнозирования показателей сбыта корпорации, дополненные блоками аналитического подхода к определению корректировок прогнозов и согласованию плановых бюджетов (выделены пунктирными линиями на схеме).

В описанном представлении типового процесса финансового планирования и прогнозирования показателей компании можно отметить несколько «узких мест», которые ограничивают потенциальные системные эффекты воздействия на качество процесса при использовании возможностей информационной системы и аналитических методов. Из всех возможных ограничений рассмотренной модели в качестве целевых для данного исследования были отмечены:

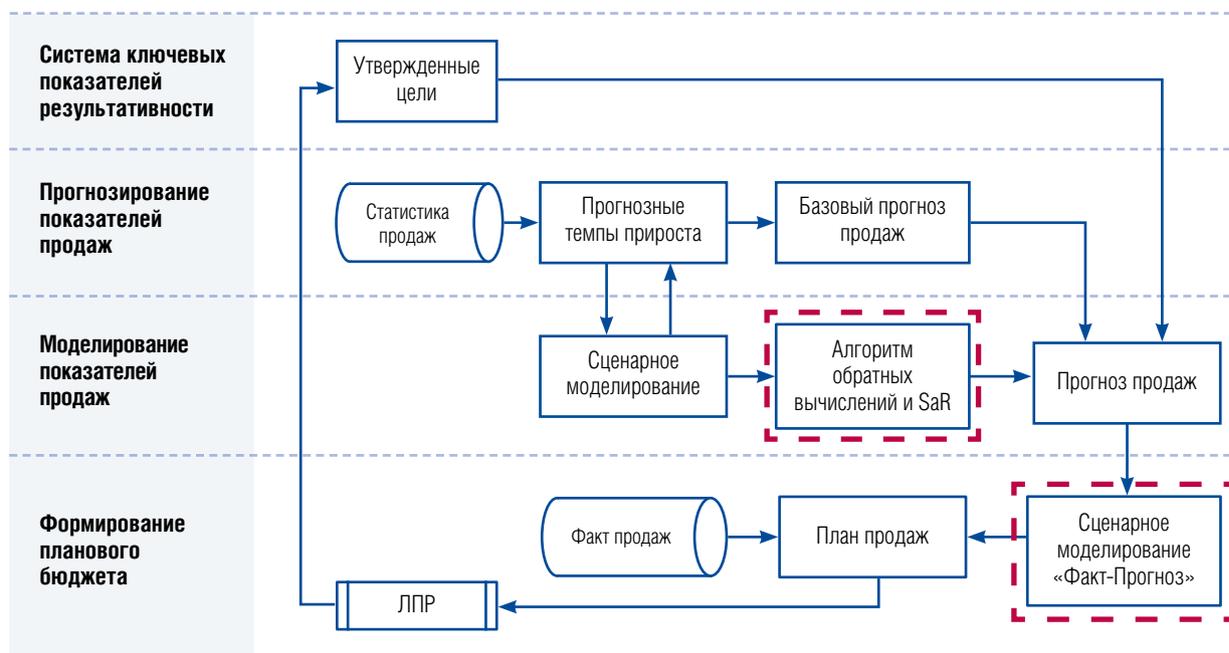


Рис. 1. Схема совершенствования типового процесса планирования.

- 1) использование экспертных оценок как на этапе формирования плановых показателей, так и на этапе согласования и формирования утвержденного планового бюджета;
- 2) из первого пункта вытекает отсутствие математической формализации задач принятия решений относительно управленческих воздействий на целевые показатели путем корректировки частных показателей;
- 3) отсутствие гибкости в части анализа прогнозов ответственных лиц на этапе формирования утвержденного планового бюджета, поскольку сценарный анализ не обеспечивает достаточных возможностей агрегации различных оценок, видений и прогнозов по различным направлениям продаж, товарным номенклатурам.

В литературе отмечается, что улучшение качества типового процесса в целом необходимо осуществлять посредством информационно-аналитической поддержки в каждой зоне процесса, обеспечивая оркестровку эволюционирующих методов информационной аналитики при согласовании управления корпоративной результативностью с управлением бизнес-процессами [2, 19–21].

Обеспечение системного улучшения качества в каждой зоне процесса достигается за счет различных факторов. Согласно работе С.Н. Брускина [7], совместное использование трендовых, информационных и сценарных методов прогнозного моделирования способно улучшить качество процесса планирования. В зоне планирования – за счет формирования статистики продаж на основе факта и пересчета плановых показателей приростным методом с учетом корректировки скользящего тренда экспертным путем. В зоне прогнозирования – за счет сценарного анализа средствами OLAP. В работе О.В. Китовой [6] представлена попытка преодолеть субъективность и формализовать процедуру выбора многочисленных вариантов стратегического развития компании по различным факторам (KPI, стратегические инициативы, распределение инвестиций) путем разработки модели целевого планирования и алгоритма выбора оптимального решения. Алгоритм выбора альтернатив предполагает проведение анализа акционерной стоимости компании и сравнительной оценки с использованием метода анализа иерархий. В работе М.И. Никишовой [9] описывается применение технологий искусственного интеллекта для преодоления информационной асимметрии при принятии реше-

ний советом директоров. Несмотря на то, что результаты направлены на стратегический уровень управления, их изучение полезно для реализации решений тактического уровня, поскольку автор предлагает подход аналитической поддержки процесса выработки управленческих воздействий, выполняемый параллельно основному процессу, не заменяющий решающей роли совета директоров. Применимость концепции обратных вычислений в задаче управления корпоративной результативностью рассматривал Б.Е. Одинцов [8]. При этом определение весов частных показателей, формирующих интегральный ключевой показатель, в работе автора осуществляется посредством экспертных методов.

Работы отдельных зарубежных авторов направлены на повышение качества процесса планирования, в первую очередь, его оперативности, за счет разработки комплексных моделей и систем, объединяющих в себе бизнес- и технический аспекты процесса. Так, Maryška и Doucek [1] описывают модель REMONA, преимущества которой заключаются в эталонной модели проектирования СPM и BI, определения методологий и моделей, сопровождающих систему, определения системы показателей, измерений, внедрения решения в компании. Prilianti и Hikmat [22] разработали платформу Integrated performance management system (IPMS), которая способна транслировать видение компании на более низкий уровень управления, повышая производительность компании, за счет определения взаимосвязей между переменными, методов оценки и приоритетов улучшения.

Кроме того, в зарубежной литературе представлены примеры аналитической поддержки процесса планирования с применением искусственного интеллекта. Zhang [4] в своей работе применил методы анализа главных компонент и анализа искусственной нейронной сети для построения модели RMT-VP, предназначенной для прогнозирования показателей результативности компаний. В качестве входных параметров используются различные финансовые и нефинансовые показатели, характеризующие результативность деятельности компаний, а соответствующие веса показателей определяются методом анализа иерархий. Magi и др. [5] показывают, как ключевые показатели могут эффективно определяться с помощью линейных ограничений, а затем как управленческие воздействия могут вычисляться методами математического программирования.

Несмотря на значительное продвижение в исследованиях технологий CRM и поддерживающих их методов, литература по-прежнему имеет открытые ниши для изучения и совершенствования процесса планирования. Рассматриваемые в литературе случаи использования искусственного интеллекта в процессе планирования основаны на модели «черного ящика», что ограничивает их применимость для бизнес-задач, которые требуют прозрачности, ясности и контролируемости со стороны менеджмента решений. Полезность машинного обучения зависит от многих факторов, которые еще недостаточно изучены, таких как тип задачи, стадия принятия решения, используемая технология [23]. При этом фрагментарность рассмотрения различных методов идет в разрез с тезисами авторов поисковых и обзорных работ о необходимости согласования методов в их взаимодействии и с особенностями процесса управления. В результате анализа научных работ и практического опыта компаний можно сделать вывод, что перед учеными стоит проблема создания эффективных подходов к практическому применению современных аналитических методов для повышения качества принятия управленческих решений в бизнесе.

Процедура согласования планов на предмет соответствия стратегии компании существующим возможностям с учетом внутренних и внешних ограничений является трудоемкой, но при этом необходимой задачей. Эта процедура выполняется путем пошаговой корректировки целевых значений показателей. На каждом этапе корректировки текущие значения стратегических показателей рассматриваются и анализируются руководством компании. По результатам анализа руководство определяет оперативные цели, которые являются компромиссом между фактическим положением дел структурного подразделения и стратегическим планом компании. Оперативные цели служат отправной точкой для расчета плановых ключевых показателей на будущие периоды. Важно отметить, что этот расчет выполняется с учетом ограничений, установленных различными структурными подразделениями, которые принимают участие в процессе составления бюджета. Например, отдел прогнозирования устанавливает ограничение целевой рентабельности на основании прогноза, отдел продаж устанавливает иное ограничение показателя, учитывая объемы ресурсов компании. Таким образом бюджет продаж дополняется элементами: целевое значение, прогнозное значение, минимальное и

максимальное значения [24]. Указанные элементы используются в качестве ограничений при осуществлении обратных вычислений значений плановых ключевых показателей.

Ввиду исключительной важности этапа планирования, задачи, связанные с развитием теоретико-методологической базы процедуры согласования планов, требуют повышенного внимания со стороны исследователей и практиков. Поэтому данное исследование направлено на разработку аналитического подхода к осуществлению процедуры согласования планов в процессе корпоративного планирования показателей, который позволит расширить использование возможностей CRM-системы и обеспечить математическую формализацию задачи выработки управленческих воздействий при корректировке значений плановых ключевых показателей.

Основываясь на результатах предшествующих авторов, в данной работе предлагается дополнить типовой процесс планирования блоками «Алгоритм обратных вычислений и SaR», позволяющий формализовать и автоматизировать процедуру формирования управленческих воздействий на частные плановые показатели, и «Сценарное моделирование Факт–Прогноз», обеспечивающий расширенные возможности, относительно сценарного моделирования, многомерного анализа и моделирования планов ПЭУ в параллель основной модели бюджетирования.

Преимущества предлагаемого подхода заключаются в повышении качества процесса планирования по всем приоритетным критериям качества: оперативность, точность, адаптивность. Оперативность достигается за счет сокращения времени и трудозатрат ПЭУ на осуществление процедуры выработки управленческих воздействий, так как менеджер в данном случае имеет поддержку в виде алгоритма расчетов, продвинутого инструментария компьютерного моделирования планов. Кроме того, сокращается количество последующих формированию бюджета корректировок плановых значений со стороны подразделений компании. Повышение точности планирования обеспечивается методичностью и системностью реализации процесса, которые предполагают взаимодействие разработанных аналитических блоков, способствующее эффективному согласованию стратегических целей и действий по их достижению. Адаптивность планирования повышается за счет сокращения времени на корректировки данных для новых шагов планирования и увеличения скорости адап-

тации системы к наблюдаемым изменениям внешней среды, поскольку в алгоритм заложено условие аппроксимации поведения ЛПР.

3. Модель процесса бюджетирования

3.1. Разработка инструментария целевого управления

Используя методы и инструменты моделирования бизнес-процессов, были построены диаграммы, отражающие суть предлагаемого подхода. Диаграммы, представленные ниже, описывают процесс, так называемой, целереализации, которая предполагает разработку инструментария (система целей и показателей, процедуры формирования предписаний, правила корректировки показателей) и технологии (цепь процедур и операций) целевого управления.

Первая диаграмма (рис. 2) отражает ключевые изменения в предлагаемой модификации модели процесса бюджетирования относительно типового решения, которые заключаются в применении методологии SaR и теории обратных вычислений для формализации и автоматизации процедуры

принятия управленческих решений в процессе бюджетирования.

Этап процесса, предваряющий алгоритм вычислений в модели – построение иерархических структур целей. Работа по формированию системы «цели-показатели» выполняется на стороне бизнес-аналитиков. Исходя из практики планирования, построение структур целей, охватывающих различные стороны деятельности компании, и впоследствии воспроизведение их показателей, требуют применения более общих структур (сетевые структуры), не обладающих свойством строгой соподчиненности узлов. Однако, для формальной обработки подобные структуры малоприспособны, поскольку порождают значительные трудности и неоднозначность в интерпретации результатов. Поэтому при отсутствии явных ограничений сетевые структуры «целей-показателей» преобразовываются в иерархические структуры. Процедура формальной обработки иерархической структуры поддерживается средствами программного обеспечения СРМ-системы, в которой измерения показателей представляются в виде иерархий.

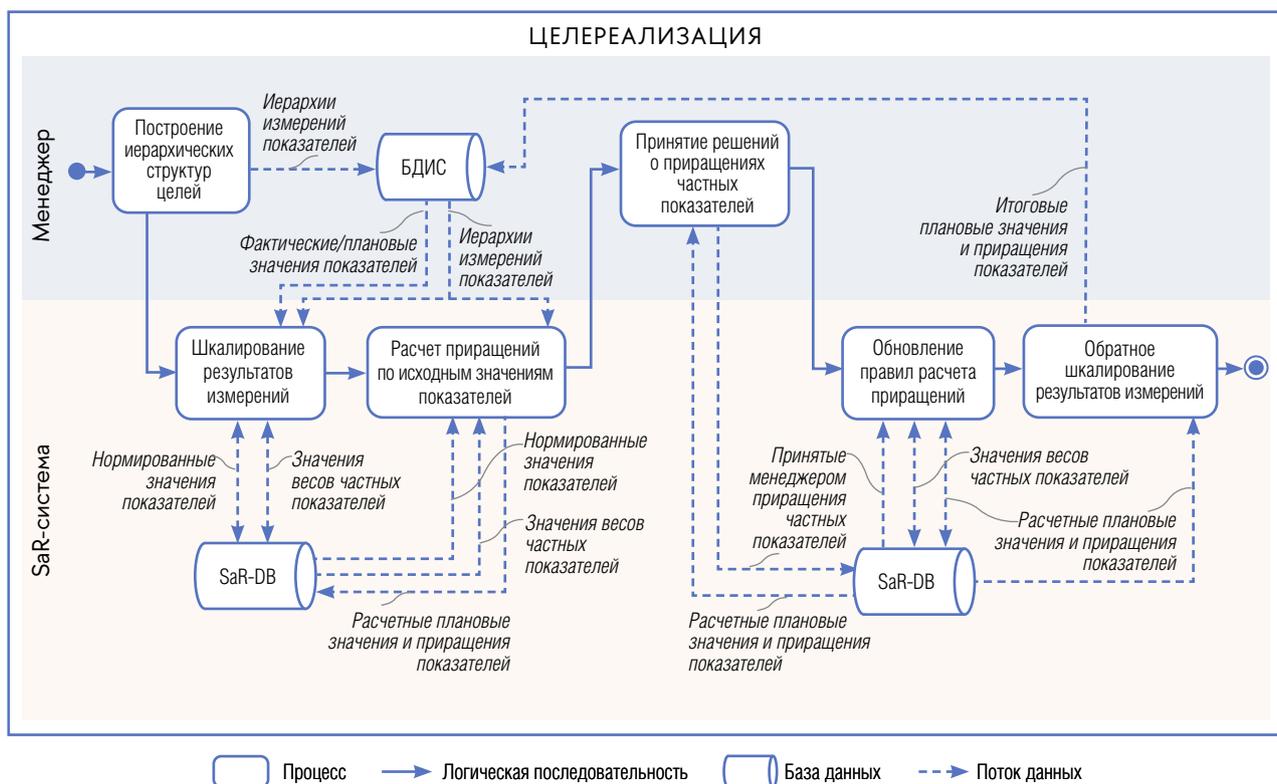


Рис. 2. Схема работы алгоритма обратных вычислений и SaR.

Первым шагом в алгоритме формальной обработки показателей является шкалирование результатов измерений. Различия в шкалах измерений отдельных частных показателей, определяющих уровень достижимости целевых показателей, накладывают ограничения на их совместную обработку. Исходя из этого в модели решается задача приведения отдельных показателей к единой шкале измерения и формирования общего интегрального показателя.

Процедура шкалирования включает предварительное нормирование исходных данных и вычисление индивидуальных весов частных показателей. Исходные данные передаются из базы данных информационной системы, а результаты вычислений хранятся в отдельном хранилище данных, связанном с системой SaR.

3.2. Математическая основа обратных вычислений и SaR

На начальном этапе вычислений значения совместно обрабатываемых частных показателей приводятся к единой шкале измерения, чтобы получить целевой безразмерный интегральный показатель. В качестве интервала нормирования берется диапазон от 0 до 1 включительно для упрощения обработки данных. Для расчета интегрального показателя используется аддитивная форма свертки, которая представляет собой сумму взвешенных значений однонаправленных частных показателей. Приведение значений частных показателей к единой шкале осуществляется методом линейного нормирования данных.

Значения частных показателей в расчете интегрального показателя должны быть однонаправленными: либо увеличиваться, либо уменьшаться. Кроме того, значения конкретного частного показателя должны равномерно заполнять интервал, определенный эмпирическим размахом его значений. В случае, если данные содержат выбросы, нормирование осуществляется на основе расчетов среднего и дисперсии.

Согласно подходу, каждый частный показатель описывается вещественным числом $h_i(x, u) \in [0, 1]$, определяющим значение действия u в состоянии x . Массив чисел $h(x, u)$ может состоять из q возможных показателей, так что $h(x, u) = [h_1(x, u), \dots, h_q(x, u)]$. При этом, обозначенный массив индуцирует лишь частичный порядок на управляющие воздействия в случае нескольких показателей. Общий порядок связывается с частным

порядком через действия менеджера, что равносильно ассоциированию весов относительной важности показателей.

Следующим шагом выполняется операция переоценки весов частных показателей в расчете интегрального целевого показателя. Поскольку процесс вычисления весов предполагается циклическим и непрерывным, на данном шаге учитываются решения, принятые в предыдущем цикле, а в нулевой точке система использует заранее определенные значения весов с равномерным распределением. Таким образом, в случае отклонения предложенных SaR управляющих воздействий, система пересматривает веса частных показателей на основе фактически выбранных менеджером управляющих воздействий. Затем на основе этих значений система вычисляет новые веса, которые будут использоваться в новой итерации ($t + 1$).

Начиная с итерации t задается скользящее окно длиной k , такое что $t - k + 1 \geq 0$. Значения весов $w(t + 1)$ зависят от значений переменной состояния системы на итерации τ , обозначается как $x(\tau)$, значений оптимальных управляющих воздействий, вычисленных системой на итерации τ , обозначается как $u(\tau)$, значений управляющих воздействий, фактически выбранных менеджером на итерации τ , после отказа от $u(\tau)$, обозначается как $u'(\tau)$, для $\tau \in [t - k + 1, t] = \{t - k + 1, \dots, t\}$.

Для каждой прошедшей итерации τ действия, фактически выбранные менеджером в итерации τ ($u'(\tau)$), должны быть предпочтительнее действий, вычисленных системой ($u(\tau)$), в зависимости от критерия оптимизации. Это условие обеспечивает соответствие новых весов знаниям, накопленным системой в течение скользящего окна. Математически данное условие формулируется следующим образом (рассматривается случай максимизации интегрального показателя):

$$w(t + 1)(h(x(\tau), u'(\tau)) - h(x(\tau), u(\tau))) \geq \delta \quad (1)$$

где значение δ должно быть малым, но не нулевым.

Дополнительное ограничение вводится на разницу между новыми значениями весов $w(t + 1)$ и значениями весов на текущей итерации $w(t)$. Разница между этими величинами должна быть минимальна, чтобы избежать колебаний весов, нарушающих сходимость алгоритма. Математически данное условие формулируется следующим образом:

$$W(t) = \sum_{\tau=t-k+1}^t \sum_{i=1}^q |w_i(t+1) - w_i(\tau)|. \quad (2)$$

Перечисленные особенности процесса взвешивания частных показателей могут быть определены в терминах линейных ограничений в задаче МИЛР. Оптимизационная задача, за исключением дополнительных условий конкретного интегрального показателя, в таком случае будет иметь вид:

$$Z(\delta, w) = \delta + W(t) \rightarrow \min_{\delta, w(t+1)}, \quad (3)$$

при условии:

$$\begin{cases} w(t+1)(h(x(\tau), u'(\tau)) - h(x(\tau), u(\tau))) - \delta \geq 0, \\ \delta \geq 0,001, \\ AbsVal(W(t), w_i(t+1), w_i(\tau)), \end{cases}$$

где δ , $w(t+1)$ – решающие переменные, $\tau \in [t-k+1, t]$, $i \in [1, q]$.

После вычисления весов частных показателей, формирующих функцию целевого интегрального показателя, наступает фаза расчета приращений аргументов функции с помощью процедуры обратных вычислений.

Целевой показатель эффективности задается в виде функции $y = f(x_1(w_1), \dots, x_q(w_q))$. Функция использует весовые коэффициенты (приоритетности), полученные на предыдущем этапе, их сумма всегда равна единице. Целевая функция и ее аргументы могут требовать положительного прироста или отрицательного приращения в зависимости от целевых установок. От характера изменения функции зависит знак приращения в целевой функции в системе уравнений.

В качестве метода обратных вычислений используется метод решения с помощью индивидуальных коэффициентов. Согласно указанному методу условия менеджера по воздействию на величины частных показателей результативности выражаются путем указания дополнительных коэффициентов (k_q) при определении искомым аргументов целевой функции:

$$x_q + \Delta x_q = k_q x_q. \quad (4)$$

От знака приращения аргументов x_q зависят значения индивидуальных коэффициентов, которые вычисляются для каждого аргумента функции: если прирост положительный, индивидуальный коэффициент умножается на свой аргумент ($k_q x_q$), если отрицательный – аргумент делится на него $\left(\frac{x_q}{k_q}\right)$.

Количество дополнительных уравнений в системе зависит от количества аргументов функции: количеству аргументов функции q будет соответствовать количество дополнительных уравнений $q - 1$. Ограничения на значения индивидуальных коэффициентов k_q устанавливаются на основе их семантики.

Таким образом, задача вычисления приращений частных показателей на основе известного приращения целевого показателя в случае с тремя аргументами и положительным приростом как целевой функции, так и аргументов, описывается системой уравнений:

$$\begin{cases} y + \Delta y = f(k_1 x_1, k_2 x_2, k_3 x_3), \\ \frac{k_1 x_1 - x_1}{k_2 x_2 - x_2 + k_3 x_3 - x_3} = \frac{w_1}{w_2 + w_3}, \\ \frac{k_2 x_2 - x_2}{k_1 x_1 - x_1 + k_3 x_3 - x_3} = \frac{w_2}{w_1 + w_3}. \end{cases} \quad (5)$$

Решив данную систему уравнений относительно k_1 , k_2 и k_3 , можно получить значения приращений аргументов целевой функции x_1 , x_2 и x_3 . При этом результаты решения должны удовлетворять условию: $k_1, k_2, k_3 \geq 1$.

На последующих шагах алгоритма целереализации менеджер получает информацию о рассчитанных приращениях частных показателей и принимает решение либо инициировать предложенное системой действие, либо инициировать иное действие. Принятое менеджером решение соответствует оптимальному решению на текущей итерации и будет использовано системой в новой итерации вычислений при обновлении весов. Поскольку предварительно вычислениям осуществлялась нормализация данных, следующий шаг алгоритма предусматривает процедуру обратного шкалирования с помощью алгоритма обратного шкального преобразователя.

Индикаторы деятельности компании являются факторами, на основе которых формируются стратегические показатели деятельности компании, определяемые руководством. Рост одних индикаторов может приводить к росту стратегических показателей, а других к их уменьшению, поэтому используя динамическое программирование можно получить сбалансированную систему стратегических показателей развития компании, удовлетворяющую требованиям руководства, изменяя значения индикаторов деятельности компании. Как

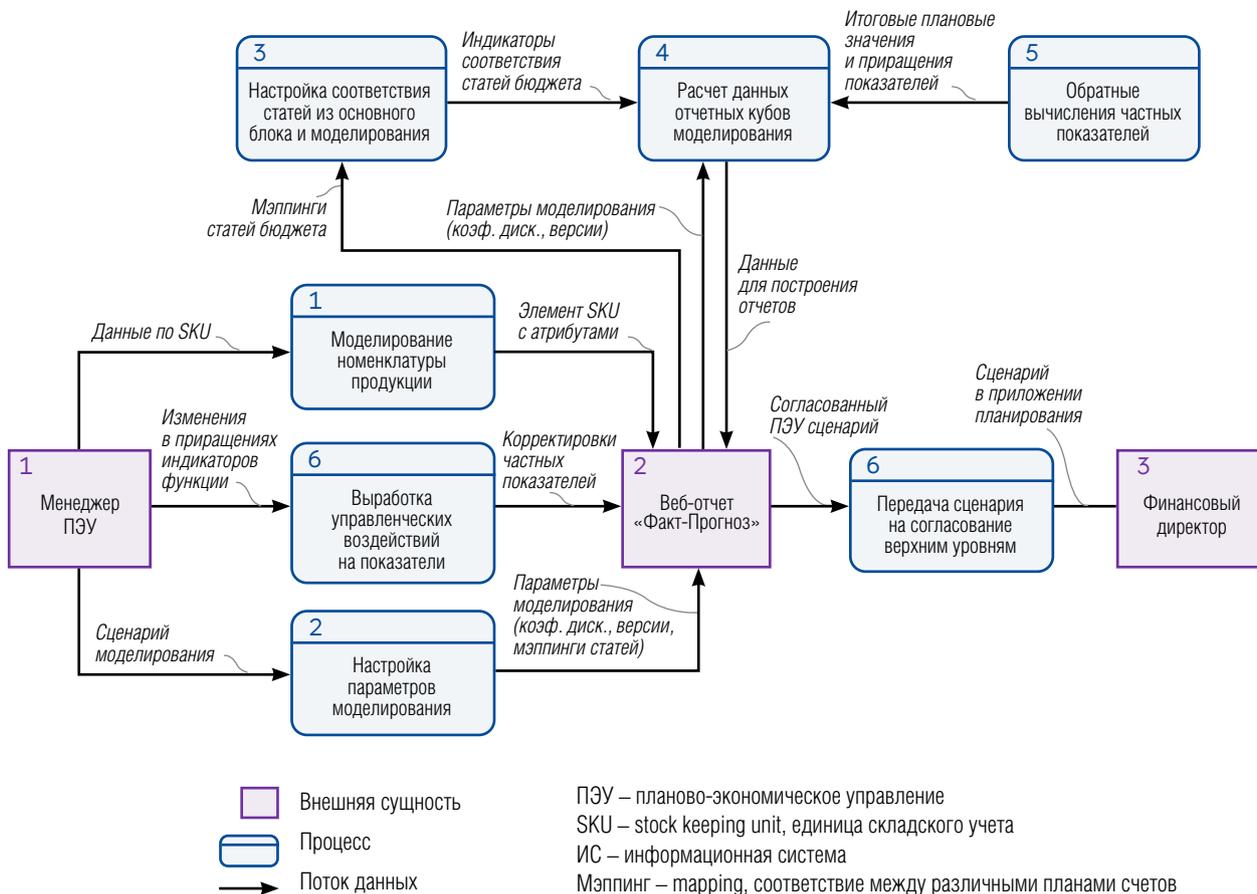


Рис. 3. Схема потоков данных блока сценарного моделирования «Факт–Прогноз».

результат, на основе стратегических показателей деятельности компании можно определить ее корпоративную результативность как свертку этих показателей.

3.3. Разработка технологии целевого управления

Вторая часть процесса целереализации – разработка технологии целевого управления, описана в контексте информационного обмена в системе управления корпоративной результативностью. Для этого используется диаграмма потоков данных, которая демонстрирует потоки данных между отдельными операциями специализированного блока сценарного моделирования «Факт–Прогноз», включенного в основную модель деятельности компании (рис. 3). Указанный блок позволяет ПЭУ параллельно основному процессу бюджетирования осуществлять выработку управленческих воздей-

ствий на основе реальных данных от ЦФО, формируя, в качестве результата, сценарии бюджетов.

В качестве начальной точки процесса моделирования бюджета на следующий плановый период рассматриваются действия менеджера ПЭУ по анализу и формированию сценария бюджета в интерактивной отчетной форме «Факт–Прогноз» после ввода в систему фактических и плановых данных по отдельным ЦФО. Обозначенные действия менеджера включают: моделирование номенклатуры и соответствующих характеристик продукции, которую необходимо учесть в сценарии; настройка параметров моделирования – указание соответствующих значений коэффициента дисконтирования, выбор рабочей версии, указание соответствия между различными планами счетов, которые необходимо анализировать и отражать в отчетности.

Как по мере обновления данных в отчетной форме, так и при непосредственном запросе менедже-

ра (запуск функции), правила в соответствующих многомерных кубах системы осуществляют расчеты значений показателей и выводят в отчетную форму. Процедура обратных вычислений по обновлению правил расчета для запуска требует решения со стороны менеджера. Процесс выработки управленческих воздействий на показатели означает выбор между значениями корректировок, предложенными системой, и ручным вводом корректировок.

На выходе процесса формируется план, согласованный менеджером ПЭУ, и передается процедурами системы на верхний уровень согласования – финансовому директору.

Заключение

В работе представлена усовершенствованная модель процесса целевого бюджетирования, которая обеспечивает автоматизированное формирование управленческих воздействий отдела бюджетирования и руководства подразделений в направлении достижения стратегических целей. В отличие от результатов, полученных в предыдущих работах, помимо параллельной оценки состояния компании, формирования управленческих воздействий с помощью искусственного интеллекта и скользящего планирования, модель процесса бюджетирования в данном исследовании включает блок моделирования и направлена на снижение частоты корректировок оперативных целей, устанавливаемых в начале планового периода.

Второй аспект теоретического вклада работы – разработка методической основы обратных вычислений при поддержке методологии SaR в процессе бюджетирования при управлении корпоративной результативностью. Применение обратных вычислений обеспечивает математическую постановку задачи вычисления индикаторов плановых ключевых значений, а система SaR позволяет дополнить мате-

матическую постановку весовыми коэффициентами первичных показателей эффективности, рассчитанными алгоритмически на основе управленческих решений менеджера, вместо экспертной оценки.

Применение предложенного подхода на практике позволяет значительно снизить величину транзакционных издержек, возникающих на этапе многочисленных корректировок версий, формируемых ЦФО, и сценариев, формируемых ПЭУ, из неточностей планирования и трудозатрат по их выявлению и исправлению. Кроме того, разработанный подход обеспечивает повышение реального качества оперативного планирования по наиболее приоритетным критериям оперативности, точности и адаптивности планирования за счет системности и методичности составления бюджетов с привлечением современных информационных технологий.

Направление дальнейших исследований предполагает интеграцию разработанной информационной модели бюджетирования в процесс определения комплексного индикатора, характеризующего корпоративную результативность компании, с использованием метода динамического программирования для расчета стратегических показателей деятельности компании на основе свертки индикаторов.

Кроме того, в качестве направления дальнейших исследований по изучаемой проблематике рассматривается проектирование информационной инфраструктуры системы управления корпоративной результативностью для повышения эффективности выработки управленческих воздействий при реализации стратегических целей компании. Основываясь на методологии Design science research, блоки аналитической поддержки рассматриваются как артефакт, который описывается путем проектирования информационно-логических схем и оценивается посредством экспериментальной апробации подхода в рабочей среде. ■

Литература

1. Maryska M., Doucek P. REMONA model for improving quality of corporate informatics performance management: How to cut costs in corporate informatics // *Quality Innovation Prosperity*. 2017. Vol. 21. No. 3. P. 15–35. <https://doi.org/10.12776/QIP.V21I3.939>
2. Richards G., Yeoh W., Chong A.Y.L., Popovič A. Business intelligence effectiveness and corporate performance management: An empirical analysis // *Journal of Computer Information Systems*. 2019. Vol. 59. No. 2. P. 188–196. <https://doi.org/10.1080/08874417.2017.1334244>
3. Bourne M., Neely A., Mills J., Platts K. Implementing performance measurement systems: A literature review // *International Journal of Business Performance Management*. 2003. Vol. 5. No. 1. P. 1–24. <https://doi.org/10.1504/IJBPM.2003.002097>
4. Zhang J. A neural network model for business performance management based on random matrix theory // *Mathematical Problems in Engineering*. 2022. Vol. 2022. Article ID 9170666. <https://doi.org/10.1155/2022/9170666>

5. Mari F., Massini A., Melatti I., Tronci E. A constraint optimization–based sense and response system for interactive business performance management // *Applied Artificial Intelligence*. 2021. Vol. 35. No. 5. P. 353–372. <https://doi.org/10.1080/08839514.2020.1843833>
6. Китова О.В. Концепции и информационная инфраструктура обеспечения управления результативностью маркетинга: теория и методология. Дис. ... д-ра экон. наук. С.-Петербург., 2012.
7. Брускин С.Н. Информационно-аналитическая система поддержки принятия решений в области планирования сбытовой деятельности корпорации. Дис. ... канд. экон. наук. М., 2016.
8. Одинцов Б.Е. Информационные системы управления эффективностью бизнеса: учебник и практикум для вузов. Москва: Юрайт, 2023.
9. Никишова М.И. Применение технологий искусственного интеллекта в системе корпоративного управления. Дис. ... канд. экон. наук. М., 2021.
10. Neely A., Gregory M., Platts K. Performance measurement system design—a literature review and research agenda // *International Journal of Operations and Production Management*. 1995. Vol. 15. No. 4. P. 80–116. <https://doi.org/10.1108/01443570510633639>
11. Kaplan R.S., Norton D.P. (1996) *The balanced scorecard. Translating strategy into action*. Boston, MA: Harvard Business School Press, 1996.
12. Kaplan R.S., Norton D.P. Using the balanced scorecard as a strategic management system // *Harvard Business Review*. 1996. P. 75–85.
13. Neely A.D. et al. *Getting the measure of your business*. London: Findlay, 1996.
14. Dixon J.R., Nanni A.J., Vollmann T.E. *The new performance challenge: Measuring operations for world-class competition*. Homewood, IL: Business One Irwin, 1990.
15. Krause O., Mertins K. Performance management // *Global Production Management: IFIP WG5.7 International Conference on Advances in Production Management Systems*, Berlin, Germany, September 6–10, 1999 (eds. K. Mertins, O. Krause, B. Schallock). P. 243–251.
16. Kaplan R.S., Norton D.P. Putting the balanced scorecard to work // *Harvard Business Review*. 1993. P.134–147. <https://doi.org/10.1016/B978-0-7506-7009-8.50023-9>
17. Рыбалко О.А., Шалаева Л.В. Стратегическое планирование и бюджетирование как базовые элементы современной системы управления // *Международный бухгалтерский учет*. 2012. №. 28. С. 25–38.
18. Dresner H. *The performance management revolution: Business results through insight and action*. Hoboken, NJ: Wiley, 2007.
19. Taticchi P., Tonelli F., Cagnazzo L. Performance measurement and management: A literature review and a research agenda // *Measuring Business Excellence*. 2010. Vol. 14. No. 1. P. 4–18. <https://doi.org/10.1108/13683041011027418>
20. Medeiros M.M., Maçada A.C.G., Hoppen N. The role of big data stewardship and analytics as enablers of corporate performance management // *RAM. Revista de Administração Mackenzie*. 2021. Vol. 22. <https://doi.org/10.1590/1678-6971/eram210063>
21. Jaklič J., Bosilj-Vukšić V., Mendling J., Štemberger M.I. The orchestration of corporate performance management and business process management and its effect on perceived organizational performance // *SAGE Open*. 2021. Vol. 11. No. 3. <https://doi.org/10.1177/21582440211040126>
22. Prilianti E., Hikmat M.T. Proposed corporate performance management using integrated performance management system (IPMS) at PT Pos Indonesia (Persero) // *International Journal of Engineering & Technology*. 2018. Vol. 7. No. 3. P. 71–83. <https://doi.org/10.14419/ijet.v7i3.25.17472>
23. Merkert J., Mueller M., Hubl M. A survey of the application of machine learning in decision support systems // *ECIS Completed Research Papers*. 2015. Paper 133. <https://doi.org/10.18151/7217429>
24. Дугельный А.П. Бюджетное управление предприятием. М.: Дело, 2003.

Об авторе

Ощепков Максим Евгеньевич

аспирант, департамент бизнес-информатики, Высшая школа бизнеса, Национальный исследовательский университет «Высшая школа экономики», Россия, 119049, г. Москва, ул. Шаболовка, д. 26–28;

E-mail: moshchepkov@hse.ru

ORCID: 0000-0003-0327-3285

Improving target budgeting in a corporate performance management system

Maxim E. Oshchepkov

E-mail: moshchepkov@hse.ru

Graduate School of Business, HSE University

Address: 26–28, Shabolovka Str., Moscow 119049, Russia

Abstract

Currently the portrayal of the procedure for developing management actions during the planning process in the scientific and professional community does not align with the practice of systematic and consistent plan creation supported by an informational analytical system. Alternatively, the non-formalized decision-making activity of the planner, which involves a situational expert approach, becomes a dependency in the planning process. This study is aimed at developing an analytical approach for implementing the plan reconciliation procedure in the process of corporate performance planning. This will increase the utilization of capabilities of the corporate performance management system and formalize the task of generating managerial actions by adjusting targeted budgeting values using mathematical methods. For this purpose, the standard planning process is enhanced by analytical support units, including the algorithm of inverse calculations of individual key performance indicators (KPI) and an advanced module for scenario modeling. The improved model of target budgeting process presented here delivers automated formation of management actions of the budgeting department and subdivision management, guided towards accomplishing strategic goals. The application of inverse calculations provides a mathematical formulation of the task of calculating indicators of planned key values, and the Sense and Respond (SaR) system allows you to supplement the mathematical formulation with weighting coefficients of key performance indicators calculated algorithmically, relying on the manager's decisions rather than expert evaluation. The implementation of the approach we developed will improve the quality of planning by the highest priority criteria of operability, accuracy and adaptability due to the consistency and methodology of budgeting with the use of modern information technology.

Keywords: corporate performance management, budgeting, information model, inverse calculations, mathematical programming

Citation: Oshchepkov M.E. (2024) Improving target budgeting in a corporate performance management system. *Business Informatics*, vol. 18, no. 1, pp. 22–35. DOI: 10.17323/2587-814X.2024.1.22.35

References

1. Maryska M., Doucek P. (2017) REMONA model for improving quality of corporate informatics performance management: How to cut costs in corporate informatics. *Quality Innovation Prosperity*, vol. 21, no. 3, pp. 15–35. <https://doi.org/10.12776/QIP.V21I3.939>
2. Richards G., Yeoh W., Chong A.Y.L., Popovič A. (2019) Business intelligence effectiveness and corporate performance management: An empirical analysis. *Journal of Computer Information Systems*, vol. 59, no. 2, pp. 188–196. <https://doi.org/10.1080/08874417.2017.1334244>
3. Bourne M., Neely A., Mills J., Platts K. (2003) Implementing performance measurement systems: A literature review. *International Journal of Business Performance Management*, vol. 5, no. 1, pp. 1–24. <https://doi.org/10.1504/IJBPM.2003.002097>
4. Zhang J. (2022) A neural network model for business performance management based on random matrix theory. *Mathematical Problems in Engineering*, vol. 2022, article ID 9170666. <https://doi.org/10.1155/2022/9170666>

5. Mari F., Massini A., Melatti I., Tronci E. (2021) A constraint optimization–based sense and response system for interactive business performance management. *Applied Artificial Intelligence*, vol. 35, no. 5, pp. 353–372. <https://doi.org/10.1080/08839514.2020.1843833>
6. Kitova O.V. (2012) *Concepts and information infrastructure of marketing performance management: theory and methodology*. St. Petersburg: UNECON (in Russian).
7. Bruskin S.N. (2016) *Information-analytical system of decision-making support in the field of planning sales activities of the corporation*. Moscow: MSU (in Russian).
8. Odintsov B.E. (2023) *Information systems of business performance management. Textbook and Practice for Bachelor's and Master's Degrees*. Moscow: Urait (in Russian).
9. Nikishova M.I. (2021) *Application of artificial intelligence technologies in the corporate governance system*. Moscow: Financial University (in Russian).
10. Neely A., Gregory M., Platts K. (1995) Performance measurement system design—a literature review and research agenda. *International Journal of Operations and Production Management*, vol. 15, no. 4, pp. 80–116. <https://doi.org/10.1108/01443570510633639>
11. Kaplan R.S., Norton D.P. (1996) *The balanced scorecard. Translating strategy into action*. Boston, MA: Harvard Business School Press.
12. Kaplan R.S., Norton D.P. (1996) Using the balanced scorecard as a strategic management system. *Harvard Business Review*, pp. 75–85.
13. Neely A.D. et al. (1996) *Getting the measure of your business*. London: Findlay.
14. Dixon J.R., Nanni A.J., Vollmann T.E. (1990) *The new performance challenge: Measuring operations for world-class competition*. Homewood, IL: Business One Irwin.
15. Krause O., Mertins K. (1999) Performance management. *Global Production Management: IFIP WG5.7 International Conference on Advances in Production Management Systems, Berlin, Germany, September 6–10, 1999* (eds. K. Mertins, O. Krause, B. Schallcock), pp. 243–251.
16. Kaplan R.S., Norton D.P. (1993) Putting the balanced scorecard to work. *Harvard Business Review*, pp. 134–147. <https://doi.org/10.1016/B978-0-7506-7009-8.50023-9>
17. Rybalko O.A., Shalaeva L.V. (2012) Strategic planning and budgeting as basic elements of modern management system. *International Accounting*, no. 28, pp. 25–38 (in Russian).
18. Dresner H. (2007) *The performance management revolution: business results through insight and action*. Hoboken, NJ: Wiley.
19. Taticchi P., Tonelli F., Cagnazzo L. (2010) Performance measurement and management: a literature review and a research agenda. *Measuring Business Excellence*, vol. 14, no. 1, pp. 4–18. <https://doi.org/10.1108/13683041011027418>
20. Medeiros M.M., Maçada A.C.G., Hoppen N. (2021) The role of big data stewardship and analytics as enablers of corporate performance management. *RAM. Revista de Administração Mackenzie*, vol. 22. <https://doi.org/10.1590/1678-6971/eram210063>
21. Jaklič J., Bosilj-Vukšić V., Mendling J., Štemberger M.I. (2021) The orchestration of corporate performance management and business process management and its effect on perceived organizational performance. *SAGE Open*, vol. 11, no. 3. <https://doi.org/10.1177/21582440211040126>
22. Prilianti E., Hikmat M.T. (2018) Proposed corporate performance management using integrated performance management system (IPMS) at PT Pos Indonesia (Persero). *International Journal of Engineering & Technology*, vol. 7, no 3, pp. 71–83. <https://doi.org/10.14419/ijet.v7i3.25.17472>
23. Merkert J., Mueller M., Hubl M. (2015) A Survey of the Application of Machine Learning in Decision Support Systems. *ECIS Completed Research Papers*, paper 133. <https://doi.org/10.18151/7217429>
24. Dugelny A.P. (2003) *Budgetary management of the enterprise*. Moscow: Delo (in Russian).

About the author

Maxim E. Oshchepkov

Doctoral Student, Department of Business Informatics, Graduate School of Business, HSE University, 26–28, Shabolovka St., Moscow 119049, Russia

E-mail: moshchepkov@hse.ru

ORCID: 0000-0003-0327-3285

DOI: 10.17323/2587-814X.2024.1.36.51

Агентное моделирование и оптимизация характеристик научно-производственных кластеров

Г.Л. Бекларян 

E-mail: glbeklaryan@gmail.com

Центральный экономико-математический институт, Российская академия наук
Адрес: 117418, г. Москва, Нахимовский пр-т, д. 47

Аннотация

В работе представлена разработанная агент-ориентированная имитационная модель развития научно-производственных кластеров России, реализованная на примере высокотехнологичных предприятий, расположенных в четырех наукоградах (г. Троицка, г. Обнинска, г. Пушкино и г. Протвино). Предложен новый подход к моделированию и оптимизации валового городского продукта (GMP), учитывающий влияние «гравитационного эффекта» на перераспределение трудовых ресурсов между развивающимися наукогородами и соответствующими предприятиями, объединенными в единые научно-производственные кластеры. Важным элементом такого подхода является формирование различных сценариев стратегического развития оцениваемых научно-производственных кластеров и поддержка возможности выбора наиболее предпочтительного сценария с использованием эволюционного оптимизационного алгоритма. Разработана и реализована в AnyLogic двухуровневая имитационная модель, описывающая возможные траектории развития научно-производственных кластеров с соответствующим изменением значений важнейших характеристик: численности экономически активного населения, количества научно-производственных предприятий, объема продукции выпускаемой в высокотехнологичных отраслях экономики, GMP и др. Спроектированный программный комплекс предназначен, в первую очередь, для управления научно-производственными кластерами, реализующими стратегию инновационного развития. Такой комплекс использует методы системной динамики и агентного имитационного моделирования, поддерживаемые в системе AnyLogic, генетические оптимизационные алгоритмы и ГИС-карты наукоградов и др. для реализации требуемой функциональности. Апробация программного комплекса выполнена с использованием реальных данных, опубликованных в утвержденных стратегиях развития соответствующих наукоградов. В результате проведенных численных экспериментов предложены некоторые рекомендации по развитию изучаемых научно-производственных кластеров с учетом их взаимовлияния и имеющейся ресурсной базы.

Ключевые слова: научно-производственный кластер, высокотехнологичное предприятие, наукоград, производственные характеристики, имитационное моделирование предприятий, валовый городской продукт, агентное моделирование, системная динамика, гравитационный эффект, производственная функция, AnyLogic

Цитирование: Бекларян Г.Л. Агентное моделирование и оптимизация характеристик научно-производственных кластеров // Бизнес-информатика. 2024. Т. 18. № 1. С. 36–51. DOI: 10.17323/2587-814X.2024.1.36.51

Введение

В настоящее время весьма актуальной проблемой повышения эффективности системы государственного управления, является развитие научно-производственных кластеров с использованием имеющегося кадрового потенциала, в том числе, сформированного на базе научно-исследовательских институтов (НИИ) и научно-производственных объединений (НПО), имеющих опыт применения собственных разработок в реальном секторе экономики.

Определение наилучших сценариев развития для высокотехнологичных предприятий, расположенных в наукоградах, разработка и внедрение систем поддержки принятия решений для управления научно-производственными кластерами является важнейшим направлением бизнес-информатики, нацеленным на повышение эффективности управления процессами инновационного развития.

В качестве примера, рассмотрим высокотехнологичные предприятия, расположенные в наукоградах, которые можно объединить в рамках единой концептуальной модели инновационного развития, в рамках которой достижимо получение значительного синергетического эффекта, обусловленного, во-первых, возможностями кадровой кооперации, а также использованием общей высокотехнологичной ресурсной базы (например, суперкомпьютеров, центров обработки данных, промышленных систем 3D-печати интегральных микросхем и т.д.). В частности, высокотехнологичные предприятия, расположенные в таких городах как Троицк, Обнинск, Пущино, Протвино вблизи Московской агломерации и являющиеся по сути соседними, могут быть объединены в единые научно-производственные кластеры, в частности, для производства микроэлектронной элементной базы, продукции точного приборостроения и др. При этом важно обеспечить сбалансированное развитие высокотехнологичных предприятий всех малых городов, объединенных в

соответствующие научно-производственные кластеры, чтобы избежать одностороннего перетока высококвалифицированных кадров и обеспечить гармоничное стратегическое развитие всех агентов-предприятий (НИИ, НПО и др.), как существующих, так и формирующихся в соответствующих наукоградах.

Поэтому необходима разработка системы поддержки принятия решений (программного комплекса) для управления высокотехнологичными предприятиями, расположенными в наукоградах и реализующими стратегию инновационного развития. Подобный комплекс использует методы системной динамики и агентного имитационного моделирования, поддерживаемые в системе AnyLogic, генетические оптимизационные алгоритмы и ГИС-карты научно-производственных кластеров и др. для реализации требуемой функциональности.

Первой фундаментальной работой по исследованию динамики развития города и его предприятий является книга [1], в которой Дж. Форрестер сформулировал основные принципы построения имитационных моделей, описывающих различные сценарии эволюции городской среды с учетом сложной системы имеющихся обратных связей, лаговых зависимостей, внутренних городских взаимодействий и др.

Среди современных работ в данной области следует выделить [2–4], в которых методы системной динамики используются для изучения возможностей достижения устойчивого экономического развития крупных городов. В частности, в [2] для г. Пекин с использованием системно-динамического подхода смоделированы четыре сценария возможного развития. Показано, что динамика устойчивого развития города по-разному реализуется в течение трех периодов (от роста экономики города в начале к ее замедлению и последующему росту).

Среди недостатков использования методов системной динамики следует отметить трудности

моделирования внутренних взаимодействий, процессов перераспределения трудовых ресурсов, сложности с анализом пространственного развития города и др. Для преодоления подобных трудностей целесообразно использовать методы агентного имитационного моделирования, позволяющие смоделировать взаимодействия на микроскопическом уровне с учетом индивидуальных систем принятия решений каждого экономического агента [5, 6].

Среди важных работ в данной области следует выделить [7–10]. В частности, в [7] представлена укрупненная агент-ориентированная имитационная модель миграционных потоков стран Европейского Союза и описано влияние «гравитационного эффекта» на межстрановое перераспределение людских потоков. Подобный эффект возникает из-за факторов, влияющих на миграционное поведение людей, в частности, при наличии существенной разницы в соотношении заработной платы, размера рынка труда и валового внутреннего продукта (ВВП) на душу населения между данной страной (регионом) и соседними странами (регионами). Например, относительно низкий уровень заработной платы в сочетании с малым количеством имеющихся рабочих мест, как правило, приводит к оттоку населения в более благоприятные регионы. Также «гравитационный эффект» влияет и на внутреннюю миграцию населения между городами, когда население (и соответствующие трудовые ресурсы) перемещается из менее развитых агломераций в экономически более развитые. В [8] представлена агентная модель популяционной динамики двух взаимодействующих сообществ: мигрантов и коренных жителей. В подобной модели коренные жители нацелены на поиск высокотехнологичных рабочих мест, а мигранты – низкотехнологичных, что также актуально и для предприятий наукоградов, активно привлекающих внешние трудовые ресурсы. Агентное моделирование можно комбинировать с другими методами, например, системной динамикой, дискретно-событийном моделированием и др. для создания цифровых двойников сложных социально-экономических систем [9, 10]. Разрабатываемые имитационные модели могут быть агрегированы по целевым функционалам и ограничениям с эволюционными оптимизационными алгоритмами, в частности, генетическими алгоритмами [11, 12] для оптимизации характеристик подобных крупномасштабных систем. Агентные модели

территориального развития региональных и городских агломераций могут быть нацелены на решение важных экологических проблем, таких как снижение концентрации вредных выбросов в социально-значимых городских районах (например, вблизи детских садов и школ) за счет озеленения [13], сокращения выбросов со стороны промышленных предприятий за счет их экологической модернизации [14], экологической трансформации городской среды [15]. Для моделирования движения людских потоков и трудовых ресурсов в городской среде целесообразно использование феноменологического подхода, ранее предложенного в [16], который позволяет учесть различные сценарии взаимодействия различных агентов друг с другом, смоделировать поведение толпы и т.д. Также следует отметить работы [17–21], посвященные как методологическим вопросам [17], так и конкретным тематическим исследованиям, выполненным для изучения возможностей развития городской архитектуры [19], улучшения трафика [20], повышения отдачи в отдельных отраслях городской экономики [21].

Таким образом, комбинированное использование различных методов имитационного моделирования, эвристических оптимизационных алгоритмов и др., в частности, поддерживаемых в системе AnyLogic, позволяет спроектировать систему поддержки принятия решений для устойчивого экономического развития научно-производственных кластеров и соответствующих наукоградов.

Цель данной статьи состоит в разработке экономико-математического и компьютерного инструментария для исследования динамики развития отдельных научно-производственных кластеров РФ при различных сценарных условиях, в рамках которых обеспечивается сбалансированное развитие всех взаимодействующих экономических агентов-предприятий (НИИ, НПО и др.).

1. Имитационная модель развития научно-производственных кластеров

В данном разделе представлена разработанная имитационная модель развития научно-производственных кластеров с реализацией на примере высокотехнологичных предприятий, расположенных в наукоградах Троицка, Обнинска, Пущино и Протвино. Предлагаемая агент-ориентированная модель состоит из нескольких взаимосвязанных уровней (*рис. 1*).

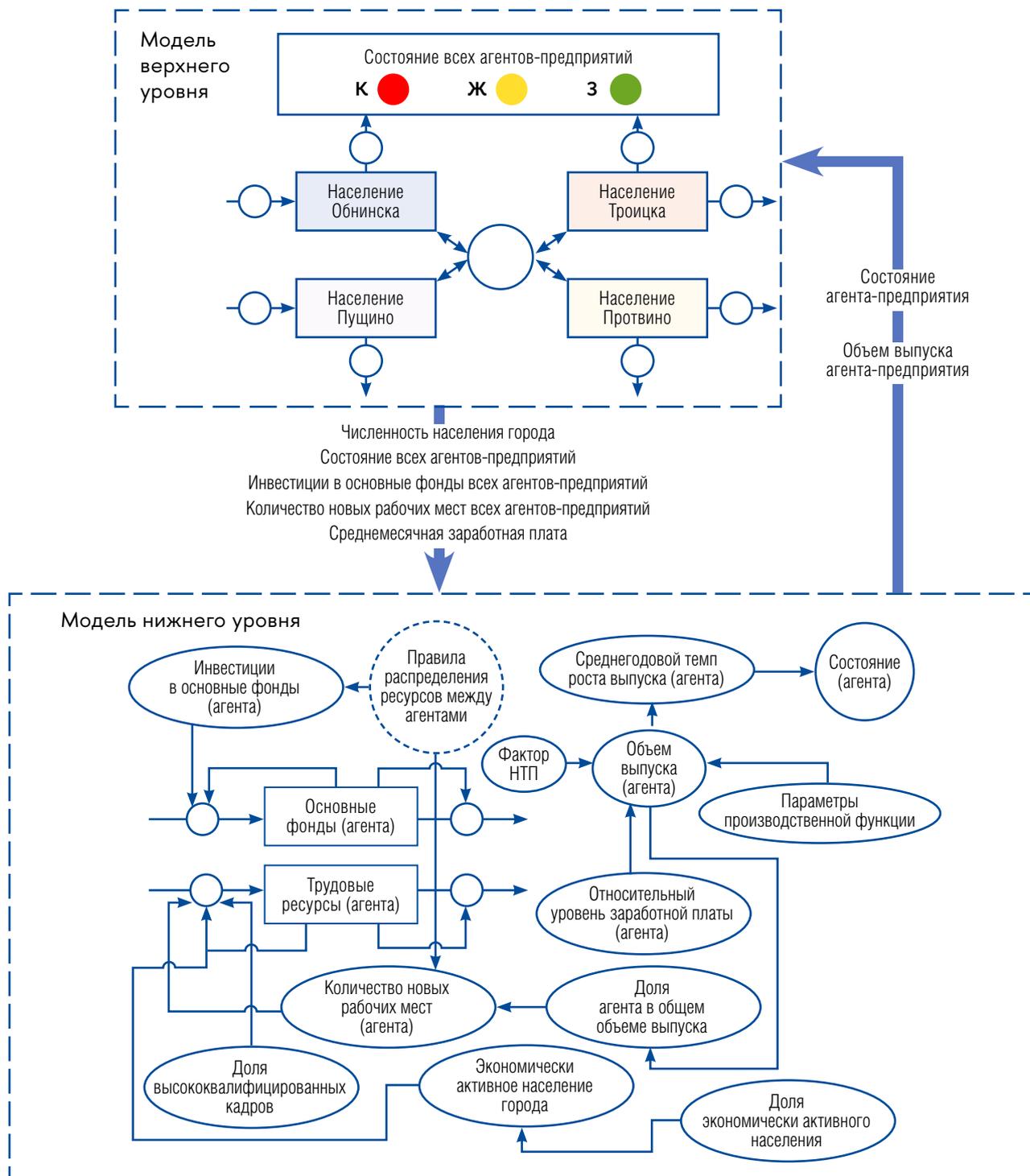


Рис. 1. Укрупненная архитектура имитационной модели научно-производственных кластеров.

Модель верхнего уровня – среда, в которой функционируют агенты-предприятия, относящиеся к различным городам, между которыми реализуется потоковое взаимодействие, в частности, перераспределение (внутренняя миграция) населения с

соответствующим изменением трудовых ресурсов, которые могут быть задействованы в высокотехнологическом секторе экономики (в НИИ, НПО и др.) в случае формирования новых рабочих мест (рис. 1). Подобное потоковое взаимодействие осуществля-

ется из-за влияния «гравитационного» эффекта [7], обусловленного различиями в уровне заработной платы, размера рынка труда и ВВП на душу населения в соответствующих наукоградах.

Модель нижнего уровня обеспечивает реализацию логики поведения агентов-предприятий, характеристики которых описываются с использованием методов системной динамики (рис. 1). Такие внутренние модели представляют собой системы конечно-разностные уравнений со своими ресурсными и целевыми характеристиками [11, 22]. При этом, для определения значений отдельных параметров требуется интеграция с верхним уровнем имитационной модели. В частности, основные управляющие параметры модели, значения которых должны быть одинаковыми для всех агентов, задаются на уровне среды. При этом, отдельные характеристики, вычисляемые на уровне агентов, используются в модели верхнего уровня, например, для вычисления GMP формируемого всеми предприятиями.

Для каждого агента-предприятия заданы три возможных состояния:

- ◆ первое, соответствующее низким темпам экономического роста и не превышающим заданное нижнее пороговое значение, выделяется «красным»;
- ◆ второе, соответствующее средним темпам экономического роста, значения которых находятся в интервале между нижним и верхним пороговыми значениями, выделяется «желтым»;
- ◆ третье, соответствующее высоким темпам экономического роста, значения которых превышают верхнее пороговое значение, выделяется «зеленым».

На темпы экономического роста агентов-предприятий (динамику GMP) влияют различные факторы, среди которых наиболее важными являются инвестиции в основной капитал, количество создаваемых рабочих мест, среднемесячная заработная плата и др. Также важны уже имеющиеся у предприятий производственные ресурсы (основные фонды, персонал), уровень научно-технического прогресса (НТП), способность привлечения высококвалифицированных кадров и др. В рамках разработанной модели, предлагается централизованное выделение ресурсов (инвестиций, новых рабочих мест и др.) с их последующим распределением по предприятиям каждого наукограда с учетом большей поддержки проблемных организаций, характеризующихся низкими темпами экономического роста. Подобный подход актуален, прежде всего, для государственных

предприятий, например, институтов РАН, НИИ, а также коммерческих организаций с существенным государственным участием.

Таким образом, важнейшей задачей рассматриваемой системы является минимизация общего количества проблемных («красных») предприятий при минимально необходимых инвестиционных и операционных затратах.

1.1. Модель перераспределения трудовых ресурсов

Модель верхнего уровня, представленная на рис. 1, обеспечивает реализацию логики перераспределения населения и трудовых ресурсов с учетом влияния «гравитационного эффекта». Впервые подобный эффект был описан в работах [23, 24] посвященных анализу поведения людей под воздействием социальных сил. В частности, на принятие решения о миграции в другой город оказывают множественные факторы: разница в заработной плате, размер рынка труда, ВВП на душу населения, климатические характеристики, уровень социальной обеспеченности и т.д. В большей степени «гравитационное» влияние подобной дифференциации проявляется на уровне стран, существенно отличающихся по уровню и качеству жизни. Однако, схожие факторы, способствующие миграции, действуют и в отношении городов, в том числе наукоградов, особенности которых позволяют их жителям перемещаться, оставаясь занятыми в высокотехнологичных отраслях экономики.

Пусть

- ◆ $T = \{t_1, t_2, \dots, t_{|T|}\}$ – набор моментов модельного времени (по годам), $|T|$ – горизонт стратегического планирования (10 лет);
- ◆ $I = \{i_1, i_2, \dots, i_{|I|}\}$ – набор индексов наукоградов, $|I|$ – общее количество наукоградов;
- ◆ $\{\nu_i, \eta_i\}$, $i \in I$ – коэффициенты рождаемости и смертности в i -ом наукограде;
- ◆ $\{GMP_i(t_k), P_i(t_k)\}$, $i \in I$ – валовый городской продукт и численность населения i -го наукограда в момент t_k , ($t_k \in T$);
- ◆ $\{W_i(t_k), L_i(t_k)\}$, $i \in I$ – уровень заработной платы и количество трудовых ресурсов i -го наукограда, занятых в высокотехнологичных отраслях экономики в момент t_k , ($t_k \in T$);
- ◆ $\{\tilde{w}_1, \tilde{w}_2, \tilde{w}_3\} \in [0, 1]$ – весовые коэффициенты, определяющие уровень влияния отдельных факторов на миграционные потоки, $\tilde{w}_1 + \tilde{w}_2 + \tilde{w}_3 = 1$;

♦ $\varkappa_i \in [0, 1]$ – коэффициент, определяющий уровень влияния «гравитационного эффекта» на эмиграцию в i -ом наукограде.

Темп эмиграции из i -го наукограда ($i \in I$) в соседний ξ -й наукоград ($\xi \in I, \xi \neq i$) в момент t_k , ($t_k \in T$) равен

$$E_{i\xi}(t_k) = \begin{cases} \varkappa_i \left(\frac{1}{g_{i\xi}(t_k)} - 1 \right) P_i(t_{k-1}), & \text{если } g_{i\xi}(t_k) < 1, \\ 0, & \text{если } g_{i\xi}(t_k) \geq 1, \end{cases} \quad (1)$$

где

$$\ln(g_{i\xi}(t_k)) = \tilde{w}_1 \ln\left(\frac{W_i(t_{k-1})}{W_\xi(t_{k-1})}\right) + \tilde{w}_2 \ln\left(\frac{L_i(t_{k-1})}{L_\xi(t_{k-1})}\right) + \tilde{w}_3 \ln\left(\frac{\frac{GMP_i(t_{k-1})}{P_i(t_{k-1})}}{\frac{GMP_\xi(t_{k-1})}{P_\xi(t_{k-1})}}\right). \quad (2)$$

Формулы (1)–(2) описывают влияние «гравитационного эффекта» на миграционные потоки. Выражение (2) отражает влияние таких ключевых факторов, как заработная плата $W_i(t_{k-1})$, количество занятых трудовых ресурсов $L_i(t_{k-1})$ и ВВП на душу населения

$$\frac{GMP_i(t_{k-1})}{P_i(t_{k-1})}$$

на миграционное поведение людей (т.е. темпы миграции) в i -ом наукограде ($i \in I$). Выбор таких факторов обусловлен, прежде всего, наблюдаемой разницей в их значениях применительно к рассматриваемым наукографам (т.е. г. Троицк, г. Обнинск и др.). Влияние других потенциально важных характеристик, например, различий в климате, транспортной доступности, экологии и др., не существенно из географической близости и пространственной однородности изучаемых агломераций. Использование относительных и масштабированных (логарифмированных) значений влияющих факторов позволяет выполнить их агрегирование с использованием (2) для последующего вычисления темпов миграции в соседние наукограды.

Из (1)–(2) следует, что, если данный наукоград менее привлекателен по сравнению с соседним, в котором более высокая заработная плата, большее количество рабочих мест и больше ВВП на душу населения, то формируется положительный эмиграционный поток в данный наукоград. Возмож-

на и обратная ситуация, когда данный наукоград является более привлекательным для жителей соседних городов. В этом случае будет обеспечено формирование положительных иммиграционных потоков.

Численность населения i -го наукограда ($i \in I$) с учетом эмиграции в соседние ξ -ые наукограды ($\xi \in I, \xi \neq i$), иммиграции из соседних ξ -ые наукоградов ($\xi \in I, \xi \neq i$), а также рождаемости и смертности в момент t_k , ($t_k \in T$) равна

$$P_i(t_k) = P_i(t_{k-1}) - \sum_{\xi=1}^{|I|} E_{i\xi}(t_k) + \sum_{\xi=1}^{|I|} E_{\xi i}(t_k) + v_i P_i(t_{k-1}) - \eta_i P_i(t_{k-1}). \quad (3)$$

Вместе с перераспределением численности населения между соседними наукографами, соответствующим образом реализуется перераспределение высококвалифицированной рабочей силы.

Количество трудовых ресурсов, которые могут быть задействованы на высокотехнологичных предприятиях i -го наукограда ($i \in I$) в момент t_k , ($t_k \in T$) равно

$$L_i^*(t_k) = \mu_i \omega_i P_i(t_k), \quad (4)$$

где

$\mu_i \in [0, 1]$ – доля экономически активного населения в i -ом наукограде ($i \in I$);

$\omega_i \in [0, 1]$ – доля высококвалифицированных трудовых ресурсов от экономического активного населения в i -ом наукограде ($i \in I$).

1.2. Модель поведения агентов-предприятий

Модель нижнего уровня, представленная на рис. 1 используется для вычисления динамики среднегодовых темпов роста выпуска агентов-предприятий с соответствующей оценкой состояний таких агентов. Наиболее важной характеристикой подобных предприятий является производственная функция, описывающая влияние основных фондов, трудовых ресурсов, уровня научно-технического развития и других факторов на динамику объема выпуска. Подробное описание принципов построения производственных функций в том числе типа Кобба-Дугласа, используемых в данной работе, представлено в [25, 26]. Также имеются примеры построения мультипликативных производственных функций, учитывающих влияние заработной платы на объем выпуска, как фактора,

существенно влияющего на производительность трудовых ресурсов (например, [27]). Особенностью производственной функции, предлагаемой в данной статье, является принятие во внимание относительного (по отношению ко всем соседним наукоградом) уровня заработной платы. Чем выше этот уровень, тем более квалифицированные трудовые ресурсы могут быть задействованы на предприятиях данного города.

Пусть,

- ◆ $J_i = \{j_{i1}, j_{i2}, \dots, j_{i|J_i|}\}$, $i \in I$ – набор индексов агенто-предприятий расположенных в i -ом наукограде, $|J_i|$ – общее количество предприятий;
- ◆ $s_{j_i}(t_k) \in \{1, 2, 3\}$ – состояния j_i -го агента-предприятия i -го наукограда в момент t_k , ($t_k \in T$): $s_{j_i}(t_k) = 1$ – первое состояние, соответствующее низким темпам экономического роста («красное»), $s_{j_i}(t_k) = 2$ – второе состояние, соответствующее средним темпам экономического роста («желтое»), $s_{j_i}(t_k) = 3$ – третье состояние, соответствующее высоким темпам экономического роста («зеленое»);
- ◆ $\{\tilde{K}_{j_i}(t_k), \tilde{L}_{j_i}(t_k)\}$, $i \in I, j_i \in J_i$ – основные фонды и трудовые ресурсы j_i -го агента-предприятия i -го наукограда в момент t_k , ($t_k \in T$);
- ◆ $\{\delta\tilde{K}_{j_i}(t_k), \delta\tilde{L}_{j_i}(t_k)\}$, $i \in I, j_i \in J_i$ – темпы ввода новых основных фондов и трудовых ресурсов j_i -го агента-предприятия i -го наукограда в момент t_k , ($t_k \in T$);
- ◆ $\{Q_i(t_k), Y_i(t_k)\}$, $i \in I$ – инвестиции в основные фонды и количество новых рабочих мест, выделяемых i -ому наукограду в момент t_k , ($t_k \in T$);
- ◆ $\{\alpha_{j_i}, \beta_{j_i}\}$ ($0, 1$), $i \in I, j_i \in J_i$ – коэффициенты эластичности выпуска по отношению к основным фондам и трудовым ресурсам j_i -го агента-предприятия i -го наукограда, $\alpha_{j_i} + \beta_{j_i} = 1$;
- ◆ A_{j_i} , $i \in I, j_i \in J_i$ – фактор (коэффициент), отражающий влияние научно-технического прогресса (НТП) на динамику объема выпуска;
- ◆ $\{\mu, \iota\}$ – известные коэффициенты выбытия основных фондов и трудовых ресурсов, соответственно.

Тогда, объем выпуска j_i -го агента-предприятия ($j_i \in J_i$) i -го наукограда ($i \in I$) может быть задан с использованием производственной функции типа Кобба-Дугласа в момент t_k , ($t_k \in T$):

$$\tilde{V}_{j_i}(t_k) = A_{j_i} S_{j_i} \left(\tilde{K}_{j_i}(t_k) \right)^{\alpha_{j_i}} \left(\tilde{L}_{j_i}(t_k) \right)^{\beta_{j_i}}, \quad (5)$$

где

$$S_{j_i} = \frac{W_i}{\frac{1}{|I|} \sum_{i=1}^{|I|} W_i}, \quad (6)$$

$$\tilde{K}_{j_i}(t_k) = \tilde{K}_{j_i}(t_{k-1}) + \delta\tilde{K}_{j_i}(t_k) - \mu\tilde{K}_{j_i}(t_{k-1}), \quad (7)$$

$$\tilde{L}_{j_i}(t_k) = \tilde{L}_{j_i}(t_{k-1}) + \delta\tilde{L}_{j_i}(t_k) - \iota\tilde{L}_{j_i}(t_{k-1}), \quad (8)$$

$$\delta\tilde{K}_{j_i}(t_k) = \begin{cases} Q_i(t_k) \frac{\tilde{V}_{j_i}(t_{k-1})}{\sum_{j_i=1}^{|J_i|} \tilde{V}_{j_i}(t_{k-1})} \left(1 + \gamma \frac{\sum_{j_i=1}^{|J_i|} r_{j_i}(t_{k-1})}{|J_i|} \right), & \text{если } s_{j_i}(t_{k-1}) = 1, \\ Q_i(t_k) \frac{\tilde{V}_{j_i}(t_{k-1})}{\sum_{j_i=1}^{|J_i|} \tilde{V}_{j_i}(t_{k-1})} \left(1 - \gamma \frac{\sum_{j_i=1}^{|J_i|} r_{j_i}(t_{k-1})}{|J_i|} \right), & \text{если } s_{j_i}(t_{k-1}) = 3, \end{cases} \quad (9)$$

$$\delta\tilde{L}_{j_i}(t_k) = \begin{cases} Y_i(t_k) \frac{\tilde{V}_{j_i}(t_{k-1})}{\sum_{j_i=1}^{|J_i|} \tilde{V}_{j_i}(t_{k-1})} \left(1 + \gamma \frac{\sum_{j_i=1}^{|J_i|} r_{j_i}(t_k)}{|J_i|} \right), & \text{если } s_{j_i}(t_{k-1}) = 1, \\ Y_i(t_k) \frac{\tilde{V}_{j_i}(t_{k-1})}{\sum_{j_i=1}^{|J_i|} \tilde{V}_{j_i}(t_{k-1})} \left(1 - \gamma \frac{\sum_{j_i=1}^{|J_i|} r_{j_i}(t_k)}{|J_i|} \right), & \text{если } s_{j_i}(t_{k-1}) = 3, \end{cases} \quad (10)$$

$$r_{j_i}(t_k) = \begin{cases} 1, & \text{если } s_{j_i}(t_k) = 1, \\ 0, & \text{если } s_{j_i}(t_k) \neq 1. \end{cases} \quad (11)$$

На рис. 2 показана карта состояний (state-flow chart) j_i -го агента-предприятия ($j_i \in J_i$) i -го наукограда ($i \in I$), реализующая правила перехода к новым состояниям $s_{j_i}(t_k) \in \{1, 2, 3\}$.

Переходы к новым состояниям (обозначены знаком $\textcircled{?}$ на рис. 2) осуществляются с использованием следующего правила:

$$s_{j_i}(t_k) = \begin{cases} 1, & \text{если } \frac{1}{t_k} \sum_{\tilde{t}_k=1}^{t_k} \frac{\tilde{V}_{j_i}(\tilde{t}_k)}{\tilde{V}_{j_i}(\tilde{t}_{k-1})} < \varphi_1, \\ 2, & \text{если } \varphi_1 \leq \frac{1}{t_k} \sum_{\tilde{t}_k=1}^{t_k} \frac{\tilde{V}_{j_i}(\tilde{t}_k)}{\tilde{V}_{j_i}(\tilde{t}_{k-1})} < \varphi_2, \\ 3, & \text{если } \frac{1}{t_k} \sum_{\tilde{t}_k=1}^{t_k} \frac{\tilde{V}_{j_i}(\tilde{t}_k)}{\tilde{V}_{j_i}(\tilde{t}_{k-1})} \geq \varphi_2. \end{cases} \quad (12)$$

Здесь $\{\varphi_1, \varphi_2\}$ – заданные пороговые значения, определяющие условия соответствия агентов-предприятий состояниям низкого, среднего и высокого темпов экономического роста.

Предлагаемая модель учитывает влияние относительного уровня заработной платы (6) на объем выпуска (5). Использование соотношений (9)–(11) позволяет обеспечить приоритетную поддержку агентов-предприятий, находящихся в первом состоянии низких темпов экономического роста. Здесь $\gamma \in [0, 1]$ – коэффициент, определяющий общий уровень подобной поддержки. Таким образом, инвестиции в основные фонды и новые рабочие места, выделяемые для каждого наукограда, распределяются между «проблемными» агентами-предприятиями пропорционально их вкладу в совокупный выпуск.

1.3. Постановка оптимизационной задачи

Итак, рассматривается система, в которой требуется обеспечить минимизацию общего количества предприятий, характеризующихся низкими темпами экономического роста для всех выбранных наукоградов. Для устойчивого развития научных кластеров желательно обеспечить переход боль-

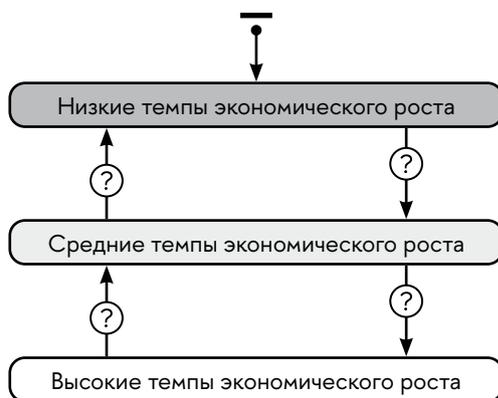


Рис. 2. Карта состояний (state-flow chart) агента-предприятия.

шей части проблемных предприятий к состояниям среднего и высокого темпов экономического роста и сохранение таких благоприятных состояний на максимально длительном временном интервале. Однако, в текущих условиях дефицита инвестиционного капитала, а также ограниченных возможностей по созданию новых высокотехнологичных рабочих мест, важно найти такие состояния и режимы функционирования системы, при которых минимизация числа проблемных организаций достигается с минимальными затратами хотя бы к некоторому заданному моменту модельного времени. В результате будет обеспечено качественное улучшение характеристик соответствующих предприятий, практически гарантирующее сохранение достигнутых ими состояний на достаточно длительном временном интервале в силу инерционности значений производственных факторов. При отсутствии статистической информации об имеющихся бюджетных и инвестиционных ограничениях, обосновано сведение подобной оптимизационной задачи к однокритериальной со следующим целевым функционалом, оцениваемым в конечный момент модельного времени $t_{|T|}$, ($t_{|T|} \in T$):

$$N(t_{|T|}) = n_1 \sum_{i=1}^{|I|} \sum_{j=1}^{|J_i|} r_{j_i}(Q_i(t_k), Y_i(t_k), W_i(t_k)) + n_2 \sum_{i=1}^{|I|} Q_i(t_k) + n_3 \sum_{i=1}^{|I|} (Y_i(t_k)W_i(t_k)), \quad (13)$$

где

$$r_{j_i}(Q_i(t_k), Y_i(t_k), W_i(t_k)) = \begin{cases} 1, & \text{если } s_{j_i}(t_k) = 1, \\ 0, & \text{если } s_{j_i}(t_k) \neq 1. \end{cases} \quad (14)$$

Здесь,

- ♦ $\{Q_i(t_k), Y_i(t_k), W_i(t_k)\}$, $i \in I$ – инвестиции в основные фонды, количество новых рабочих мест и заработная плата в i -ом наукограде в момент t_k , ($t_k \in T$) – управляющие параметры модели;
- ♦ $\{n_1, n_2, n_3\}$ – коэффициенты нормализации факторов, используемые для приведения значений элементов целевого функционала к одному масштабу.

Выбор значений управляющих параметров модели осуществляется в каждый момент времени t_k , ($t_k \in T$) и приводит к изменению производственных характеристик агентов-предприятий, т.е. основных фондов, трудовых ресурсов и объемов выпуска в соответствии с (5)–(11). В результате, обеспечивается воздействие на среднегодовые темпы эконо-

номического роста и состояния соответствующих предприятий, см. (12)–(14).

Тогда можно сформулировать оптимизационную задачу в следующем виде.

Задача А. Требуется минимизировать значение общего количества предприятий, характеризующихся низкими темпами экономического роста в конечный момент модельного времени $t_{|T|}$, ($t_{|T|} \in T$), а также инвестиционные и операционные расходы, связанные с реализацией данной стратегии:

$$\min_{\{Q_i(t_k), Y_i(t_k), W_i(t_k)\}_{k=1}^{|T|}} N(t_{|T|}) \quad (15)$$

при ограничениях

$$\underline{Q} \leq Q_i(t_k) \leq \bar{Q}, \quad \underline{Y} \leq Y_i(t_k) \leq \bar{Y}, \quad \underline{W} \leq W_i(t_k) \leq \bar{W}, \\ i \in I, k = 1, 2, \dots, |T|.$$

Здесь $\{Q, Y, W\}$, $\{\bar{Q}, \bar{Y}, \bar{W}\}$ – известные нижние и верхние значения управляющих параметров имитационной модели.

Задача А может быть решена с использованием генетического оптимизационного алгоритма, агрегированного по целевому функционалу с предложенной имитационной моделью. Реализация подобных алгоритмов основана на методах эволюционного поиска наилучших потенциальных реше-

ний и подробно описана в [12, 28]. Для упрощения рассматриваемой задачи, значения управляющих параметров в подобной системе могут быть заданы для каждого i -го наукограда ($i \in I$) как константы в начальный момент модельного времени.

2. Программная реализация модели

Программная реализация разработанной имитационной модели (1)–(15) выполнена в системе AnyLogic. Особенностью предлагаемого подхода является использование ГИС-карт для визуализации состояний агентов-предприятий, размещаемых в заданных пространственных координатах, а также интеграция модели с ранее разработанным генетическим оптимизационным алгоритмом [28], что, в частности, обеспечивает решение **Задачи А**.

Разработанный программный комплекс обеспечивает возможность навигации по всем изучаемым наукоградом, варьирования значениями управляющих параметров модели, визуализации состояний агентов-предприятий на карте.

На *рис. 3* показана реализация модели перераспределения населения и трудовых ресурсов между рассматриваемыми наукоградом (модели верхнего уровня) в AnyLogic.

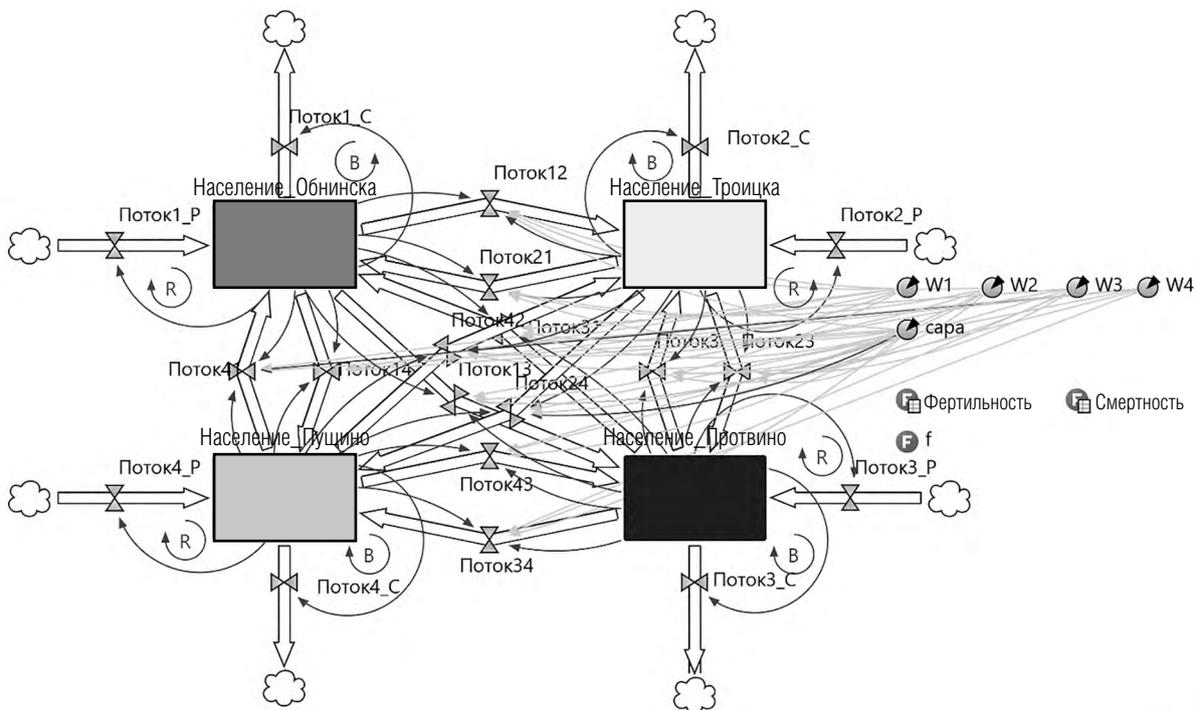


Рис. 3. Модель перераспределения трудовых ресурсов между в AnyLogic.

Из рис. 3 видно, что для построения данной модели используются методы системной динамики (см. например, [11, 22]). Численность населения Обнинска, Троицка, Протвино и Пушкино моделируется с использованием четырех системных уровней («резервуаров»), взаимодействующих посредством потоков с темпами, значения которых зависят от разницы в уровне заработной платы, размера рынка труда и др. Таким образом, моделируется влияние «гравитационного эффекта» на темпы внутренней миграции. Также учитываются коэффициенты фертильности и смертности со своими известными значениями. Имеющиеся в модели усиливающие и балансирующие обратные связи отмечены буквами **R** и **B**, соответственно (рис. 3).

На рис. 4 показана реализация модели поведения агентов-предприятий (модели нижнего уровня) в AnyLogic.

Из рис. 4 видно, что для построения данной модели используются как методы системной динамики, так и агентного моделирования (см. например, [5, 6, 13, 14, 16]). В частности, для моделирования возможных изменений состояний агентов, в соответствии с заданными правилами, используется карта состояний (элемент **statechart** на рис. 4). При

этом, для определения значений среднегодового темпа роста выпуска, значение которого сравнивается с заданными пороговыми значениями для реализации переходов между состояниями каждого агента, вычисляется с использованием модели системной динамики (левая часть рис. 4). Подобная модель использует функцию типа Кобба-Дугласа (см. (5)–(11)) для вычисления объема выпуска с учетом значений управляющих параметров. Результаты имитационного моделирования, в частности, состояние каждого агента-предприятия, объем выпуска, трудовые ресурсы и др. автоматически выгружаются в модель верхнего уровня и используется для симуляции «гравитационного эффекта», определяющего процесс перераспределения населения и трудовых ресурсов (рис. 3).

3. Результаты оптимизационных экспериментов

Для улучшения состояний всего ансамбля агентов-предприятий, в частности, минимизации количества организаций с низкими темпами экономического роста (менее 2% в год), были проведены оптимизационные эксперименты, нацеленные на поиск рекомендуемых значений управляющих па-

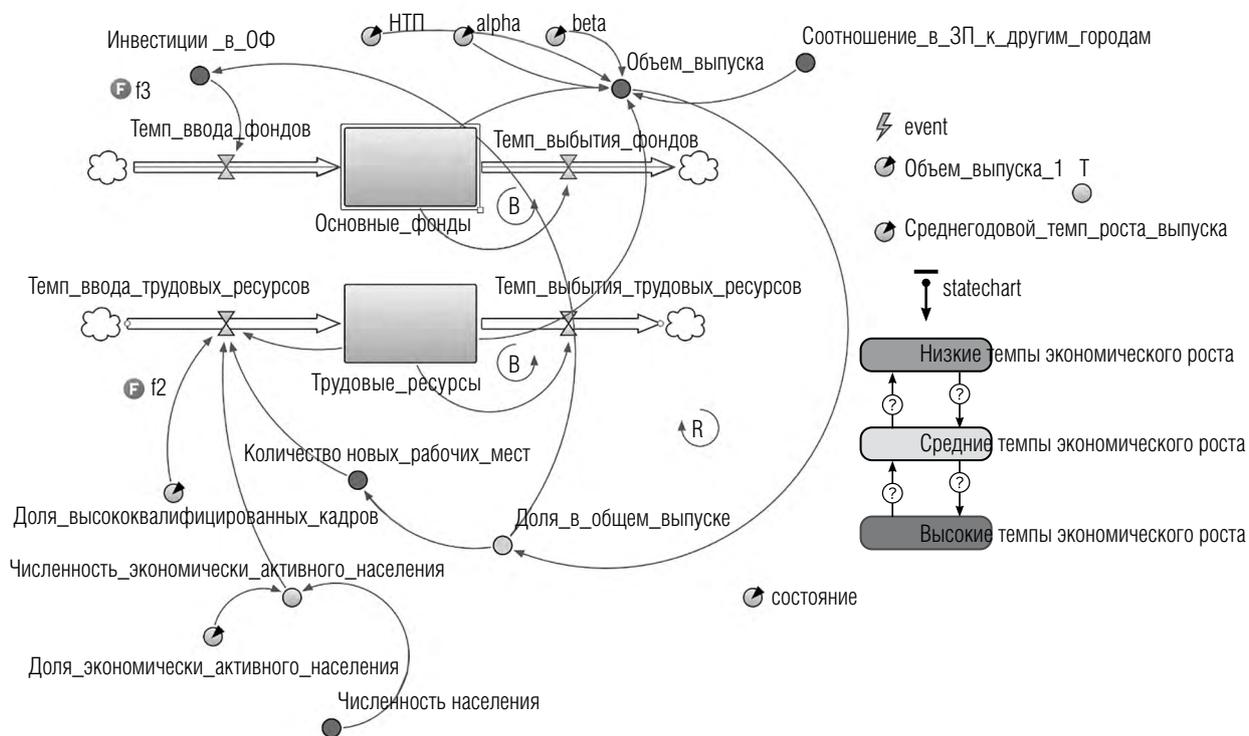


Рис. 4. Модель поведение агентов-предприятий в AnyLogic.

Таблица 1.

Пример исходных данных по агентам-предприятия для г. Обнинск (за 2016 г.)

№	Предприятие	Долгота	Широта	Трудовые ресурсы, чел.	Инвестиции в основные фонды, тыс. руб.	Основные фонды, тыс. руб.	Объем выпуска, тыс. руб.
1	АО «ГНЦ РФ-ФЗИ»	55.089851	36.591257	2612	14782	594124	3199431
2	АО «ОНПП «Технология»	55.105885	36.636379	2641	488089	6527163	5856985
3	АО «ГНЦ РФ НИФХИ»	55.063435	36.623261	528	103226	1300044	803209
4	МРНЦ им. А.Ф. Цыба – филиал ФГБУ «НМИРЦ» МЗ РФ	55.111844	36.617534	1559	136132	2230550	2052357
5	НПО «Тайфун»	55.104454	36.609027	580	1551	80820	293792
6	ВНИИРАЭ	55.106651	36.638687	270	3058	110706	242231
7	Публичное акционерное общество «Приборный завод «Сигнал»	55.11013	36.59102	308	48217	996640	1061128
8	Общество с ограниченной ответственностью «Хемофарм»	55.13496	36.6424	209	169412	2247154	1374894
9	Акционерное общество «Прогресс-экология»	55.12122	36.58286	236	14758	696989	1709746
10	Общество с ограниченной ответственностью «Ниармедик Фарма»	55.11442	36.61703	125	2810003	29994145	3506284

раметров модели. В качестве исходных данных были использованы укрупненные статические данные наукоградов указанные в [29], а также стратегии социально-экономического развития наукоградов, опубликованные на сайтах городских администраций. Пример исходных данных по агентам-предприятиям для г. Обнинск представлен в *таблице 1*.

На *рис. 5* представлена визуализация состояний агентов-предприятий, полученная в результате оптимизации значений управляющих параметров имитационной модели («сценарий на основе рекомендуемых параметров») в сравнении с ранее утвержденными инвестиционными и кадровыми стратегиями («базовый сценарий (до оптимизации)»).

Состояния агентов-предприятий, представленные на *рис. 5*, соответствуют десятилетнему горизонту стратегического планирования. Сценарий изменений подобных состояний на основе рекомендуемых параметров (*рис. 5*) был сформирован с помощью генетического оптимизационного алгоритма, агрегированного по целевому функционалу с разработанной имитационной моделью, реализованной в AnyLogic. В *таблице 2* представлены значения управляющих параметров модели, советуемые рассматриваемым сценариям.

Из *таблицы 2* следует, что для улучшения состояний агентов-предприятий в изучаемых наукоградах необходимо существенно увеличить инвестиции в основные фонды (преимущественно в

г. Обнинск	Базовый сценарий (до оптимизации)	Сценарий на основе рекомендуемых параметров
ООО «Хемофарм»	К	Ж
АО «Прогресс-экология»	З	З
ООО «Ниармедик Фарма»	К	Ж
МРНЦ им. А.Ф. Цыба – филиал ФГБУ «НМИРЦ» МЗ РФ	Ж	Ж
ПАО «Приборный завод «Сигнал»	Ж	З
АО «ОНПП «Технология»	Ж	Ж
НПО «Тайфун»	З	З
АО «ГНЦ РФ-ФЗИ»	З	З
ВНИИРАЭ	З	З
АО «ГНЦ РФ НИФХИ»	К	Ж

К Ж З
 Первое состояние (низкие темпы роста) Второе состояние (средние темпы роста) Третье состояние (высокие темпы роста)

Рис. 5. Состояния агентов-предприятий до и после оптимизации.

Таблица 2.

Значения управляющих параметров модели

Управляющие параметры	Наукоград	Базовый сценарий (до оптимизации)	Сценарий на основе рекомендуемых параметров
Инвестиции в основные фонды (млн руб./год)	г. Обнинск	3320	4341
	г. Троицк	453	3108
	г. Протвино	246	2252
	г. Пущино	374	2100
Количество новых рабочих мест (ед./год)	г. Обнинск	1000	3020
	г. Троицк	500	2550
	г. Протвино	300	1551
	г. Пущино	200	1121
Среднемесячная заработная плата (руб./мес.)	г. Обнинск	62 000	72 000
	г. Троицк	66 000	66 000
	г. Протвино	60 000	60 000
	г. Пущино	45 000	55 000

Троицке, Протвино и Пущино), а также кратно увеличить количество создаваемых рабочих мест. При этом, рекомендуется увеличить уровень заработной платы в Обнинске и Пущино, чтобы предотвратить отток кадров в соседние города (в частности, в г. Троицк). В результате, будет обеспечено принципиальное улучшение состояний агентов-предприятий (см. рис. 5), большинство из которых перейдет к третьему состоянию высоких темпов экономического роста.

Заключение

В данной работе представлена новая агент-ориентированная имитационная модель развития отдельных научно-производственных кластеров России с реализацией в системе AnyLogic. Разработанная модель состоит из двух уровней (рис. 1). Модель верхнего уровня реализует среду, в которой функционируют агенты-предприятия, относящиеся к различным наукоградам, между которыми реализуется потоковое взаимодействие с соответствующим перераспределением численности населения и трудовых ресурсов. Модель нижнего уровня обеспечивает реализацию логики поведения агентов-предприятий, характеристики которых описываются с использованием методов системной динамики. Сформулирована важная оптимизационная задача по минимизации общего

количества предприятий, характеризуемых низкими темпами экономического роста с учетом инвестиционных и операционных расходов. Предложенная двухуровневая имитационная модель реализована в виде программного комплекса (на примере научно-производственных кластеров Обнинска, Троицка, Протвино и Пущино), предназначенного, в первую очередь для органов муниципального и регионального управления, реализующих стратегию инновационного развития. Использование такой системы, позволило сформировать рекомендуемые значения управляющих параметров (таких как, инвестиции в основные фонды, количество новых рабочих мест и др.), обеспечивающие существенное улучшение состояний агентов-предприятий, расположенных в изучаемых наукоградах на десятилетнем горизонте стратегического планирования (рис. 5 и табл. 2). Разработанный экономико-математический и компьютерный инструментарий предназначен, в первую очередь для управления научно-производственными кластерами реализующих стратегию инновационного развития.

Дальнейшие исследования будут направлены на изучение проблем, связанных с преодолением дефицита высококвалифицированной рабочей силы, а также созданием детализированных агент-ориентированных моделей развития научно-производственных кластеров. ■

Литература

1. Forrester J. W. *Urban Dynamics*. M.I.T. Press, Cambridge, 1969.
2. Li G., Kou C., Wang Y., Yang H. System dynamics modelling for improving urban resilience in Beijing, China // *Resources, Conservation and Recycling*. 2020. Vol. 161. Art. No. 104954. <https://doi.org/10.1016/j.resconrec.2020.104954>
3. Diemer A., Nedelciu C.E. System dynamics for sustainable urban planning // *Sustainable Cities and Communities*. Encyclopedia of the UN Sustainable Development Goals (eds. W. Leal Filho, A. Marisa Azul, L. Brandli, P. Gökçin Özuyar, T. Wall). Springer, Cham. 2020. P. 760–773. https://doi.org/10.1007/978-3-319-95717-3_115
4. Armenia S., Barnabè F., Pompei A., Scolozzi R. System dynamics modelling for urban sustainability // *Urban Sustainability*. Springer Texts in Business and Economics (eds. J. Papathanasiou, G. Tsaples, A. Blouchoutzi). Springer, Cham. 2021. P. 131–173. https://doi.org/10.1007/978-3-030-67016-0_4
5. Макаров В.Л., Бахтизин А.Р., Эпштейн Д. Агент-ориентированное моделирование для сложного мира. Часть 1 // *Экономика и математические методы*. 2022. Т. 58. № 1. С. 5–26. <https://doi.org/10.31857/S042473880018970-6>
6. Макаров В.Л., Бахтизин А.Р., Эпштейн Д. Агент-ориентированное моделирование для сложного мира. Часть 2 // *Экономика и математические методы*. 2022. Т. 58. № 2. С. 7–21. <https://doi.org/10.31857/S042473880020009-8>
7. Макаров В.Л., Бахтизин А.Р., Бекларян Г.Л., Акопов А.С., Ровенская Е.А., Стрелковский Н.В. Укрупненная агент-ориентированная имитационная модель миграционных потоков стран Европейского Союза // *Экономика и математические методы*. 2019. Т. 55. № 1. С. 3–15. <https://doi.org/10.31857/S042473880004044-7>
8. Макаров В.Л., Бахтизин А.Р., Бекларян Г.Л., Акопов А.С., Ровенская Е.А., Стрелковский Н. Агентное моделирование популяционной динамики двух взаимодействующих сообществ: мигрантов и коренных жителей // *Экономика и математические методы*. 2020. Т. 56. № 2. С. 5–19. <https://doi.org/10.31857/S042473880009217-7>
9. Макаров В.Л., Бахтизин А.Р., Бекларян Г.Л., Акопов А.С. Цифровой завод: методы дискретно-событийного моделирования и оптимизации производственных характеристик // *Бизнес-информатика*. 2021. Т. 15. № 2. С. 7–20. <https://doi.org/10.17323/2587-814X.2021.2.7.20>
10. Макаров В.Л., Бахтизин А.Р., Бекларян Г.Л., Акопов А.С. Разработка программной платформы для крупномасштабного агент-ориентированного моделирования сложных социальных систем // *Программная инженерия*. 2019. Т. 10. № 4. С. 167–177. <https://doi.org/10.17587/prin.10.167-177>
11. Akopov A.S. Designing of integrated system-dynamics models for an oil company // *International Journal of Computer Applications in Technology*. 2012. Vol. 45. No. 4. P. 220–230. <https://doi.org/10.1504/IJCAT.2012.051122>
12. Akopov A.S. Parallel genetic algorithm with fading selection // *International Journal of Computer Applications in Technology*. 2014. Vol. 49. Nos. 3–4. P. 325–331. <https://doi.org/10.1504/IJCAT.2014.062368>
13. Akopov A.S., Beklaryan L.A., Saghatelian A.K. Agent-based modelling of interactions between air pollutants and greenery using a case study of Yerevan, Armenia // *Environmental Modelling & Software*. 2019. Vol. 116. P. 7–25. <https://doi.org/10.1016/j.envsoft.2019.02.003>
14. Akopov A.S., Beklaryan L.A., Saghatelian A.K. Agent-based modelling for ecological economics: a case study of the republic of Armenia // *Ecological Modelling*. 2017. Vol. 346. P. 99–118. <https://doi.org/10.1016/j.ecolmodel.2016.11.012>
15. Кислицын Е.В., Гоголин В.В. Имитационное моделирование экологической ситуации в мегаполисе // *Модели, системы, сети в экономике, технике, природе и обществе*. 2021. № 1(37). С. 92–106. <https://doi.org/10.21685/2227-8486-2021-1-8>
16. Акопов А.С., Бекларян Л.А. Агентная модель поведения толпы при чрезвычайных ситуациях // *Автоматика и телемеханика*. 2015. № 10. С. 131–143. <https://doi.org/10.1134/S0005117915100094>
17. Crooks A., Heppenstall A., Malleon N., Manley E. Agent-based modeling and the city: A gallery of applications // *Urban Informatics*. The Urban Book Series (eds. W. Shi, M.F. Goodchild, M. Batty, M.P. Kwan, A. Zhang). Springer, Singapore. 2021. P. 885–910. https://doi.org/10.1007/978-981-15-8983-6_46
18. Chen L. Agent-based modeling in urban and architectural research: A brief literature review // *Frontiers of Architectural Research*. 2012. Vol. 1. No. 2 P. 166–177. <https://doi.org/10.1016/j.foar.2012.03.003>
19. Tian G., Qiao Z. Modeling urban expansion policy scenarios using an agent-based approach for Guangzhou Metropolitan Region of China // *Ecology and Society*. 2014. Vol. 19. No. 3. Art. No. 52. <https://doi.org/10.5751/ES-06909-190352>
20. Yun, T.-S., Kim D., Moon, I.-C., Bae J.W. Agent-based model for urban administration: A case study of bridge construction and its traffic dispersion effect // *Journal of Artificial Societies and Social Simulation*. 2022. Vol. 25. No. 4. Art. No. 5. <https://doi.org/10.18564/jasss.4923>
21. Ghandar A., Theodoropoulos, G., Zhong M., Zhen B., Chen S., Gong Y., Ahmed A. An agent-based modelling framework for urban agriculture // *2019 Winter Simulation Conference (WSC)*, National Harbor, MD, USA. 2019. P. 1767–1778. <https://doi.org/10.1109/WSC40007.2019.9004854>

22. Акопов А.С. Системно-динамическое моделирование стратегии банковской группы // Бизнес-информатика. 2012. № 2(20). С. 10–19.
23. Stewart Q.J. The development of social physics // American Journal of Physics. 1950. Vol. 18. P. 239–253. <https://doi.org/10.1119/1.1932559>
24. Yap Y.L. The attraction of cities: A review of the migration literature // Journal of Development Economics. 1977. Vol. 4. No. 3. P. 239–264. [https://doi.org/10.1016/0304-3878\(77\)90030-x](https://doi.org/10.1016/0304-3878(77)90030-x)
25. Клейнер Г.Б. Производственные функции. Теория, методы, применение. М.: Финансы и статистика, 1986.
26. Суворов Н.В., Ахунов Р.Р., Губарев Р.В., Дзюба Е.И., Файзуллин Ф.С. Применение производственной функции Кобба – Дугласа для анализа промышленного комплекса региона // Экономика региона. 2020. Т. 16. № 1. С. 187–200. <https://doi.org/10.17059/2020-1-14>
27. Hellerstein J.K., Neumark D. Production function and wage equation estimation with heterogeneous labor: Evidence from a new matched employer-employee data set // Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches (eds. Ernst R. Berndt and Charles R. Hulten). Chicago: University of Chicago Press. 2007. P. 31-72. [Электронный ресурс]: <https://www.nber.org/system/files/chapters/c0873/c0873.pdf> (дата обращения 21.02.2024).
28. Beklaryan G.L., Акопов А.С., Khachatryan N.K. Optimisation of system dynamics models using a real-coded genetic algorithm with fuzzy control // Cybernetics and Information Technologies. 2019. Vol. 19. No. 2. P. 87–103. <https://doi.org/10.2478/cait-2019-0017>
29. Справка о результатах анализа соответствия показателей научно-производственных комплексов наукоградов Российской Федерации требованиям, установленным пунктом 8 статьи 2.1 Федерального закона № 70-ФЗ «О статусе наукограда Российской Федерации», и достижения результатов, предусмотренных планами мероприятий по реализации стратегий социально-экономического развития наукоградов Российской Федерации в 2022 году [Электронный ресурс]: <https://minobrnauki.gov.ru/upload/iblock/77c/cemgzf9g61hhktvme7dfmm9feddbfzv.pdf> (дата обращения 21.02.2024).

Об авторе

Бекларян Гаянэ Леоновна

кандидат экономических наук;

старший научный сотрудник, отделение теоретической экономики и математических исследований, лаборатория компьютерного моделирования социально-экономических процессов, Центральный экономико-математический институт, Российская академия наук, Россия, 117418, г. Москва, Нахимовский проспект, д. 47;

E-mail: glbeklaryan@gmail.com

ORCID: 0000-0002-1286-0345

Agent-based modeling and optimization of the characteristics for research-and-production clusters

Gayane L. Beklaryan

E-mail: glbeklaryan@gmail.com

Central Economics and Mathematics Institute, Russian Academy of Sciences
Address: 47, Nakhimovsky Ave., Moscow 117418, Russia

Abstract

This paper presents a developed agent-based simulation model for the development of research-and-production clusters in Russia implemented with the use of high-tech enterprises located in four science cities (Troitsk, Obninsk, Pushchino and Protvino) as the case study. A new approach to modeling and optimization of gross metropolitan product (GMP) is proposed, taking into account the influence of the “gravity effect” on the redistribution of labor resources between developing science cities and appropriate enterprises united in single research and research-and-production clusters. An important element of this approach is the formation of various scenarios for the strategic development of the research-and-production clusters being assessed and support for the possibility of choosing the most preferable scenario using an evolutionary optimization algorithm. An enlarged simulation model has been developed and implemented in AnyLogic describing the possible development trajectories of science cities with a corresponding change in the values of the most important characteristics: the number of economically active population, the number of research-and-production enterprises, the volume of products produced in high-tech sectors of the economy, GMP, etc. The designed framework is intended primarily for the management of research-and-production clusters implementing the strategy of innovative development. Such a framework uses methods of system dynamics and agent-based simulation modeling supported in the AnyLogic system, genetic optimization algorithms and GIS mapping for science cities, etc. to implement the required functionality. The approbation of the framework was completed with the use of real data published in the approved strategies of the relevant science cities development. As a result of the numerical experiments carried out, some recommendations were proposed for the development of the research-and-production clusters under study considering their mutual influence and the existing base of resources.

Keywords: research-and-production cluster, high-tech enterprise, science city, production characteristics, simulation modeling for enterprises, gross metropolitan product, agent-based modeling, system dynamics, gravity effect, production function, AnyLogic

Citation: Beklaryan G.L. (2024) Agent-based modeling and optimization of the characteristics for research-and-production clusters. *Business Informatics*, vol. 18, no. 1, pp. 36–51. DOI: 10.17323/2587-814X.2024.1.36.51

References

1. Forrester J. W. (1969) *Urban dynamics*. M.I.T. Press, Cambridge.
2. Li G., Kou C., Wang Y., Yang H. (2020) System dynamics modelling for improving urban resilience in Beijing, China. *Resources, Conservation and Recycling*, vol. 161, 104954. <https://doi.org/10.1016/j.resconrec.2020.104954>
3. Diemer A., Nedelciu C.E. (2020) System dynamics for sustainable urban planning. *Sustainable Cities and Communities. Encyclopedia of the UN Sustainable Development Goals* (eds. W. Leal Filho, A. Marisa Azul, L. Brandli, P. Gökçin Özuyar, T. Wall). Springer, Cham, pp. 760–773. https://doi.org/10.1007/978-3-319-95717-3_115
4. Armenia S., Barnabè F., Pompei A., Scolozzi R. (2021) System dynamics modelling for urban sustainability. *Urban Sustainability. Springer Texts in Business and Economics* (eds. J. Papathanasiou, G. Tsaples, A. Blouchoutzi). Springer, Cham, pp. 131–173. https://doi.org/10.1007/978-3-030-67016-0_4
5. Makarov V., Bakhtizin A., Epstein J. (2022) Agent-based modeling for a complex world. Part 1. *Economics and the Mathematical Methods*, vol. 58, no. 1, pp. 5–26 (in Russian). <https://doi.org/10.31857/S042473880018970-6>
6. Makarov V., Bakhtizin A., Epstein J., (2022) Agent-based modeling for a complex world. Part 2. *Economics and the Mathematical Methods*, vol. 58, no. 2, pp. 7–21 (in Russian). <https://doi.org/10.31857/S042473880020009-8>
7. Makarov V., Bakhtizin A., Beklaryan G., Akopov A., Rovenskaya E., Strelkovskiy N. (2019) Aggregated agent-based simulation model of migration flows of the European Union countries. *Economics and the Mathematical Methods*, vol. 55, no. 1, pp. 3–15 (in Russian). <https://doi.org/10.31857/S042473880004044-7>
8. Makarov V., Bakhtizin A., Beklaryan G., Akopov A. (2020) Agent-based modelling of population dynamics of two interacting social communities: migrants and natives. *Economics and the Mathematical Methods*, vol. 56, no. 2, pp. 5–19 (in Russian). <https://doi.org/10.31857/S042473880009217-7>
9. Makarov V.L., Bakhtizin A.R., Beklaryan G.L., Akopov A.S. (2021) Digital plant: methods of discrete-event modeling and optimization of production characteristics. *Business Informatics*, vol. 15, no. 2, pp. 7–20. <https://doi.org/10.17323/2587-814X.2021.2.7.20>

10. Makarov V.L., Bakhtizin A.R., Beklaryan G.L., Akopov A.S. (2019) Development of software framework for large-scale agent-based modeling of complex social systems. *Programmaya Ingeneria (Software Engineering)*, vol. 10, no. 4, pp. 167–177 (in Russian). <https://doi.org/10.17587/prin.10.167-177>
11. Akopov A.S. (2012) Designing of integrated system-dynamics models for an oil company. *International Journal of Computer Applications in Technology*, vol. 45, no. 4, pp. 220–230. <https://doi.org/10.1504/IJCAT.2012.051122>
12. Akopov A.S. (2014) Parallel genetic algorithm with fading selection. *International Journal of Computer Applications in Technology*, vol. 49, no. 3–4, pp. 325–331. <https://doi.org/10.1504/IJCAT.2014.062368>
13. Akopov A.S., Beklaryan L.A., Saghatelian A.K. (2019) Agent-based modelling of interactions between air pollutants and greenery using a case study of Yerevan, Armenia. *Environmental Modelling & Software*, vol. 116, pp. 7–25. <https://doi.org/10.1016/j.envsoft.2019.02.003>
14. Akopov A.S., Beklaryan L.A., Saghatelian A.K. (2017) Agent-based modelling for ecological economics: A case study of the republic of Armenia. *Ecological Modelling*, vol. 346, pp. 99–118. <https://doi.org/10.1016/j.ecolmodel.2016.11.012>
15. Kislitsyn E.V., Gogulin V.V. (2021) Simulation of the environmental situation in a megalopolis. *Modeli, sistemy, seti v ekonomike, tekhnike, prirode i obshchestve (Models, systems, networks in economics, technology, nature and society)*, vol. 1, no. 37, pp. 92–106 (in Russian). <https://doi.org/10.21685/2227-8486-2021-1-8>
16. Akopov A.S., Beklaryan L.A. (2015) An agent model of crowd behavior in emergencies. *Automation and Remote Control*, vol. 76, no. 10, pp. 1817–1827. <https://doi.org/10.1134/S0005117915100094>
17. Crooks A., Heppenstall A., Malleon N., Manley E. (2021) Agent-based modeling and the city: A gallery of applications. *Urban Informatics. The Urban Book Series* (eds. W. Shi, M.F. Goodchild, M. Batty, M.P. Kwan, A. Zhang). Springer, Singapore, pp. 885–910. https://doi.org/10.1007/978-981-15-8983-6_46
18. Chen L. (2012) Agent-based modeling in urban and architectural research: A brief literature review. *Frontiers of Architectural Research*, vol. 1, no. 2, pp. 166–177. <https://doi.org/10.1016/j.foar.2012.03.003>
19. Tian G., Qiao Z. (2014) Modeling urban expansion policy scenarios using an agent-based approach for Guangzhou Metropolitan Region of China. *Ecology and Society*, vol. 19, no. 3, art. 52. <https://doi.org/10.5751/ES-06909-190352>
20. Yun T.-S., Kim D., Moon I.-C., Bae J.W. (2022) Agent-based model for urban administration: A case study of bridge construction and its traffic dispersion effect. *Journal of Artificial Societies and Social Simulation*, vol. 25, no. 4, art. 5. <https://doi.org/10.18564/jasss.4923>
21. Ghandar A., Theodoropoulos G., Zhong M., Zhen B., Chen S., Gong Y., Ahmed A. (2019) An agent-based modelling framework for urban agriculture. 2019 *Winter Simulation Conference (WSC)*, National Harbor, MD, USA, pp. 1767–1778. <https://doi.org/10.1109/WSC40007.2019.9004854>
22. Akopov A.S. (2012) System dynamics modeling of banking group strategy. *Business Informatics*, vol. 2, no. 20, pp. 10–19 (in Russian).
23. Stewart Q.J. (1950) The development of social physics. *American Journal of Physics*, vol. 18, pp. 239–253. <https://doi.org/10.1119/1.1932559>
24. Yap Y.L. (1977) The attraction of cities: A review of the migration literature. *Journal of Development Economics*, vol. 4, no. 3, pp. 239–264. [https://doi.org/10.1016/0304-3878\(77\)90030-x](https://doi.org/10.1016/0304-3878(77)90030-x)
25. Kleiner G.B. (1986) *Production functions. Theory, methods, application*. Moscow: Finance and Statistics (in Russian).
26. Suvorov N.V., Akhunov R.R., Gubarev R.V., Dzyuba E.I., Fayzullin F.S. (2020) Applying the Cobb-Douglas production function for analysing the region's industry. *Economy of Regions*, vol. 16, no. 1, pp. 187–200 (in Russian). <https://doi.org/10.17059/2020-1-14>
27. Hellerstein J.K., Neumark D. (2007) Production function and wage equation estimation with heterogeneous labor: Evidence from a new matched employer-employee data set. *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, (eds. Ernst R. Berndt and Charles R. Hulten). Chicago: University of Chicago Press, pp. 31–72. <https://www.nber.org/system/files/chapters/c0873/c0873.pdf>
28. Beklaryan G.L., Akopov A.S., Khachatryan N.K. (2019) Optimisation of system dynamics models using a real-coded genetic algorithm with fuzzy control. *Cybernetics and Information Technologies*, vol. 19, no. 2, pp. 87–103. <https://doi.org/10.2478/cait-2019-0017>
29. *Guide on the results of the analysis of the compliance of the indicators of scientific and industrial complexes of science cities of the Russian Federation with the requirements established by paragraph 8 of Article 2.1 of Federal Law No. 70-FZ "On the status of the science city of the Russian Federation", and the achievement of the results provided for by the action plans for the implementation of the strategies of socio-economic development of science cities of the Russian Federation in 2022* (in Russian). Available at: <https://minobrnauki.gov.ru/upload/iblock/77c/cemgzf9g61hhktvme7dfmm9feddbfzv.pdf> (accessed 21 February 2023).

About the author

Gayane L. Beklaryan

Cand. Sci. (Econ.);

Senior Researcher, Department of Theoretical Economics and Mathematics Research, Laboratory of Computer Modeling of Social and Economic Processes, Central Economics and Mathematics Institute, Russian Academy of Sciences, 47, Nakhimovsky Ave., Moscow 117418, Russia;

E-mail: glbeklaryan@gmail.com

ORCID: 0000-0002-1286-0345

DOI: 10.17323/2587-814X.2024.1.52.64

Непараметрическая процедура сравнения эффективности работы подразделений сетевой организации

П.А. Колданов ^a 

E-mail: pkoldanov@hse.ru

В.А. Колданов ^b 

E-mail: vlad.kold@gmail.com

^a Лаборатория алгоритмов и технологий анализа сетевых структур, Национальный исследовательский университет «Высшая школа экономики»

Адрес: Россия, 603093, г. Н. Новгород, ул. Родионова, 136

^b Нижегородский государственный инженерно-экономический университет

Адрес: Россия, 606340, Нижегородская область, г. Княгинино, ул. Октябрьская, д. 22А, корпус 2

Аннотация

Для решения задачи сравнительного анализа эффективности работы филиалов по небольшому объему наблюдений случайной природы актуальным является непараметрический подход, не требующей вероятностной модели наблюдений. Актуальной также является задача сравнения результатов непараметрического подхода с результатами, полученными в рамках традиционно применяемой Гауссовской модели. Кроме того, актуальной является проблема получения непротиворечивого сравнения группы (не менее трех) филиалов. В настоящее время непараметрический подход и соответствующее сравнение с известными результатами решения рассматриваемой в настоящей работе задачи, полученного в рамках нормальной модели, отсутствуют. Кроме того, поиску методов получения непротиворечивых решений уделяется недостаточно внимания. Настоящая работа в определенном степени закрывает эти пробелы. Для решения этих задач в настоящей работе используются методы непараметрической статистики и теории одновременной проверки многих гипотез. В работе предлагается процедура сравнительного анализа эффективности работы нескольких подразделений сетевой организации по небольшому объему наблюдений, основанная на тестах Манна–Уитни. Проводится сравнение результатов предлагаемой непараметрической

процедуры с результатами, основанными на обобщениях тестов Стьюдента. Предлагается способ уменьшения числа возникающих проблем несовместимости, основанный на поиске подходящего уровня значимости. Приводится пример полностью непротиворечивого сравнения эффективности работы филиалов.

Ключевые слова: сетевая организация, эффективность деятельности филиалов, тесты Манна–Уитни, проблема несовместимости

Цитирование: Колданов П.А., Колданов В.А. Непараметрическая процедура сравнения эффективности работы подразделений сетевой организации // Бизнес-информатика. 2024. Т. 18. № 1. С. 52–64.
DOI: 10.17323/2587-814X.2024.1.52.64

Введение

Различные аспекты проблемы сравнения эффективности деятельности организаций обсуждаются во многих работах, например [1–3]. При этом, как правило, сравнения проводятся по многим показателям. Очевидно, успех таких сравнений зависит от того, насколько адекватно и качественно удастся провести сравнения по отдельным показателям, особенно если они имеют случайную природу. В этом последнем случае общепринятым является применение методов математической статистики. Такие методы делятся на:

- ◆ параметрические, которые основаны на конкретной вероятностной модели анализируемых показателей. При этом в качестве вероятностной модели чаще всего используется нормальное распределение [4, 5],
- ◆ непараметрические, свободные от детальной вероятностной модели, а иногда и от предположения о случайном характере анализируемых данных [3, 6–9].

Многие задачи, в том числе и задача настоящей работы, могут быть рассмотрены в рамках обоих подходов. В этом случае возникает необходимость сравнения выводов процедур, опирающихся на параметрическую модель, с выводами непараметрических процедур. Такое сравнение является одной из целей настоящей работы.

Известно много результатов сравнения параметрических и непараметрических тестов проверки гипотезы против альтернативы. При конечном объеме наблюдений такое сравнение проводится на основе анализа функции мощности тестов, опреде-

ляемой вероятностью отвержения гипотезы. При неограниченном объеме наблюдений сравнение тестов проверки гипотезы против альтернативы основано на вычислении показателей асимптотической эффективности [10]. Специфика рассматриваемой в настоящей работе задачи заключается в необходимости выбора одного из многих решений по небольшому объему наблюдений.

В настоящей работе предлагается процедура сравнительного анализа эффективности деятельности подразделений сетевой организации. Результаты применения такой процедуры могут быть использованы для принятия обоснованных управленческих решений руководством сетевой организации. При этом под эффективностью деятельности подразделения понимается отношение числа продаж определенного товара (например, числа автомобилей) к числу потенциальных покупателей. Под сетевой организацией понимается совокупность подразделений, работающих по общей схеме. Примерами таких организаций являются сеть филиалов крупной автомобильной компании или сеть магазинов «Пятерочка» и т.д.

В настоящей работе в качестве иллюстративного примера рассматривается задача поддержки управленческих решений руководством сети филиалов крупного университета. На подготовительные курсы таких филиалов принимаются все желающие. Естественной характеристикой эффективности работы персонала таких филиалов является отношение числа слушателей подготовительных курсов к числу потенциальных абитуриентов. Информация о сравнительной эффективности является основой для принятия стратегических решений по развитию сети филиалов.

Сетевую организацию удобно представить в виде графа. Вершины этого графа соответствуют подразделениям. Специфика рассматриваемого нами графа заключается в том, что в нем могут быть как направленные, так и ненаправленные ребра. Ненаправленное ребро между вершинами i, j добавляется в граф тогда и только тогда, когда принимается решение, что i -ое и j -ое подразделения работают одинаково эффективно. Направленное ребро от вершины i к вершине j добавляется в граф тогда и только тогда, когда принимается решение, что i -ое подразделение работает более эффективно, чем j -ое. Заметим, что обычно в графах используются либо только направленные, либо только ненаправленные ребра [11–15]. Мы будем пользоваться и теми, и другими просто потому, что они лучше позволяют подчеркнуть некоторые структуры анализируемого графа, которые характеризуют специфику анализируемой сетевой организации. К числу таких структур мы будем относить клики [16] (совокупность вершин, любые две из которых соединены ненаправленным ребром), которые характеризуют совокупность подразделений, относительно которых принимается решение о том, что они работают одинаково эффективно. В дальнейшем такие клики мы будем называть классами эквивалентности. Другим примером структуры является полный подграф только с направленными ребрами. Такую структуру мы будем называть структурой упорядочения или доминирования.

Во многих задачах, в частности, в анализируемом нами примере, число продаж естественно рассматривать как случайную величину. При этом анализ реальных данных, особенно когда их мало, может приводить к появлению противоречивых выводов. В этом случае соответствующий граф содержит некоторые логически противоречивые структуры, например подграфы из трех вершин, два ребра между которыми ненаправленные, а одно – направленное. Такого типа проблема возникла в [17] при обсуждении задачи проверки гипотез однородности не менее трех совокупностей и называлась проблемой несовместимости. Прикладные задачи, в которых возникает проблема несовместимости (проблема непротиворечивого объединения результатов сравнения эффективности пар подразделений), рассматривались в работах [4, 5] в рамках нормальной модели. Решение проблемы несовместимости было основано на введении дополнительного параметра Δ и переходе к задачам сравнения эффективности двух подразделений с точностью Δ . При этом

если эффективности подразделений i, j отличались меньше, чем на Δ , то принималось решение, что их эффективность одинакова (с точностью Δ). Такой прием позволяет решить проблему несовместимости, но приводит к дополнительной проблеме выбора Δ . Другая цель настоящей работы заключается в поиске способов решения проблемы несовместимости без введения вспомогательного параметра Δ .

В настоящей работе, в отличие от [4, 5], предположение о нормальном распределении числа продаж не используется. Попарное сравнение эффективности двух подразделений основано на применении тестов Манна–Уитни. Процедура сравнительного анализа подразделений по эффективности основана на комбинировании непараметрических тестов попарного сравнения двух подразделений. При этом используется графическое представление, которое оказывается удобным для визуализации возникающих проблем несовместимости. Предлагаемая непараметрическая процедура применяется к анализу данных, приведенных в [4], и приводится пример, в котором проблему несовместимости удается решить за счет анализа p -значений и подходящего выбора уровней значимости тестов попарных сравнений. Проводится сравнение с результатами, полученными в рамках нормальной модели.

Статья организована следующим образом: в разделе 1 приведены основные обозначения и постановка задачи; в разделе 2 описана непараметрическая процедура сравнительного анализа подразделений по эффективности и ее графическое представление; в разделе 3 приводятся иллюстративный пример, пример решения проблемы несовместимости и проводится сравнение с результатами, полученными в [4].

1. Постановка задачи

Данные о числе продаж удобно представить в виде матрицы $\|x_{ji}\|$, где $\|x_{ji}\|$ – отношение числа продаж к числу потенциальных покупателей в подразделении j в i -ый временной период, $j = 1, \dots, N$, где N – количество подразделений сетевой организации, $i = 1, \dots, m_j$, где m_j – количество анализируемых временных периодов работы j -ого подразделения. Будем считать, что наблюдения x_{ji} представляют собой значения случайных величин X_{ji} , которые описывают отношение численности продаж к числу потенциальных покупателей в подразделении j во временной период i . Предположим, что все вре-

менные периоды одинаковы, а случайные величины X_{ji} независимы при всех $j = 1, \dots, N$; $i = 1, \dots, m_j$ и при фиксированном j одинаково распределены, как X_j . Пусть $F_j(x)$ – функция распределения случайной величины X_j .

Задача, рассматриваемая в настоящей работе, заключается в построении и применении для анализа конкретных данных статистической процедуры различения гипотез вида:

$$\begin{aligned} H_1 : F_1(x) = F_2(x) = \dots = F_N(x), \quad \forall x \\ H_2 : F_1(x) < F_2(x) = \dots = F_N(x), \quad \forall x \\ \vdots \\ H_L : F_1(x) < F_2(x) < \dots < F_N(x), \quad \forall x \end{aligned} \quad (1)$$

При этом гипотеза H_1 означает, что эффективность всех подразделений одинакова, гипотеза H_2 означает, что подразделение 1 работает более эффективно, чем остальные подразделения, эффективность которых одинакова, и т.д. Заметим, что соотношения (1) не описывают все возможные соотношения между функциями распределения $F_j(x)$, $j = 1, \dots, N$. Мы ограничиваемся рассмотрением только этих гипотез, так как нас интересует только наличие систематического сдвига, к которому может приводить разная эффективность работы различных подразделений сетевой организации.

Как и в [4] будем использовать метод построения процедур со многими решениями, предложенный в [17]. Этот метод основан на сведении многоальтернативной задачи к совокупности соответствующим образом подобранных двухальтернативных порождающих задач. В нашем случае для различения (1) естественно рассматривать двухальтернативные задачи проверки гипотез $h_{ij} : F_j(x) \geq F_i(x), \forall x, \forall i, j = 1, \dots, N$.

При фиксированных i, j комбинация тестов φ_{ij} , φ_{ji} одновременной проверки гипотез h_{ij} и h_{ji} с ненулевой вероятностью может приводить к логически несостоятельному (при данном x) решению об отклонении обеих гипотез, т.е. к проблеме несовместимости. Как показано в [17], для исключения такого противоречия достаточно потребовать, чтобы уровни значимости α_{ij} , α_{ji} тестов φ_{ij} , φ_{ji} удовлетворяли условию $\alpha_{ij} + \alpha_{ji} < 1$. В этом случае комбинация тестов φ_{ij} , φ_{ji} приводит к совместной процедуре различения трех гипотез

$$\begin{aligned} H_{ij}^1 : F_i(x) < F_j(x); \\ H_{ij}^2 : F_i(x) = F_j(x); \\ H_{ij}^3 : F_i(x) > F_j(x). \end{aligned} \quad (2)$$

Однако объединение таких процедур с тремя решениями при различных i, j может привести к противоречию, а именно: с ненулевой вероятностью может быть принято решение о том, что (например):

$$F_1(x) = F_2(x) \text{ и } F_2(x) = F_3(x), \text{ но } F_1(x) \neq F_3(x).$$

Для исключения указанного противоречия в [4], следуя предложению [17], рассматривалась несколько измененная система порождающих гипотез. При этом исследования были ограничены случаем, когда $F_i(x)$ представляет собой нормальное распределение. В обозначениях настоящей работы измененная система порождающих гипотез имеет вид:

$$h'_{ij} : F_j(x) + \Delta \geq F_i(x), \forall x, \forall i, j = 1, \dots, N.$$

При комбинировании тестов φ'_{ij} , φ'_{ji} одновременной проверки гипотез h'_{ij} , h'_{ji} получаем процедуру различения трех гипотез:

$$\begin{aligned} H_{ij}^{\prime 1} : F_i(x) + \Delta < F_j(x); \\ H_{ij}^{\prime 2} : |F_i(x) - F_j(x)| < \Delta; \\ H_{ij}^{\prime 3} : F_i(x) > F_j(x) + \Delta. \end{aligned} \quad (3)$$

При этом проблема получения противоречивых выводов не возникает. Вместе с тем введение параметра Δ формально изменяет исходную задачу.

В настоящей работе параметр Δ не вводится и предположение о нормальном распределении не делается. При этом один из интересных вопросов заключается в поиске вариантов непротиворечивого объединения статистических правил с тремя решениями без введения параметра Δ .

Заметим, что в настоящее время в интенсивно развивающейся теории одновременной проверки многих гипотез не делается акцент на необходимости решения проблемы несовместимости [18–21]. При этом начиная с работы [22] проблема несовместимости рассматривается как слишком сильное требование, накладываемое на процедуру одновременной проверки многих гипотез. В рамках теории одновременной проверки многих гипотез в основном изучаются подходы к построению процедур, контролирующих вероятность хотя бы одной ошибки первого рода, доли ошибок первого рода и некоторые другие. В настоящей работе напротив, мы делаем акцент на решении проблемы несовместимости, что позволяет провести адекватное сравнение с результатами, полученными в [4].

2. Непараметрическая процедура сравнительного анализа и ее визуализация

2.1. Процедура с тремя решениями

Одна из наиболее эффективных непараметрических процедур различения гипотез (2) основана на статистике Манна–Уитни [6, 9]. Статистика Манна–Уитни имеет вид:

$$W_{ij}(x_i, x_j) = \sum_{r=1}^{m_i} \sum_{s=1}^{m_j} I(x_{is} < x_{jr}), \quad (4)$$

где $I(A)$ – индикатор события A .

При фиксированных i, j процедура с тремя решениями различения гипотез (2) в терминах p -значений может быть записана в виде

$$\varphi(i, j) = \begin{cases} d_{ij}^1 & p_{ij}^1 \leq \alpha_1 \\ d_{ij}^3 & p_{ij}^3 \leq \alpha_3 \\ d_{ij}^2 & p_{ij}^1 > \alpha_1, p_{ij}^3 > \alpha_3, \end{cases} \quad (5)$$

где d_{ij}^k – решение о принятии гипотезы H_{ij}^k ($k = 1, 2, 3$);

p_{ij}^1, p_{ij}^3 – соответствующие p -значения, а именно:

$$\begin{aligned} p_{ij}^1 &= P_{F_i=F_j}(W_{ij}(X_i, X_j) \leq W_{ij}(x_i, x_j)) \\ p_{ij}^3 &= P_{F_i=F_j}(W_{ij}(X_i, X_j) > W_{ij}(x_i, x_j)), \end{aligned} \quad (6)$$

где α_1, α_3 – уровни значимости тестов проверки гипотез H_{ij}^1 и H_{ij}^3 соответственно.

При этом предполагается, что среди наблюдаемых значений x_{j1}, \dots, x_{jm_j} и x_{i1}, \dots, x_{im_i} нет равных. Необходимые корректировки в случае равных наблюдений можно внести, опираясь на методику, изложенную в [9].

Таблицы распределения статистики (4) при малых m_i, m_j приведены в [9]. При больших m_i, m_j можно использовать нормальное распределение

$$N\left(\frac{m_i \cdot m_j}{2}, \frac{m_i \cdot m_j \cdot (m_i + m_j + 1)}{12}\right),$$

которое рекомендуется использовать при

$$\min(m_i, m_j) > 50 \text{ [9]}.$$

Из (6) очевидно, что при фиксированных i, j справедливо $p_{ij}^1 + p_{ij}^3 = 1$. Поэтому для применения процедуры с тремя решениями (5) достаточно информации о минимальном p -значении:

$$p_{ij} = \min(p_{ij}^1, p_{ij}^3). \quad (7)$$

2.2. Процедура со многими решениями и ее графическое представление

Процедуру различения гипотез (1) будем получать комбинированием процедур (5).

Такая процедура может быть записана в виде:

$$\delta = (\varphi(1, 2), \varphi(2, 3), \dots, \varphi(N-1, N)). \quad (8)$$

Через $f_i, i = 1, \dots, N$ обозначим эффективность работы i -го подразделения. В рассматриваемой задаче между эффективностями работы подразделений сетевой организации (f_1, \dots, f_N) возможны два типа отношений (доминирование или эквивалентность). Запись $f_i > f_j$ означает, что i -ое подразделение работает более эффективно (доминирует), чем j -ое подразделение. Запись $f_i = f_j$ означает, что i -ое и j -ое подразделения работают одинаково эффективно (эквивалентность). Для наглядного анализа результатов применения процедуры (8) воспользуемся методикой, предложенной в [23–25].

При данном векторе (f_1, \dots, f_N) введем матрицу D с элементами

$$d_{ij} = \begin{cases} 1, & f_i > f_j \\ 0, & f_i = f_j \\ -1, & f_i < f_j \end{cases}$$

и матрицу B с элементами

$$b_{ij} = \begin{cases} 1, & f_i > f_j \\ 0, & \text{иначе.} \end{cases}$$

Легко показать [24], что матрица B связана с матрицей D соотношением

$$D = B - B^T,$$

где B^T – транспонированная матрица B .

Соотношения, описываемые матрицей B , удобнее интерпретировать, если строки и столбцы матрицы B переставлены таким образом, чтобы получить верхнюю треугольную форму (т.е. собрать по возможности все единицы над главной диагональю матрицы B). Для получения верхней треугольной формы матрицы B можно упорядочить по убыванию строки (и столбцы) матрицы D по суммам элементов строк. Заменяя элементы -1 полученной матрицы на 0, получим матрицу B , наиболее согласованную с верхней треугольной формой. Матрица B , наиболее согласованная с верхней треугольной формой, позволяет выделить так называ-

емые «классы эквивалентности» («undifferentiated classes» [25]). Термин «класс эквивалентности» будет использоваться для обозначения наибольшего множества подразделений, эффективности работы которых значимо неотличимы друг от друга. Термин «наибольший» означает, что для любого подразделения i , которое не принадлежит данному классу эквивалентности, существует по крайней мере одно подразделение j из этого класса, такое что эффективности работы подразделений i и j значимо различимы. Матрица B , наиболее согласованная с верхней треугольной формой, показывает классы эквивалентности, как квадратные подматрицы, симметричные относительно главной диагонали, все элементы которых равны 0.

Очевидно, что пересекающиеся классы эквивалентности, получающиеся при применении процедуры δ , означают существование проблемы несовместимости. Матрицу D удобно визуализировать в виде графа $G = (V, E)$, где $V = \{1, 2, \dots, N\}$ – множество вершин графа, $E = \{e_{ij}\}$ – множество ребер графа. Если элемент d_{ij} матрицы D равен 1, то в граф G добавляется направленное ребро от вершины i к вершине j . При этом вершина i доминирует вершину j . В [11] вершина i называется родителем вершины j , а вершина j – ребенком вершины i . Если вершина i соединена направленным путем длины больше 1 с некоторой вершиной k , то вершина i называется предком вершины k , а вершина k – потомком вершины i . Если элемент d_{ij} матри-

цы D равен 0, то в граф G добавляется ненаправленное ребро между вершинами i и j . Если элемент $d_{ij} = -1$, то, так как $d_{ij} = -d_{ji}$, граф G уже содержит направленное ребро от вершины j к вершине i . При этом очевидно, что все вершины этого графа, соответствующие подразделениям из некоторого класса эквивалентности, соединены между собой ненаправленными ребрами и поэтому представляют собой клики графа G . Ниже мы будем отдельно изображать подграфы только с направленными ребрами и только с ненаправленными ребрами.

Заметим, что предлагаемое представление будет явно отображать существование проблем несовместимости полученных выводов, в случае их наличия. Очевидно, что если в представлении различные классы эквивалентности содержат одинаковые вершины, то проблема несовместимости имеет место.

3. Иллюстративный пример

Рассмотрим задачу сравнительного анализа эффективностей работы филиалов университета, которая была кратко описана во введении. Обозначим $1f$ – *первый* филиал университета, $2f$ – *второй* филиал университета и т.д. Данные для анализа заимствованы из [4] и приведены в *таблице 1*.

Минимальные p -значения (7) тестов (5) приведены в *таблице 2*.

Таблица 1.

Данные о числе слушателей
подготовительных курсов в различных филиалах

$1f$	$2f$	$3f$	$4f$	$5f$	$6f$	$7f$	$8f$
103	131	187	154				
92	212	262	92	151	99	235	
122	197	376	129	164	268	338	77
48	143	283	146	141	217	239	63
86	95	231	125	140	231	187	59
89	70	203	127	173	175	123	78
147	92	276	183	141	137	139	82
134	95	258	213	187	242	185	28
Количество потенциальных абитуриентов в i -ом филиале							
6390	7090	28900	6320	6320	11130	4660	2530

Таблица 2.

Минимальные p -значения

	1f	2f	3f	4f	5f	6f	7f
2f	p_{12}^1 0,4392						
3f	p_{13}^3 0,0023	p_{23}^3 0,0103					
4f	p_{14}^1 0,0364	p_{24}^1 0,0652	p_{34}^1 0,0001				
5f	p_{15}^1 0,0006	p_{25}^1 0,0469	p_{35}^1 0,0002	p_{45}^1 0,1678			
6f	p_{16}^1 0,3063	p_{26}^1 0,4775	p_{36}^1 0,0070	p_{46}^3 0,0760	p_{56}^3 0,0020		
7f	p_{17}^1 0,0002	p_{27}^1 0,0011	p_{37}^1 0,0002	p_{47}^1 0,0200	p_{57}^1 0,0012	p_{67}^1 0,0003	
8f	p_{18}^3 0,0147	p_{28}^3 0,0539	p_{38}^1 0,0007	p_{48}^1 0,1725	p_{58}^1 0,1830	p_{68}^1 0,0256	p_{78}^3 0,0175

3.1. Построение классов эквивалентности

Рассмотрим прежде всего традиционный уровень значимости $\alpha_j = 0,05, \forall i, j = 1, \dots, 8$. Матрица $D_{0,05}$ приведена в таблице 3.

Таблица 3.

Матрица $D_{0,05}$

	1f	2f	3f	4f	5f	6f	7f	8f	Σ
1f	-	0	1	-1	-1	0	-1	-1	-3
2f	0	-	1	0	-1	0	-1	0	-1
3f	-1	-1	-	-1	-1	-1	-1	-1	-7
4f	1	0	1	-	0	0	-1	0	1
5f	1	1	1	0	-	1	-1	0	3
6f	0	0	1	0	-1	-	-1	-1	-2
7f	1	1	1	1	1	1	-	1	7
8f	1	0	1	0	0	1	-1	-	2

Графическое представление матрицы $D_{0,05}$ показано на рисунке 1.

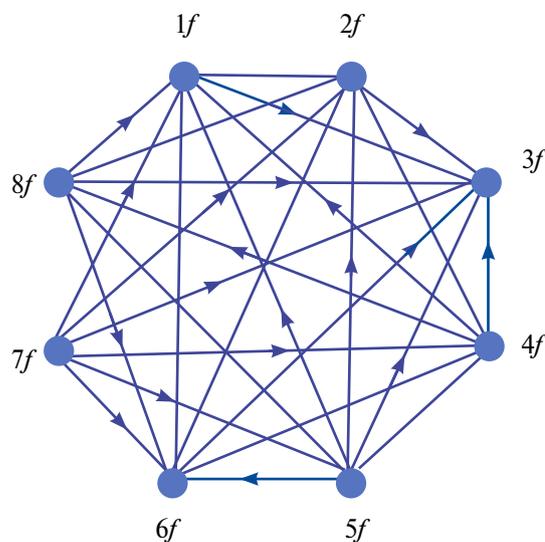


Рис. 1. Графическое представление матрицы $D_{0,05}$.

Матрица $B_{0,05}$, полученная из матрицы $D_{0,05}$, приведенная к верхней треугольной форме, показана в таблице 4.

Графическое представление матрицы $B_{0,1}$, приведенной к верхней треугольной форме, показано на рисунке 4.

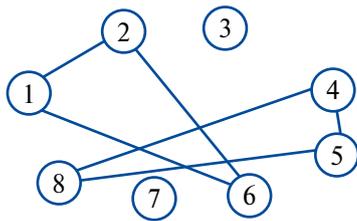


Рис. 4. Графическое представление матрицы $B_{0,1}$, приведенной к верхней треугольной форме. Номер филиала обозначен цифрой.

На рисунке 4 легко видеть, что в данном графе четыре клики $\{7\}$, $\{4, 5, 8\}$, $\{2, 6, 1\}$, $\{3\}$ причем эти клики не имеют общих вершин. Это свидетельствует о том, что при $\alpha_j = 0,1, \forall i, j = 1, \dots, 8$ проблема несовместимости не возникает, т.е. классы эквивалентности не пересекаются.

3.2. Построение структур упорядочения

Графическое представление матрицы $D_{0,05}$ показывает структуру упорядочения филиалов по эффективности их работы, приведенную на рисунке 5.

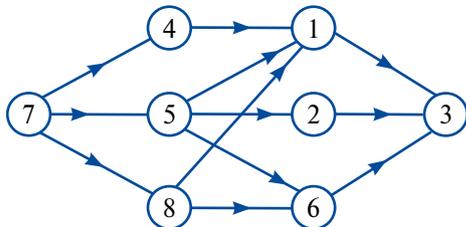


Рис. 5. Непараметрическое упорядочение $\alpha = 0,05$. Номер филиала обозначен цифрой.

Показаны только связи доминирования, связи эквивалентности и связи между предками и потомками не показаны. В частности, не проведено направленное ребро между вершинами 7 и 3, т.к. существует направленный путь $7 \rightarrow 4 \rightarrow 1 \rightarrow 3$ от предка $7f$ к потомку $3f$, означающий строгое упорядочение: $7f$ эффективнее $4f$, $4f$ эффективнее $1f$, $1f$ эффективнее $3f$. Заметим, что в данном графе нет направленных путей $7 \rightarrow 4 \rightarrow 2 \rightarrow 3$; $7 \rightarrow 4 \rightarrow 6 \rightarrow 3$; $7 \rightarrow 8 \rightarrow 2 \rightarrow 3$. Это свидетельствует об отсутствии полного упорядочения между эффективностями работы филиалов при $\alpha = 0,05$. Учитывая, что

$\{4f, 5f, 8f\}$ принадлежат одному классу эквивалентности, отсутствие таких путей приводит к логическим противоречиям. Подчеркнем, что при $\alpha = 0,1$ полное упорядочение имеет место (см. рис. 6) и логические противоречия не возникают.

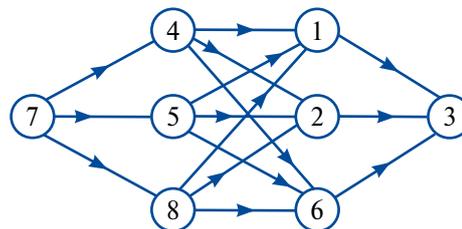


Рис. 6. Непараметрическое упорядочение $\alpha = 0,1$. Номер филиала обозначен цифрой.

3.3. Сравнение

На рисунке 7 показаны графы, построенные по результатам сравнения эффективностей работы филиалов как в предположении нормального распределения изучаемых случайных величин [4], так и в непараметрической постановке.

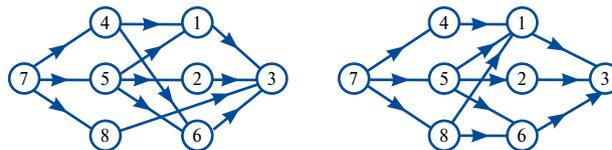


Рис. 7. Параметрическое (слева), непараметрическое (справа) упорядочения при $\alpha = 0,05$. Номер филиала обозначен цифрой.

Графы упорядочения, представленные на рисунке 7, отличаются тремя ребрами, а именно: ребро $(4, 6)$ присутствует при параметрическом упорядочении, и отсутствует при непараметрическом упорядочении; ребра $(8, 1)$; $(8, 6)$ присутствует при непараметрическом упорядочении, и отсутствует при параметрическом упорядочении.

Линейное упорядочение, построенное по схеме, предложенной в [4], соответствующее непараметрическому упорядочению (рис. 7, справа), имеет вид:

$$f_3 < f_1 \leq f_6 \leq f_2 \leq f_4 \leq f_8 \leq f_5 < f_7 \text{ с точностью } \Delta. \quad (9)$$

Получение такого линейного упорядочения, формально предложенного в [4] в несколько иной постановке (упорядочение с точностью Δ), осно-

вано на анализе числа направленных связей, выходящих из конкретной вершины или входящих в конкретную вершину. А именно, так как вершины 4, 5, 8 (рис. 7, справа) не соединяются направленными ребрами, то на первый взгляд кажется, что может быть вынесено решение $f_4 = f_8 = f_5$ с точностью Δ . Однако, так как вершина 4 доминирует только одну вершину 1 (из вершины 4 выходит одно направленное ребро), вершина 8 доминирует две вершины 1 и 6 (из вершины 8 выходит два направленных ребра), и вершина 5 доминирует три вершины 1, 2, 6 (из вершины 5 выходит три направленных ребра), то выносятся решение $f_4 \leq f_8 \leq f_5$ с точностью Δ . Запись $f_4 \leq f_8$ с точностью Δ означает, что $f_4 + \Delta < f_8$ или $|f_4 - f_8| < \Delta$. Аналогично, так как вершины 1, 2, 6 (рис. 7, справа) не соединяются направленными ребрами, то на первый взгляд кажется, что может быть вынесено решение $f_1 = f_2 = f_6$ с точностью Δ . Однако, так как вершина 1 – потомок всех вершин 4, 5, 8 (в вершину 1 входит три направленных ребра), вершина 6 – потомок двух вершин 5 и 8 (в вершину 6 входит два направленных ребра), вершина 2 – потомок одной вершины 5 (в вершину 2 входит одно направленное ребро), то выносятся решение $f_1 \leq f_6 \leq f_2$ с точностью Δ .

Линейное упорядочение, полученное в [4] имеет вид:

$$f_3 < f_1 = f_6 \leq f_2 \leq f_8 \leq f_4 \leq f_5 \leq f_7 \text{ с точностью } \Delta. \quad (10)$$

Упорядочения (9) и (10) отличаются очень незначительно. В самом деле, крайние элементы упорядочений совпадают, изменился знак $f_1 \leq f_6$ по сравнению со знаком $f_1 = f_6$, и нестрогое упорядочение $f_4 \leq f_8$ изменилось на упорядочение $f_8 \leq f_4$. Но в обоих случаях значимого различия эффективности работы филиалов f_8 и f_4 не обнаружено. Это косвенно свидетельствует о приемлемости нормальной модели, предложенной в [4].

Классы эквивалентности, представленные на рисунке 8, отличаются тремя ребрами, а именно: ребро (4, 6) присутствует при непараметрическом построении классов эквивалентности, и отсутствует при параметрическом построении; ребра (8, 1); (8, 6) присутствует при параметрическом постро-

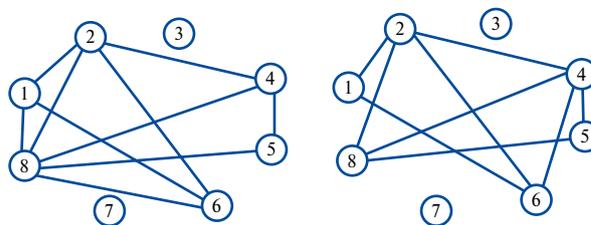


Рис. 8. Параметрические (слева), непараметрические (справа) классы эквивалентности при $\alpha = 0,05$. Номер филиала обозначен цифрой.

ении классов эквивалентности, и отсутствует при непараметрическом построении. Это вполне согласуется со сравнением структур упорядочения (см. рис. 7).

Заключение

В работе построена непараметрическая процедура сравнительного анализа эффективности работы нескольких подразделений сетевой организации по небольшому объему наблюдений. Приведен пример применения предлагаемого подхода к сравнительному анализу эффективностей работы филиалов ВУЗа. Выполнено сравнение результатов сравнительного анализа, полученных предлагаемой непараметрической процедурой с результатами, полученными в рамках нормальной модели [4]. Показано, что результаты непараметрического упорядочения без введения дополнительного параметра неопределенности Δ и результаты упорядочения, полученные в рамках нормальной модели с введением Δ , достаточно близки. Приведен пример полного непротиворечивого сравнения эффективности работы нескольких подразделений сетевой организации. ■

Благодарности

Результаты разделов 1 и 2 подготовлены в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ). Результаты раздела 3 подготовлены при поддержке гранта Российского научного фонда (проект 22-11-00073).

Литература

1. Алескерев Ф.Т., Белоусова В.Ю. Эффективное развитие филиальной сети коммерческого банка // Управление в кредитной организации. 2007. № 6. С. 23–34.
2. Крюков А.М. Анализ эффективности деятельности филиалов и подразделений – необходимое условие устойчивости бизнеса // Стратегические решения и риск-менеджмент. 2010. Т. 4. С. 84–87. <https://doi.org/10.17747/2078-8886-2010-4-84-87>

3. Мызникова М.А. Качество стратегического управления в условиях неопределенности: оценка в контексте устойчивого развития // Бизнес-информатика. 2022. Т. 16. № 3. С. 36–52. <https://doi.org/10.17323/2587-814X.2022.3.36.52>
4. Колданов А.П., Колданов П.А. Статистические процедуры со многими решениями в задаче анализа итогов приема в филиалы ВУЗа // Бизнес-информатика. 2012. Т. 19. № 1. С. 24–31.
5. Koldanov P.A. Efficiency analysis of branch network // Models, Algorithms, and Technologies for Network Analysis (eds. B.I. Goldengorin, V.A. Kalyagin, P.M. Pardalos). Springer Proceedings in Mathematics & Statistics. 2013. Vol. 59. P. 71–83. https://doi.org/10.1007/978-1-4614-8588-9_5
6. Хеттманспергер Т. Статистические выводы, основанные на рангах. М.: Финансы и статистика, 1987.
7. Кендалл М. Ранговые корреляции. М.: Статистика, 1975.
8. Холлендер М., Вулф Д. Непараметрические методы статистики. М.: Финансы и статистика, 1983.
9. Lehmann E.L. Nonparametrics: Statistical methods based on ranks. San Francisco: Hoden-Day, 1975. <https://doi.org/10.1002/zamm.19770570922>
10. Кендалл М., Стюарт А. Статистические выводы и связи. М.: Наука, 1973.
11. Jordan M.I. Graphical models // Statistical Science. 2004. Vol. 19. No. 1. P. 140–155. <https://doi.org/10.1214/088342304000000026>
12. Newman M.E.J. Networks. An introduction. New York: Oxford University Press, 2010.
13. Opsahl T., Panzarasa P. Clustering in weighted networks // Social Networks. 2009. Vol. 31. No. 2. P. 155–163.
14. Horvath S. Weighted network analysis. Applications in genomics and systems biology. New York: Springer, 2011. <https://doi.org/10.1007/978-1-4419-8819-5>
15. Li S., He J., Zhuang Y. A network model of the interbank market // Physica A: Statistical Mechanics and its Applications. 2010. Vol. 389. No. 24. P. 5587–5593. <https://doi.org/10.1016/j.physa.2010.08.057>
16. Boginski V., Butenko S., Pardalos P.M. Statistical analysis of financial networks // Computational Statistics & Data Analysis. 2005. Vol.48. No. 2. P. 431–443. <https://doi.org/10.1016/j.csda.2004.02.004>
17. Lehmann E.L. A theory of some multiple decision problems, I // The Annals of Mathematical Statistics. 1957. Vol. 28. No. 1. P. 1–25. <https://doi.org/10.1214/aoms/1177707034>
18. Bretz F., Hothorn T., Westfall P. Multiple comparisons using R. New York: Chapman and Hall/CRC, 2010. <https://doi.org/10.1201/9781420010909>
19. Hochberg Y., Tamhane A.C. Multiple comparison procedures. New York: John Wiley & Sons, 1987. <https://doi.org/10.1002/9780470316672>
20. Handbook of Multiple Comparisons / [Edited by X. Cui, T. Dickhaus, Y. Ding, J.C. Hsu]. New York: Chapman and Hall/CRC, 2021. <https://doi.org/10.1201/9780429030888>
21. Dickhaus T. Simultaneous statistical inference. With applications in the life sciences. Heidelberg: Springer Berlin, 2014. <https://doi.org/10.1007/978-3-642-45182-9>
22. Gabriel K.R. Simultaneous test procedures – some theory of multiple comparisons // The Annals of Mathematical Statistics. 1969. Vol. 40. No. 1. P. 224–250. <https://doi.org/10.1214/aoms/1177697819>
23. Cliff N. Complete orders from incomplete data: Interactive ordering and tailored testing // Psychological Bulletin. 1975. Vol. 82. No. 2. P. 289–302. <https://doi.org/10.1037/h0076373>
24. Cliff N. Dominance statistics: Ordinal analyses to answer ordinal questions // Psychological Bulletin. 1993. Vol. 114. No. 3. P. 494–509. <https://doi.org/10.1037/0033-2909.114.3.494>
25. Donoghue J.R. Implementing Shaffer’s multiple comparison procedure for a large number of groups // Recent Developments in Multiple Comparison Procedures (Eds. Benjamini, Y., Bretz, F. and Sarkar, S.), Institute of Mathematical Statistics, Beachwood, Ohio, 2004. P. 1–23.

Об авторах

Колданов Петр Александрович

д.ф.-м.н.;

старший научный сотрудник, лаборатория алгоритмов и технологий анализа сетевых структур, Национальный исследовательский университет «Высшая школа экономики», Россия, 603093, г. Н. Новгород, ул. Родионова, 136;

E-mail: pkoldanov@hse.ru

ORCID: 0000-0001-5961-0282

Колданов Владимир Александрович

к.ф.-м.н.;

доцент, кафедра математики и вычислительной техники, Нижегородский государственный инженерно-экономический университет, Россия, 606340, Нижегородская область, г. Княгинино, ул. Октябрьская, д. 22А, корпус 2;

E-mail: vlad.kold@gmail.com

ORCID: 0009-0006-1924-746X

Nonparametric procedure for comparing the performance of divisions of a network organization

Petr A. Koldanov^a

E-mail: pkoldanov@hse.ru

Vladimir A. Koldanov^b

E-mail: vlad.kold@gmail.com

^a Laboratory of Algorithms and Technologies for Network Analysis, HSE University
Address: 136, Rodionova St., Nizhny Novgorod 603093, Russia

^b Nizhny Novgorod State University of Engineering and Economics
Address: 22A, building 2, Oktyabrskaya St., Nizhny Novgorod Region, Knyaginino 606340, Russia

Abstract

To solve the problem of comparative efficiency analysis of branch operations for a small volume of randomly observed data, a non-parametric approach is relevant, since it does not require a probabilistic model of observations. Comparing the results of the non-parametric approach with the results obtained within the traditionally used Gaussian model is also relevant. Additionally, obtaining a consistent comparison of a group (of no less than three) branches is important. Currently, the non-parametric approach and the corresponding comparison with the known results of solving the problem considered in this work obtained within the framework of the normal model are absent. In addition, insufficient attention is paid to the search for methods of obtaining consistent solutions. This work to some extent fills these gaps. This work uses non-parametric statistical methods and theory of simultaneous hypothesis testing to address these problems. This paper proposes a procedure for comparative analysis of the efficiency of several units within a network organization with a small volume of observations based on the Mann–Whitney tests. We carry out a comparison of the results obtained from the proposed non-parametric procedure with results based on extensions of Student's t-tests. We propose a method for reducing the number of compatibility problems based on the search for an appropriate significance level. We provide an example of a fully consistent comparison of the efficiency of branch operations.

Keywords: network organization, efficiency of branch operations, Mann–Whitney tests, incompatibility problem

Citation: Koldanov P.A., Koldanov V.A. (2024) Nonparametric procedure for comparing the performance of divisions of a network organization. *Business Informatics*, vol. 18, no. 1, pp. 52–64. DOI: 10.17323/2587-814X.2024.1.52.64

References

1. Aleskerov F.T., Belousova V. (2007) Effective development of the branch network of a commercial bank. *Upravleniye v kreditnoy organizatsii (Management in a credit institution)*, no. 6, pp. 23–34 (in Russian).
2. Kryukov A.M. (2010) Efficiency analysis of branch and division activities is a prerequisite for business stability. *Strategic decisions and risk management*, vol. 4, pp. 84–87 (in Russian). <https://doi.org/10.17747/2078-8886-2010-4-84-87>
3. Myznikova M.A. (2022) Quality of strategic management under ambiguity: Assessment within the framework of sustainable development. *Business Informatics*, vol. 16, no. 3, pp. 36–52. <https://doi.org/10.17323/2587-814X.2022.3.36.52>
4. Koldanov A.P., Koldanov P.A. (2012) Multiple decision procedures for the analysis of higher school entry selection results. *Business Informatics (Biznes-informatika)*, vol. 19, no. 1, pp. 24–31 (in Russian).

5. Koldanov P.A. (2013) Efficiency analysis of branch network. *Models, Algorithms, and Technologies for Network Analysis* (eds. B.I. Goldengorin, V.A. Kalyagin, P.M. Pardalos). Springer Proceedings in Mathematics & Statistics, vol. 59, pp. 71–83. https://doi.org/10.1007/978-1-4614-8588-9_5
6. Hettmansperger T. (1987) *Statistical Inference Based on Ranks*. Moscow: Finance and Statistics (in Russian).
7. Kendall M. (1975) *Rank correlations*. Moscow: Statistics (in Russian).
8. Hollander M., Wolfe D. (1983) *Nonparametric statistical methods*. Moscow: Finance and Statistics (in Russian).
9. Lehmann E.L. (1975) *Nonparametrics: Statistical methods based on ranks*. San Francisco: Hoden-Day. <https://doi.org/10.1002/zamm.19770570922>
10. Kendall M., Stuart A. (1973) *Statistical Inferences and relationships*. Moscow: Science (in Russian).
11. Jordan M.I. (2004) Graphical models. *Statistical Science*, vol. 19, no. 1, pp. 140–155. <https://doi.org/10.1214/088342304000000026>
12. Newman M.E.J. (2010) *Networks. An introduction*. New York: Oxford University Press.
13. Opsahl T., Panzarasa P. (2009) Clustering in weighted networks. *Social Networks*, vol. 31, no. 2, pp. 155–163.
14. Horvath S. (2011) *Weighted network analysis. Applications in genomics and systems biology*. New York: Springer. <https://doi.org/10.1007/978-1-4419-8819-5>
15. Li S., He J., Zhuang Y. (2010) A network model of the interbank market. *Physica A: Statistical Mechanics and its Applications*, vol. 389, no. 24, pp. 5587–5593. <https://doi.org/10.1016/j.physa.2010.08.057>
16. Boginski V., Butenko S., Pardalos P.M. (2005) Statistical analysis of financial networks. *Computational Statistics & Data Analysis*, vol. 48, no. 2, pp. 431–443. <https://doi.org/10.1016/j.csda.2004.02.004>
17. Lehmann E.L. (1957) A theory of some multiple decision problems, I. *The Annals of Mathematical Statistics*, vol. 28, no. 1, pp. 1–25. <https://doi.org/10.1214/aoms/1177707034>
18. Bretz F., Hothorn T., Westfall P. (2010) *Multiple comparisons using R*. New York: Chapman and Hall/CRC. <https://doi.org/10.1201/9781420010909>
19. Hochberg Y., Tamhane A.C. (1987) *Multiple comparison procedures*. New York: John Wiley & Sons. <https://doi.org/10.1002/9780470316672>
20. *Handbook of Multiple Comparisons* (2021) (eds. X. Cui, T. Dickhaus, Y. Ding, J.C. Hsu). New York: Chapman and Hall/CRC. <https://doi.org/10.1201/9780429030888>
21. Dickhaus T. (2014) *Simultaneous statistical inference. With applications in the life sciences*. Heidelberg: Springer Berlin. <https://doi.org/10.1007/978-3-642-45182-9>
22. Gabriel K.R. (1969) Simultaneous test procedures – some theory of multiple comparisons. *The Annals of Mathematical Statistics*, vol. 40, no. 1, pp. 224–250. <https://doi.org/10.1214/aoms/1177697819>
23. Cliff N. (1975) Complete orders from incomplete data: Interactive ordering and tailored testing. *Psychological Bulletin*, vol. 82, no. 2, pp. 289–302. <https://doi.org/10.1037/h0076373>
24. Cliff N. (1993) Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin*, vol. 114, no. 3, pp. 494–509. <https://doi.org/10.1037/0033-2909.114.3.494>
25. Donoghue J.R. (2004) Implementing Shaffer's multiple comparison procedure for a large number of groups. *Recent Developments in Multiple Comparison Procedures* (Eds. Benjamini, Y., Bretz, F. and Sarkar, S.), Institute of Mathematical Statistics, Beachwood, Ohio, pp. 1–23.

About the authors

Petr A. Koldanov

Dr. Sci. (Phys.-Math.);

Senior Research Fellow, Laboratory of Algorithms and Technologies for Networks Analysis, HSE University (campus in Nizhny Novgorod), 136, Rodionova St., Nizhny Novgorod 603093, Russia;

E-mail: pkoldanov@hse.ru

ORCID: 0000-0001-5961-0282

Vladimir A. Koldanov

Cand. Sci. (Phys.-Math.);

Associate Professor, Department of Mathematics and Computer Science, Nizhny Novgorod State University of Engineering and Economics, 22A, building 2, Oktyabrskaya St., Nizhny Novgorod Region, Knyaginino 606340, Russia;

E-mail: vlad.kold@gmail.com

ORCID: 0009-0006-1924-746X

DOI: 10.17323/2587-814X.2024.1.65.78

Determinants of an auditor's continuance intention with respect to use of the Audit Tools and Linked Archives System (ATLAS): A model of extended expectation confirmation

Kurniasari Novi Hardanti* 

E-mail: kurniasarinovih@gmail.com

Sutrisno T. 

E-mail: sutrisno@ub.ac.id

Erwin Saraswati 

E-mail: erwin@ub.ac.id

Arum Prastiwi 

E-mail: arum@ub.ac.id

Brawijaya University

Address: Veteran Street, Malang 65145, Indonesia

Abstract

The aim of this study was to examine the determinants of the continuance intention with respect to use of the Audit Tools and Linked Archives System (ATLAS) by employing survey methods. These determinants are developed from an Expectation Confirmation Model (ECM). The sample of this study is auditors who use ATLAS in public accounting firms in Indonesia. As many as 356 data points can be processed using smartPLS. This study revealed that perceived usefulness, confirmation, information quality, top management commitment and satisfaction affected the auditor's intentions when using ATLAS. The implications of this study are (1) Public accounting firms must provide full support to auditors in using ATLAS and equip auditors through training so auditors understand that using ATLAS is very useful; (2) IAPI must pay attention to outputs that are complete, good and appropriate so that the auditor is satisfied when using ATLAS. The auditor has a tendency to continue using ATLAS if he is satisfied.

* Corresponding Author

Keywords: continuance intention, audit tools, ECM, satisfaction, perceived usefulness, confirmation, information quality, top management commitment

Citation: Hardanti K.N., Sutrisno T., Saraswati E., Prastiwi A. (2024) Determinants of an auditor's continuance intention with respect to use of the Audit Tools and Linked Archives System (ATLAS): A model of extended expectation confirmation. *Business Informatics*, vol. 18, no. 1, pp. 65–78. DOI: 10.17323/2587-814X.2024.1.65.78

Introduction

The rapid development of information systems in the digital era has brought changes in various sectors. One such sector is auditing. The audit field has used information technology in the work process. The information system used by auditors in Indonesia is the Audit Tools and Linked Archives System (ATLAS).

ATLAS is one solution to improve auditor performance. Some of the benefits of ATLAS are, first, assisting the auditor in carrying out and documenting the audit process. Second, it makes it easier for the auditor to prepare audit working papers, so that the audit working papers he prepares are more systematic. Third, it reduces auditor errors when carrying out the entire audit procedure. The use of working papers processed on a computer has existed for a long time among auditors, but the ATLAS implementation is a new thing for auditors.

Auditors are not yet required to use ATLAS in preparing working papers [1]. However, the successful implementation and success of ATLAS is expected with confidence by the Ministry of Finance and Indonesian Association of Public Accountants (IAPI). In fact, not all implementations of ATLAS are in accordance with the government's expectations. Based on the results of [2], the implementation of ATLAS does not improve auditor performance and also does not improve audit quality. One reason is the low perception of the ease of using ATLAS that makes auditors reluctant to use the ATLAS system [3].

Technological and human factors are important components in the implementation of technology adoption such as ATLAS. ATLAS implementation is not considered to be in accordance with the expectations of the Indonesian Association of Public Accountants (IAPI). This is one of the causes of the human factor. Implementation failure occurs because of the reluctance or rejection of individuals within the company towards the implementation of technology [4]. Aldholay et al. [5] showed that the failure occurred due to aspects of individual behavior in the organization. Information system implementation will be successful if the information system is accepted by its users [6].

This study will focus on the determinants of the continuance intention to use ATLAS. This model is proposed to facilitate understanding of the factors that impact ATLAS acceptance. Bhattacharjee [4] created a model, namely the Expectation Confirmation Model (ECM). ECM is different from other information system acceptance concepts. ECM focuses on the continued use of information systems, thereby providing solid explanations and long-term projections of behavior.

ECM validated in various types of studies with many sample characteristics and different countries. Several research results using the ECM model include [7–11] support for the main model of ECM, although not all of the studies that have been conducted show results consistent with previous studies. The inconsistency of the results of this study is due to differences in the characteristics of the sample and the context of the information system studied.

The ECM model has several advantages over the previous theory. However, the ECM model has not included the technological context and organizational context. The ECM model only focuses on individual contexts: confirmation constructs, satisfaction, and perceived usefulness. The researcher finds that ECM does not include the technological context and organizational context in its model.

To improve ECM, this study will incorporate both the technological context and the organizational context. The technological context is represented by system quality, information quality and service quality which had previously been formulated by [12]. In the success model, system quality measures technical success, namely the accuracy and efficiency of the system in producing information. Information quality measures semantic success, namely the success of information in conveying meaning. Service quality measures the success of the level of effectiveness, namely the influence of the information generated by the information system. The selected organizational context is top management commitment. Top management commitment is support from management (a public accountant firm) for auditors to use ATLAS.

Based on the analysis above, the aim of this study is to investigate the relationship between perceived usefulness, confirmation, information quality, design quality, service quality, top management commitment and satisfaction with the intention to continue using ATLAS among auditors in Indonesia.

This study has made its contribution in two aspects, namely the contribution to the development of theory and practice. The results of this study contribute to theory and practice related to scientific development.

- ◆ The theory contribution in this study is in terms of construct development, model development and providing empirical evidence about acceptance and ATLAS.
- ◆ The practical contribution in this study is for systems analysts and governments. The results of this study can be applied to the practice of designing and implementing ATLAS systems, so that in developing a system one can pay attention to the constructs of confirmation, satisfaction and perceived usefulness, system quality as well as top management commitment. For the government, the results of this study provide input for the Ministry of Finance and IAPI so that in making decisions to implement a system, it can pay attention to aspects of user acceptance.

1. The comprehensive theoretical basis

In this section, the concept of Expected Confirmation Theory (ECM) is introduced before the discussion regarding the relationship between related variables.

1.1. Expectation Confirmation Model

Studies on the use after adoption in the area of information systems began since [4] proposed the Expectation Confirmation Model (ECM). ECM is a model for using an information system after adoption. ECM is a development model based on Expectation Confirmation Theory (ECT) developed by [13]. The concept of ECT is integrated in the technology acceptance model (TAM) in information systems and with further refinement to address its theoretical weaknesses, [4] added a perceived usefulness variable. Perceived usefulness is felt after using an information system, not before using the system.

ECM is a theoretical model of information system sustainability hypothesizing that expectations followed by initial acceptance lead to confirmation by comparing anticipated consequences. When actual performance is

confirmed, users of the information system will be satisfied and potentially this leads to continued use. In contrast to other consumer acceptance models that focus on the first use of a new information system, ECM focuses on continuing use after acceptance of the information system, and provides solid explanations and long-term scale projections of user behavior [4].

Continuance is a form of behavior after adoption [14]. Information system continuance is basically the same as repurchasing behavior, namely following the initial decision and being influenced by initial use and potentially stopping use [4]). The use of advanced information systems can be determined by the intention resulting from certain reasons. The user's interest to continue using a technology can be referred to as continuance intention.

1.2. Research model and hypothesis

Figure 1 shows the research model developed from the previously described literature. This study is different from previous studies. This study carries out construct development in the ECM proposed by [4] which includes the technological context and organizational context. The constructs in the context of this technology are system quality, information quality and service quality. System quality and information quality will determine the attitude of system users [15]. System quality, information quality and service quality will also determine individual satisfaction in using information systems [16]. Several researchers, including [12, 17–19] in an information system success model with different indicators and measurements, have used the technological context. Organizational context is also needed because it is closely related to individual attitudes in using the system. The selected organizational context is top management commitment. Top management commitment will greatly affect employee satisfaction because with this role, management can monitor the quality of the system that is being implemented within the company [20]. This model is expected to be able to prove what factors can influence satisfaction and interest in continuing to use ATLAS.

1.2.1. The hypothesis of perceived usefulness

Perceived usefulness is defined as a person's belief that he will use the system if the system has utility value [21]. The construct of perceived usefulness was originally described through TAM by [21]. TAM relates the perceived usefulness construct to the behavioral interest construct. TAM found that perceptions of usability and ease of use

are prominent beliefs that influence information systems' acceptance behavior in the field of technology. ECM modifies the relationship between perceived usefulness and satisfaction. Perceived usefulness is expected to influence user satisfaction after working with the information system. The results of [4] study state that perceived usefulness has a positive effect on satisfaction. The more useful an information system, the stronger the individual satisfaction in using the information system.

The following studies [22–24] also strengthen ECM that the positive influence between perceived usefulness constructs on satisfaction in various types of applications. Based on the research results of [22], perceived usefulness has a positive impact towards satisfaction. Oghuma et al. [24] found perceived usefulness is the construct that has the greatest influence on their model. From the previous explanation, the formulation of the hypothesis is as follows.

H1: Perceived usefulness has a positive effect on satisfaction for using the ATLAS system.

1.2.2. The hypothesis of confirmation

Confirmation is defined as the perception between the conformity of expectations with reality after someone uses the system [4]. The research results of [4] show that confirmation has a positive effect toward satisfaction. Confirmation is the strongest predictor of satisfaction. The more in line with the performance of the information system with user expectations, the higher the individual satisfaction in using the information system. The positive relationship between confirmation and satisfac-

tion formulated and proven by [4] in ECM, has been supported by many studies. From the previous explanation, the formulation of the hypothesis is as follows.

H2: Confirmation has a positive effect toward satisfaction for using the ATLAS system.

1.2.3. The hypothesis of information quality

Information quality is characteristics system output when used by user [12]. This model shows that the quality of information has a positive impact toward satisfaction. This positive relationship formulated and proven by [12] in the DeLone and McLean Information Systems (IS) success model has been supported by many studies [25–29]. From the previous explanation, the formulation of the hypothesis is as follows.

H3: Information quality has a positive effect on satisfaction for users of the ATLAS system.

1.2.4. The hypothesis of system quality

System quality has a positive impact toward satisfaction because when a system has good performance in providing the information needed by users, users will be satisfied when working with the system [12]. This work shows that system quality has a positive impact on satisfaction. The positive relationship has been supported [25–28]. From the previous explanation, the formulation of the hypothesis is as follows.

H4: System quality has a positive effect on satisfaction for using the ATLAS system.

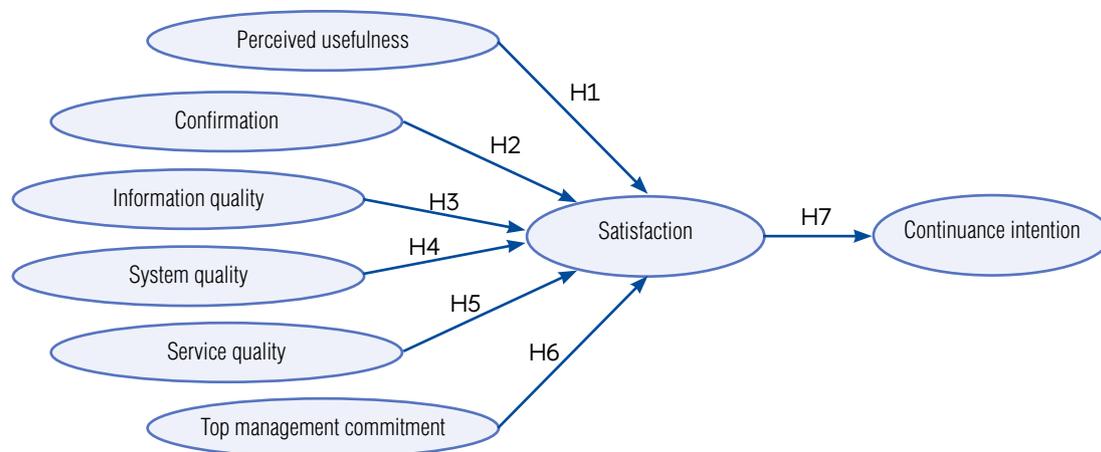


Fig. 1. Model research.

1.2.5. The hypothesis of service quality

Service quality is defined as the quality of support by users when using the system directly. There is a positive effect between service quality and satisfaction because if users get good enough support when using a system, for example when there are difficulties using the system (errors), then there are IT staff who are ready to help quickly. In this case, user satisfaction will increase in using the system [12]. This success model shows that service quality has a positive effect toward satisfaction. The positive relationship formulated by [12] in the DeLone and McLean IS success model has been supported by many studies [26–28]. From the previous explanation, the formulation of the hypothesis is as follows.

H5: Service quality has a positive effect on satisfaction for users of the ATLAS system.

1.2.6. The hypothesis of top management commitment

Sinha et al. [30] define top management commitment as top management's active involvement in establishing and monitoring policies, communicating and encouraging employees to achieve their goals. Management plays an important role in the success of an organization. Top management commitment is believed to be able to increase the level of confidence and willingness of employees to complete tasks properly. Addis et al. [20] proved that top management commitment has a positive effect on satisfaction in quality management practices of manufacturing organizations in Ethiopia. From the previous explanation, the formulation of the hypothesis is as follows.

H6: Top management commitment has a positive effect on satisfaction when using the ATLAS system.

1.2.7. The hypothesis of satisfaction on continuance intention

Bhattacharjee [4] presented satisfaction as an emotion related to previous experience by users of information systems. The satisfaction construct is explained through ECM. ECM relates the satisfaction construct to the repeated use intention construct with the logic of thinking that the user's repeated use intention will be determined mainly from the user's satisfaction when using the given system. If the user feels bad emotions when using an information system, then this could be a reason for the user to stop using the information system. ECM proved that satisfaction has a positive effect on intention to use again.

The following studies also strengthen the theory formulated by [4] about a positive influence of the satisfaction construct on intention to use repeatedly in various types of applications by [22, 31, 32]. From the previous explanation, the formulation of the hypothesis is as follows.

H7: Satisfaction has a positive effect on the intention to continue using the ATLAS system.

2. Research method

The population in this study was auditors who use the ATLAS system in their work processes in all public accounting firms (PAF) in Indonesia that were registered with the Indonesian Public Accountants Association (IAPI) in 2022.

The unit of analysis is an auditor who works in PAF and has used the ATLAS system in his work. The sampling technique in this study is non-probability sampling. Non-probability sampling is a technique in which all members of the population do not have the same opportunity to be sampled. This study uses judgment sampling in the sampling procedure, namely the sample is taken based on existing criteria on auditors who have used ATLAS in their work process.

The data collection technique used is a survey data collection technique. The survey method of this study is a mail survey. The Mail Survey conducted in this study was by distributing questionnaires via WhatsApp, Instagram, Email and Telegram. Mail surveys have a potential problem anticipated by researchers, namely low return rates. To ensure that researchers get a high rate of return on questionnaires, certain steps are taken.

The statistical method used is Partial Least Square (PLS). By using SmartPLS ver. 4.0 M3. PLS is a structural equation modeling that can solve multiple regression problems with small samples and multicollinearity.

$$S = \beta_1 PU + \beta_2 C + \beta_3 IQ + \beta_4 SYQ + \beta_5 SEQ + \beta_6 TMC + e$$

$$CI = \beta_7 S + e,$$

where:

S – satisfaction;

PU – perceived usefulness;

C – confirmation;

IQ – information quality;

SYQ – system quality;

SEQ – services quality;

TMC – top management commitment;

CI – continuance intention;

β_i – coefficient construct;

e – error.

The constructs in this study are information quality, system quality, service quality, top management commitment, confirmation, satisfaction, perceived usefulness and intention to continue using it. The instruments used to measure constructs in this study are instruments that have been used in previous studies [4, 10, 20, 27, 33], making it possible to increase the validity and reliability of the measurements. The measurement uses a Likert scale from 1 to 7 which has the following meanings: (1) Strongly Disagree, (2) Disagree, (3) Somewhat Disagree, (4) Neutral, (5) Somewhat Agree, (6) Agree, and (7) Strongly Agree.

Before the actual distribution of the questionnaire, the researcher conducted a pilot test with the aim of convincing himself that the questionnaire items were sufficient, correct and understandable to the respondents. Researchers conducted a pre-test as follows:

1. Translate the original English instrument into Indonesian.
2. Ask for help from individuals who are experts in English to translate it back into English. Next, the results of the translation into English will be compared by the researcher with the original instrument so that from this step it is hoped that there will be no differences in the meaning or significance of the translated instrument.
3. Asking negative questions on several items in the questionnaire so that the questionnaire is not biased and there is a form of control over the questionnaire.
4. Discuss the meaning of each indicator with several ATLAS users.

Then, confirmation of the public accountant's ability to accept research and confirmation that the public accounting firm is indeed using ATLAS is carried out first by phone. After the survey was approved, the researcher brought and took the questionnaire directly to the research location for the reason of increasing the response rate from those returning the questionnaire.

The pilot test was conducted on 34 students of PPAK and Master of Accounting as well as apprentice students at PAF who had used ATLAS in the audit process. After testing, the results show that all items in the questionnaire are valid and reliable. Once it is known that the items in the questionnaire are valid and reliable, the researcher distributes the questionnaires to the real respondents in the field.

3. Results

All questionnaires that can be used and processed are 356 questionnaires. The majority of respondents, 39.89%, were junior auditors, 51.40% were senior auditors and 8.71% were supervisors. The number of fe-

male respondents was 33.99% and 66.01% were male. The majority of respondents were aged 31–40 years and the majority of respondents' highest education was the bachelor's degree. The majority of respondents had more than 5–10 years of work experience and more than 10 years of experience using computers.

3.1. Validity and reliability testing

PLS model evaluation is done by evaluating the outer model and inner model. Evaluation of the outer model is carried out by conducting convergent validity tests, discriminant validity tests and reliability tests. In order to fulfil convergent validity, all constructs must have AVE value of more than 0.5. All constructs in *Table 1* have an AVE value more than 0.5. It can be concluded that convergent validity has been fulfilled.

In order to qualify for discriminant validity, all indicators must have a factor loading value of more than 0.7. The test results show that all indicators in *Table 2* have a value of more than 0.7; it means that discriminant validity is fulfilled.

Table 1.

Output quality criteria overview model with SmartPLS

	AVE	Composite reliability	R-squared	Cronbach's Alpha
Perceived usefulness	0.837	0.954		0.935
Confirmation	0.828	0.935		0.896
Information quality	0.835	0.953		0.934
System quality	0.809	0.944		0.921
Services quality	0.798	0.941		0.916
Top management commitment	0.830	0.936		0.898
Satisfaction	0.840	0.940	0.969	0.905
Continuance intention	0.807	0.926	0.577	0.881

All constructs must have a Cronbach's alpha value composite reliability must be greater than 0.7. Tests show that all constructs in *Table 2* have Cronbach's alpha values and composite reliability of more than 0.7. This means that all constructs are reliable.

Table 2.

Output cross loading with SmartPLS

Indicator	C	CI	IQ	PU	S	SEQ	SYQ	TMC
C1	0.899							
C2	0.910							
C3	0.921							
CI1		0.893						
CI2		0.890						
CI3		0.912						
IQ1			0.915					
IQ2			0.906					
IQ3			0.920					
IQ4			0.914					
PU1				0.920				
PU2				0.900				
PU3				0.921				
PU4				0.917				
S1					0.913			
S2					0.908			
S3					0.928			
SEQ1						0.889		
SEQ2						0.886		
SEQ3						0.902		
SEQ4						0.898		
SYQ1							0.891	
SYQ2							0.895	
SYQ3							0.912	
SYQ4							0.899	
TMC1								0.905
TMC2								0.910
TMC3								0.918

Evaluation of the inner model or structural model is a stage for evaluating the relationship between constructs. The Inner Model is evaluated with R2 and statistical testing. The results of the structural model testing using the R2 value and the significance test through the path coefficient values for each path. This is the results of the significance test for the path coefficient values or *t*-values for each path.

Through Fig. 2, it can be concluded that the R2 values for satisfaction and interest in continuing to use the system are 0.97 and 0.58. This explains that the constructs of perceptions of usefulness, confirmation, information quality, system quality, service quality and top management commitment are able to explain the satisfaction construct by 97%; the remaining 3% is explained through other variables outside the proposed model. Continuing intention to use ATLAS can explain the satisfaction construct by as much as 58%.

The significance test is obtained through the results of the *p*-value. The hypothesis is supported if the *p*-value has a value less than 0.05. Based on Table 3 it can be concluded that H1a, H1b, H1c, H1f, and H2 are supported because they have *p*-values less than 0.05, while H1d and H1e are rejected because they have *p*-values greater than 0.05.

This study succeeded in proving that perceptions of usefulness, confirmation, information quality and top management commitment have a positive impact on satisfaction when using the ATLAS system. Meanwhile, the satisfaction construct has a positive effect on the intention to continue using ATLAS. This study failed to prove the effect of system quality and service quality on satisfaction. The following is a discussion for each construct.

4. Discussion and conclusion

This study examines the relationship between perceived usefulness, confirmation, information quality, system quality, service quality, top management commitment, satisfaction and intention to continue using ATLAS based on the development of the ECM model [4] and [12]. The focus of this research was to investigate the auditor's continuance intention for use of ATLAS in the context of information systems. This study confirms that (a) the more useful a system is, the higher the auditor's satisfaction with its use; (b) the more expectations are confirmed in ATLAS, the higher the perceived satisfaction created; (c) the more complete the information available on the ATLAS, the higher the auditor's satisfaction; (d) the more useful the information provided in ATLAS, the higher the auditor's satisfaction in using it; (e) the more satisfied the auditor is in using ATLAS, the

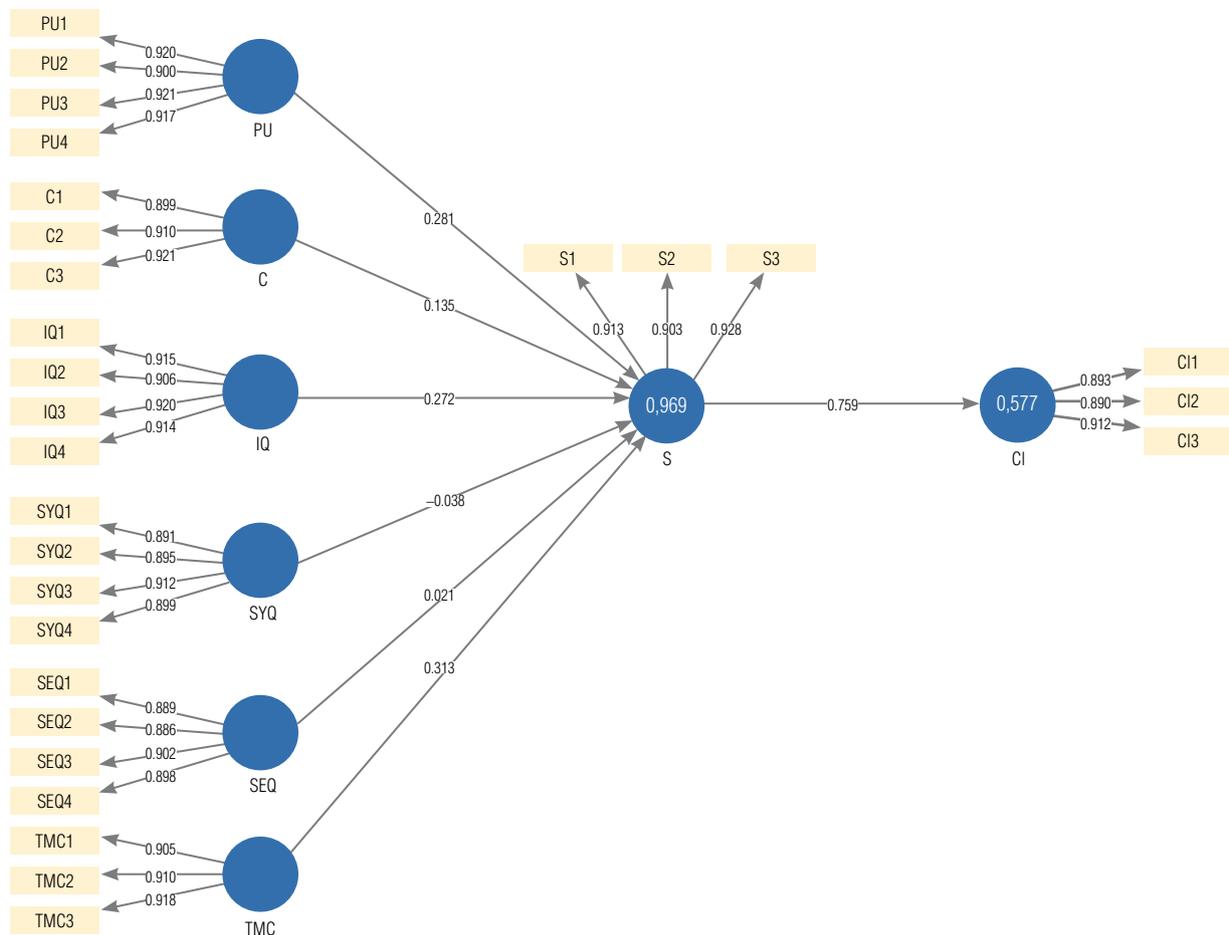


Fig. 2. Measurement model output.

higher the interest in continuing its use in the future. The findings support the following implications in this study.

First, the greater the perception of usability, the higher the satisfaction when using ATLAS. The outcomes from this research are consistent with [22, 34–36]. This empirical evidence is consistent with ECM that perceived usefulness has an effect on satisfaction. These results imply that perceived usefulness can be a predictor of perceived satisfaction in using the system for auditors. This empirical evidence has implications that the auditor is satisfied using ATLAS if the auditor feels ATLAS is useful in supporting his performance. Thus, PAF must equip auditors to provide more training so that auditors understand that the use of ATLAS is very useful and can support performance. When auditors feel they have the ability and control after receiving sufficient provision from the training process, they will tend to use the ATLAS system well and have high confidence.

Second, auditor expectations that have been confirmed in ATLAS positively influence auditor satisfac-

tion as supported by previous research [22–24]. This empirical evidence is consistent with ECM that confirmation has an impact on satisfaction. These results imply that expectations can be a predictor of perceived satisfaction when auditors use the system. Good experience and maximum service for auditors must be provided to meet the expectations of auditors when using ATLAS. Thus, IAPI as the manufacturer of ATLAS must provide the best performance from ATLAS so that ATLAS can be used continuously by auditors.

Third, satisfaction increases when the quality of information is fulfilled. The same findings were found in previous studies [25–29]. This empirical evidence is consistent with the information systems achievement model, that the quality of information affects satisfaction. These results indicate that the quality of information can play an important role in understanding satisfaction in using a technology. Thus, ATLAS must be designed to be useful, understandable, attractive and believable. High quality information must be maintained and improved to provide even more satisfying results.

Table 3.

Results of hypothesis testing with SmartPLS

Hypotesis	Path	Original sample	<i>t</i> -statistic	<i>p</i> -value	Conclusion
H1	$PU \rightarrow S$	0.281	3.629	0.000	Supported
H2	$C \rightarrow S$	0.135	2.024	0.022	Supported
H3	$IQ \rightarrow S$	0.272	3.606	0.000	Supported
H4	$SYQ \rightarrow S$	-0.038	0.678	0.249	Rejected
H5	$SEQ \rightarrow S$	0.021	0.374	0.354	Rejected
H6	$TMC \rightarrow S$	0.313	4.436	0.000	Supported
H7	$S \rightarrow CI$	0.759	23.145	0.000	Supported

Fourth, top management commitment affects auditor satisfaction when using ATLAS. In other words, when management provides full support to the auditors to use ATLAS, the auditors will be satisfied with the new system. This finding is in accordance with the findings of the previous study [20]. Top management commitment must be continuously improved. This empirical evidence has implications that auditors have a high level of satisfaction when auditors in PAF get full support from top management. Top management commitment is believed to be able to increase the level of confidence and willingness of employees to complete tasks properly. Thus, in designing a system, PAF does not only think about the greatness of the implemented system, but also must provide full support to the auditors in the application of ATLAS.

In addition, perceived satisfaction positively influences intention to continue using ATLAS. The outcomes of this study are consistent with studies accompanied by [22, 32]. This empirical evidence is consistent with the ECM and information system success models, that satisfaction influences the intention to continue using the system. The findings in this study indicate that satisfaction is a strong factor influencing intention to continue using ATLAS. That is, even though auditors initially have a positive perception of ATLAS, the auditors will not continue to use ATLAS if they are dissatisfied with the system. Moreover, the use of the ATLAS system is voluntary, so the auditor has the power to continue using ATLAS or not. Thus, PAF management needs to think about motivating its employees so that they are more motivated to use the ATLAS system properly. When auditors are motivated to use the ATLAS system, they will tend to continue using the ATLAS system.

On the other hand, system quality and service quality do not affect satisfaction when using ATLAS. This finding is inconsistent with previous research [25–28]. This empirical evidence is supported by the study of [37]. This empirical evidence is suspect because the satisfaction felt by the auditor is subjective so that the perception of each auditor can be different. This study proves that the relationship between information quality and satisfaction in the DeLone and McLean model is not sufficient to explain the phenomenon of someone accepting or rejecting use of the ATLAS system.

With the development of technology, ATLAS was introduced as an information system that combines the entire audit process from the initial engagement to the independent auditor's report. The significance of this study is to identify the factors that influence the intention to continue using ATLAS. From the results, this study suggests a strategy to increase interest in continuing the use of ATLAS by verifying the causal relationship between related factors, and the use of ATLAS for auditors is voluntary.

The limitation of this study is that it only uses one construct related to the organizational context, namely top management commitment. Based on the limitations of this study, further studies can be carried out, namely further studies can add other constructs in the organizational context that were not examined in this study, for example organizational culture. ■

Acknowledgements

The work described in this paper was supported by Educational Fund Management Institution – Ministry of Finance.

References

1. Krismonanda C., Widyastuti S., Nugraheni R. (2021) Analysis of the application of the audit tools and linked archives system (ATLAS) to the financial statement audit process. *Jurnal Penelitian Ekonomi dan Akuntansi (JPENSI)*, vol. 6, no. 3 (in Indonesian). <https://doi.org/10.30736/jpensi.v6i3.794>
2. Ismail A.M. (2019) *The effect of using information technology on auditor performance (empirical study at public accounting firms in Makassar city)*. Thesis, Economics Faculty in State University of Makassar. Available at: <http://eprints.unm.ac.id/14065/> (accessed 8 Feb. 2024).
3. Setiawan I., Alfie A.A., Astuti W.B. (2022) Application of the ATLAS application, competence and independence of auditors and audit quality of public accounting firms in Semarang. *The Journal of Accounting and Finance (JAFIN)*, vol. 1, no. 1, pp. 67–77 (in Indonesian).
4. Bhattacharjee A. (2001) Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*, vol. 25, no. 3, pp. 351–370.
5. Aldholay A.H., Isaac O., Abdullah Z., Ramayah T. (2018) The role of transformational leadership as a mediating variable in DeLone and McLean information system success model: The context of online learning usage in Yemen. *Telematics and Informatics*, vol. 35, no. 5, pp. 1421–1437. <https://doi.org/10.1016/j.tele.2018.03.012>
6. Jeyaraj A., Zadeh A.H. (2020) Evolution of information systems research: Insights from topic modelling. *Information & Management*, vol. 57, no. 4, 103207. <https://doi.org/10.1016/j.im.2019.103207>
7. Joo Y.J., Park S., Shin E.K. (2017) Students' expectation, satisfaction, and continuance intention to use digital textbooks. *Computers in Human Behavior*, vol. 69, pp. 83–90. <https://doi.org/10.1016/j.chb.2016.12.025>
8. Yousaf A., Mishra A., Taheri B., Kesgin M. (2021) A cross-country analysis of the determinants of customer recommendation intentions for over-the-top (OTT) platforms. *Information and Management*, vol. 58, no. 8, 103543. <https://doi.org/10.1016/j.im.2021.103543>
9. Hsu C.L., Chen M.C. (2021) Advocating recycling and encouraging environmentally friendly habits through gamification: An empirical investigation. *Technology in Society*, vol. 66, 101621. <https://doi.org/10.1016/j.techsoc.2021.101621>
10. Si H., Duan X., Cheng L., Zhang Z. (2022) Determinants of consumers' continuance intention to use dynamic ride-sharing services. *Transportation Research Part D: Transport and Environment*, vol. 104, 103201. <https://doi.org/10.1016/j.trd.2022.103201>
11. Maduku D.K., Thusi P. (2023) Understanding consumers' mobile shopping continuance intention: New perspectives from South Africa. *Journal of Retailing and Consumer Services*, vol. 70, 103185. <https://doi.org/10.1016/j.jretconser.2022.103185>
12. DeLone W.H., McLean E.R. (2003) The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, vol. 19, no. 4, pp. 9–30. <https://doi.org/10.1080/07421222.2003.11045748>
13. Oliver R.L. (1980) A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, vol. 17, no. 4, pp. 460–469. <https://doi.org/10.2307/3150499>
14. Limayem M., Hirt S.G., Cheung C.M.K. (2007) How habit limits the predictive power of intention: The case of information systems continuance. *MIS Quarterly*, vol. 31, no. 4, pp. 705–737. <https://doi.org/10.2307/25148817>
15. Chen Y.-C., Hu L.-T., Tseng K.-C., Juang W.-J., Chang C.-K. (2019). Cross-boundary e-government systems: Determinants of performance. *Government Information Quarterly*, vol. 36, no. 3, pp. 449–459. <https://doi.org/10.1016/j.giq.2019.02.001>
16. Mohamadali N.A.K.S., Garibaldi J.M. (2010) A novel evaluation model of user acceptance of software technology in healthcare sector. Proceedings of the *Third International Conference on Health Informatics BIOSTEC, Valencia, Spain*, pp. 392–397. <https://doi.org/10.5220/0002695703920397>
17. Alyoussef I.Y. (2023) Acceptance of e-learning in higher education : The role of task-technology fit with the information systems success model. *Heliyon*, vol. 9, no. 3, e13751. <https://doi.org/10.1016/j.heliyon.2023.e13751>
18. Roky H., Meriouh A.Y. (2015) Evaluation by users of an industrial information system (XPPS) based on the DeLone and McLean model for IS success. *Procedia Economics and Finance*, vol. 26, pp. 903–913. [https://doi.org/10.1016/s2212-5671\(15\)00903-x](https://doi.org/10.1016/s2212-5671(15)00903-x)
19. Alnemer H.A. (2022) Determinants of digital banking adoption in the Kingdom of Saudi Arabia: A technology acceptance model approach. *Digital Business*, vol. 2, no. 2, 100037. <https://doi.org/10.1016/j.digbus.2022.100037>
20. Addis S., Dvivedi A., Beshah B. (2019) Quality management as a tool for job satisfaction improvement in low-level technology organizations: The case of Ethiopia. *Production Planning & Control*, vol. 30, no. 8, pp. 665–681. <https://doi.org/10.1080/09537287.2019.1574510>
21. Davis F.D., Bagozzi R.P., Warshaw P.R. (1989) User acceptance of computer-technology: A comparison of two theoretical models. *Management Science*, vol. 35, no. 8, pp. 982–1003. Available at: <http://www.jstor.org/stable/2632151> (accessed 8 Feb. 2024).
22. Hadji B., Degoulet P. (2016) Information system end-user satisfaction and continuance intention: A unified modelling approach. *Journal of Biomedical Informatics*, vol. 61, pp. 185–193. <https://doi.org/10.1016/j.jbi.2016.03.021>
23. Sasongko D.T., Handayani P.W., Satria R. (2022) Analysis of factors affecting continuance use intention of the electronic money application in Indonesia. *Procedia Computer Science*, vol. 197, pp. 42–50. <https://doi.org/10.1016/j.procs.2021.12.116>
24. Oghuma A.P., Chang Y., Libaque-Saenz C.F., Park M.C., Rho J.J. (2015) Benefit-confirmation model for post-adoption behavior of mobile instant messaging applications: A comparative analysis of KakaoTalk and Joyn in Korea. *Telecommunications Policy*, vol. 39, no. 8, pp. 658–677. <https://doi.org/10.1016/j.telpol.2015.07.009>

25. Lee K.C., Chung N. (2009) Understanding factors affecting trust in and satisfaction with mobile banking in Korea: A modified DeLone and McLean's model perspective. *Interacting with Computers*, vol. 21, nos. 5–6, pp. 385–392. <https://doi.org/10.1016/j.intcom.2009.06.004>
26. Pamugar H., Winarno W.W., Najib W. (2014) Evaluation model of success and acceptance of e-learning information systems in government training institutions. *Scientific Journal of Informatics*, vol. 1, no. 1 (in Indonesian). <https://doi.org/10.15294/sji.v1i1.3638>
27. Tam C., Oliveira T. (2016) Understanding the impact of m-banking on individual performance: DeLone & McLean and TTF perspective. *Computers in Human Behavior*, vol. 61, pp. 233–244. <https://doi.org/10.1016/j.chb.2016.03.016>
28. Wang Y.S., Liao Y.W. (2008) Assessing eGovernment systems success: A validation of the DeLone and McLean model of information systems success. *Government Information Quarterly*, vol. 25, no. 4, pp. 717–733. <https://doi.org/10.1016/j.giq.2007.06.002>
29. Akrong G.B., Yunfei S., Owusu E. (2022) Development and validation of an improved DeLone-McLean IS success model – application to the evaluation of a tax administration ERP. *International Journal of Accounting Information Systems*, vol. 47, 100579. <https://doi.org/10.1016/j.accinf.2022.100579>
30. Sinha N., Garg A.K., Dhingra S., Dhall N. (2016) Mapping the linkage between organizational culture and TQM: The case of Indian auto component industry. *Benchmarking: An International Journal*, vol. 23, no. 1, pp. 208–235. <https://doi.org/10.1108/BIJ-12-2014-0112>
31. Mardiana S., Tjakraatmadja J.H., Aprianingsih A. (2015) Validating the conceptual model for predicting intention to use as part of Information System success model: The case of an Indonesian government agency. *Procedia Computer Science*, vol. 72, pp. 353–360. <https://doi.org/10.1016/j.procs.2015.12.150>
32. Prakash A.V., Das S., Pillai K.R. (2021) Understanding digital contact tracing app continuance: Insights from India. *Health Policy and Technology*, vol. 10, no. 4, 100573. <https://doi.org/10.1016/j.hlpt.2021.100573>
33. Huang Y.M. (2019) Examining students' continued use of desktop services: Perspectives from expectation-confirmation and social influence. *Computers in Human Behavior*, vol. 96, pp. 23–31. <https://doi.org/10.1016/j.chb.2019.02.010>
34. Alraimi K.M., Zo H., Ciganek A.P. (2015) Understanding the MOOCs continuance: The role of openness and reputation. *Computers & Education*, vol. 80, pp. 28–38. <https://doi.org/10.1016/j.compedu.2014.08.006>
35. Cheng Y.-M. (2014) Why do users intend to continue using the digital library? An integrated perspective. *Aslib Journal of Information Management*, vol. 66, no. 6, pp. 640–662. <https://doi.org/10.1108/AJIM-05-2013-0042>
36. Chen S.C., Yen D.C., Peng S.-C. (2018) Assessing the impact of determinants in e-magazines acceptance: An empirical study. *Computer Standards & Interfaces*, vol. 57, pp. 49–58. <https://doi.org/10.1016/j.csi.2017.11.004>
37. Wara L.S., Kalangi L., Gamaliel H. (2021) Testing the success model of the Delone and Mclean information systems on the examination application system (SIAP) at the supreme audit agency of the Republic of Indonesia representative for North Sulawesi province. *Jurnal Riset Akuntansi dan Auditing "GOODWILL"*, vol. 12, no. 1, pp. 1–15 (in Indonesian). <https://doi.org/10.35800/jjs.v12i1.31885>

About the authors

Kurniasari Novi Hardanti

Doctoral Student, Department of Accounting, Brawijaya University, Veteran Street, Malang 65145, Indonesia;
E-mail: kurniasarinovih@gmail.com
ORCID: 0009-0004-9505-3224

Sutrisno T.

Doctor of Accounting (PhD), Professor;
Professor, Department of Accounting, Brawijaya University, Veteran Street, Malang 65145, Indonesia;
E-mail: sutrisno@ub.ac.id
ORCID: 0000-0002-0147-5070

Erwin Saraswati

Doctor of Accounting (PhD), Professor;
Professor, Department of Accounting, Brawijaya University, Veteran Street, Malang 65145, Indonesia;
E-mail: erwin@ub.ac.id
ORCID: 0000-0002-8799-8605

Arum Prastiwi

Doctor of Accounting (PhD);
Senior Lecturer, Department of Accounting, Brawijaya University, Veteran Street, Malang 65145, Indonesia;
E-mail: arum@ub.ac.id
ORCID: 0000-0002-0993-8850

Факторы, определяющие намерение аудитора продолжить использование системы аудита и связанных архивов (ATLAS): Модель расширенного подтверждения ожиданий

Харданти К.Н.*

E-mail: kurniasarinovih@gmail.com

Сутрисно Т.

E-mail: sutrisno@ub.ac.id

Сарасвати Э.

E-mail: erwin@ub.ac.id

Прастиви А.

E-mail: arum@ub.ac.id

Университет Бравиджая

Адрес: Индонезия, 65145, Маланг, улица Ветеранов

Аннотация

Статья посвящена исследованию факторов, определяющих намерение аудиторов продолжать работу с системой аудита и связанных архивов (Audit Tools and Linked Archives System, ATLAS). В ходе исследования использованы опросы и модель подтверждения ожиданий (Expectation Confirmation Model, ECM). Выборка, использованная в ходе исследования, охватывает аудиторов, использующих систему ATLAS в бухгалтерских фирмах Индонезии. Для этого был использован инструментарий SmartPLS, позволяющий обрабатывать до 356 элементов набора данных. Исследование показало, что на намерения аудиторов в отношении использования ATLAS повлияли такие факторы, как воспринимаемая полезность, подтверждение, качество информации, удовлетворенность и приверженность высшего руководства. В результате можно сделать следующие выводы. Во-первых, бухгалтерские фирмы должны оказывать полную поддержку аудиторам в использовании ATLAS и обучать аудиторов, чтобы они ощущали полезность системы. Во-вторых, Ассоциация бухгалтеров Индонезии (Indonesian Association of Public Accountants, IAPI) должна обращать внимание на то, чтобы результаты были полными, качественными и уместными, чтобы аудитор был удовлетворен использованием ATLAS. В случае удовлетворенности аудиторы склонны продолжать использование системы ATLAS.

Ключевые слова: намерение продолжения использования, инструменты аудита, модель подтверждения ожиданий, удовлетворенность, воспринимаемая полезность, подтверждение, качество информации, приверженность высшего руководства

Цитирование: Hardanti K.N., Sutrisno T., Saraswati E., Prastiwi A. (2024) Determinants of an auditor's continuance intention with respect to use of the Audit Tools and Linked Archives System (ATLAS): A model of extended expectation confirmation // Business Informatics. 2023. Vol. 18. No. 1, P. 65–78. DOI: 10.17323/2587-814X.2024.1.65.78

* Автор, ответственный за переписку

Литература

1. Krismonanda C., Widyastuti S., Nugraheni R. Analysis of the application of the audit tools and linked archives system (ATLAS) to the financial statement audit process // *Jurnal Penelitian Ekonomi dan Akuntansi (JPENSI)*. 2021. Vol. 6. No. 3 (in Indonesian). <https://doi.org/10.30736/jpensiv6i3.794>
2. Ismail A.M. (2019) The effect of using information technology on auditor performance (empirical study at public accounting firms in Makassar city) // Thesis, Economics Faculty in State University of Makassar. Available at: <http://eprints.unm.ac.id/14065/> (accessed 8 Feb. 2024).
3. Setiawan I., Alfie A.A., Astuti W.B. Application of the ATLAS application, competence and independence of auditors and audit quality of public accounting firms in Semarang // *The Journal of Accounting and Finance (JAFIN)*. 2022. Vol. 1. No. 1. P. 67–77 (in Indonesian).
4. Bhattacherjee A. Understanding information systems continuance: An expectation-confirmation model // *MIS Quarterly*. 2001. Vol. 25. No. 3. P. 351–370.
5. Aldholay A.H., Isaac O., Abdullah Z., Ramayah T. The role of transformational leadership as a mediating variable in DeLone and McLean information system success model: The context of online learning usage in Yemen // *Telematics and Informatics*. 2018. Vol. 35. No. 5. P. 1421–1437. <https://doi.org/10.1016/j.tele.2018.03.012>
6. Jeyaraj A., Zadeh A.H. Evolution of information systems research: Insights from topic modelling // *Information & Management*. 2020. Vol. 57. No. 4, 103207. <https://doi.org/10.1016/j.im.2019.103207>
7. Joo Y.J., Park S., Shin E.K. Students' expectation, satisfaction, and continuance intention to use digital textbooks // *Computers in Human Behavior*. 2017. Vol. 69. P. 83–90. <https://doi.org/10.1016/j.chb.2016.12.025>
8. Yousaf A., Mishra A., Taheri B., Kesgin M. A cross-country analysis of the determinants of customer recommendation intentions for over-the-top (OTT) platforms // *Information and Management*. 2021. Vol. 58. No. 8. 103543. <https://doi.org/10.1016/j.im.2021.103543>
9. Hsu C.L., Chen M.C. Advocating recycling and encouraging environmentally friendly habits through gamification: An empirical investigation // *Technology in Society*. 2021. Vol. 66. 101621. <https://doi.org/10.1016/j.techsoc.2021.101621>
10. Si H., Duan X., Cheng L., Zhang Z. Determinants of consumers' continuance intention to use dynamic ride-sharing services // *Transportation Research Part D: Transport and Environment*. 2022. Vol. 104. 103201. <https://doi.org/10.1016/j.trd.2022.103201>
11. Maduku D.K., Thusi P. Understanding consumers' mobile shopping continuance intention: New perspectives from South Africa // *Journal of Retailing and Consumer Services*. 2023. Vol. 70. 103185. <https://doi.org/10.1016/j.jretconser.2022.103185>
12. DeLone W.H., McLean E.R. The DeLone and McLean model of information systems success: A ten-year update // *Journal of Management Information Systems*. 2003. Vol. 19. No. 4. P. 9–30. <https://doi.org/10.1080/07421222.2003.11045748>
13. Oliver R.L. A cognitive model of the antecedents and consequences of satisfaction decisions // *Journal of Marketing Research*. 1980. Vol. 17. No. 4. P. 460–469. <https://doi.org/10.2307/3150499>
14. Limayem M., Hirt S.G., Cheung C.M.K. How habit limits the predictive power of intention: The case of information systems continuance // *MIS Quarterly*. 2007. Vol. 31. No. 4. P. 705–737. <https://doi.org/10.2307/25148817>
15. Chen Y.-C., Hu L.-T., Tseng K.-C., Juang W.-J., Chang C.-K. (). Cross-boundary e-government systems: Determinants of performance // *Government Information Quarterly*. 2019. Vol. 36. No. 3. P. 449–459. <https://doi.org/10.1016/j.giq.2019.02.001>
16. Mohamadali N.A.K.S., Garibaldi J.M. A novel evaluation model of user acceptance of software technology in healthcare sector // *Proceedings of the Third International Conference on Health Informatics BIOSTEC, Valencia, Spain*. 2010. P. 392–397. <https://doi.org/10.5220/0002695703920397>
17. Alyoussef I.Y. Acceptance of e-learning in higher education : The role of task-technology fit with the information systems success model // *Heliyon*. 2023. Vol. 9. No. 3. e13751. <https://doi.org/10.1016/j.heliyon.2023.e13751>
18. Roky H., Meriouh A.Y. Evaluation by users of an industrial information system (XPPS) based on the DeLone and McLean model for IS success // *Procedia Economics and Finance*. 2015. Vol. 26. P. 903–913. [https://doi.org/10.1016/s2212-5671\(15\)00903-x](https://doi.org/10.1016/s2212-5671(15)00903-x)
19. Alnemer H.A. Determinants of digital banking adoption in the Kingdom of Saudi Arabia: A technology acceptance model approach // *Digital Business*. 2022. Vol. 2. No. 2. 100037. <https://doi.org/10.1016/j.digbus.2022.100037>
20. Addis S., Divedi A., Beshah B. Quality management as a tool for job satisfaction improvement in low-level technology organizations: The case of Ethiopia // *Production Planning & Control*. 2019. Vol. 30. No. 8. P. 665–681. <https://doi.org/10.1080/09537287.2019.1574510>
21. Davis F.D., Bagozzi R.P., Warshaw P.R. User acceptance of computer-technology: A comparison of two theoretical models // *Management Science*. 1989. Vol. 35. No. 8. P. 982–1003. Available at: <http://www.jstor.org/stable/2632151> (accessed 8 Feb. 2024).
22. Hadji B., Degoulet P. Information system end-user satisfaction and continuance intention: A unified modelling approach // *Journal of Biomedical Informatics*. 2016. Vol. 61. P. 185–193. <https://doi.org/10.1016/j.jbi.2016.03.021>
23. Sasongko D.T., Handayani P.W., Satria R. Analysis of factors affecting continuance use intention of the electronic money application in Indonesia // *Procedia Computer Science*. 2022. Vol. 197. P. 42–50. <https://doi.org/10.1016/j.procs.2021.12.116>
24. Oghuma A.P., Chang Y., Libaque-Saenz C.F., Park M.C., Rho J.J. Benefit-confirmation model for post-adoption behavior of mobile instant messaging applications: A comparative analysis of KakaoTalk and Joyn in Korea // *Telecommunications Policy*. 2015. Vol. 39. No. 8. P. 658–677. <https://doi.org/10.1016/j.telpol.2015.07.009>

25. Lee K.C., Chung N. Understanding factors affecting trust in and satisfaction with mobile banking in Korea: A modified DeLone and McLean's model perspective // *Interacting with Computers*. 2009. Vol. 21. Nos. 5–6. P. 385–392. <https://doi.org/10.1016/j.intcom.2009.06.004>
26. Pamugar H., Winarno W.W., Najib W. Evaluation model of success and acceptance of e-learning information systems in government training institutions // *Scientific Journal of Informatics*. 2014. Vol. 1. No. 1 (in Indonesian). <https://doi.org/10.15294/sji.v1i1.3638>
27. Tam C., Oliveira T. Understanding the impact of m-banking on individual performance: DeLone & McLean and TTF perspective // *Computers in Human Behavior*. 2016. Vol. 61. P. 233–244. <https://doi.org/10.1016/j.chb.2016.03.016>
28. Wang Y.S., Liao Y.W. Assessing eGovernment systems success: A validation of the DeLone and McLean model of information systems success // *Government Information Quarterly*. 2008. Vol. 25. No. 4. P. 717–733. <https://doi.org/10.1016/j.giq.2007.06.002>
29. Akrong G.B., Yunfei S., Owusu E. Development and validation of an improved DeLone-McLean IS success model – application to the evaluation of a tax administration ERP // *International Journal of Accounting Information Systems*. 2022. Vol. 47. 100579. <https://doi.org/10.1016/j.accinf.2022.100579>
30. Sinha N., Garg A.K., Dhingra S., Dhall N. Mapping the linkage between organizational culture and TQM: The case of Indian auto component industry // *Benchmarking: An International Journal*. 2016. Vol. 23. No. 1. P. 208–235. <https://doi.org/10.1108/BIJ-12-2014-0112>
31. Mardiana S., Tjakraatmadja J.H., Apriansih A. Validating the conceptual model for predicting intention to use as part of Information System success model: The case of an Indonesian government agency // *Procedia Computer Science*. 2015. Vol. 72. P. 353–360. <https://doi.org/10.1016/j.procs.2015.12.150>
32. Prakash A.V., Das S., Pillai K.R. Understanding digital contact tracing app continuance: Insights from India // *Health Policy and Technology*. 2021. Vol. 10. No. 4. 100573. <https://doi.org/10.1016/j.hlpt.2021.100573>
33. Huang Y.M. Examining students' continued use of desktop services: Perspectives from expectation-confirmation and social influence // *Computers in Human Behavior*. 2019. Vol. 96. P. 23–31. <https://doi.org/10.1016/j.chb.2019.02.010>
34. Alraimi K.M., Zo H., Ciganek A.P. Understanding the MOOCs continuance: The role of openness and reputation // *Computers & Education*. 2015. Vol. 80. P. 28–38. <https://doi.org/10.1016/j.compedu.2014.08.006>
35. Cheng Y.-M. Why do users intend to continue using the digital library? An integrated perspective // *Aslib Journal of Information Management*. 2014. Vol. 66. No. 6. P. 640–662. <https://doi.org/10.1108/AJIM-05-2013-0042>
36. Chen S.C., Yen D.C., Peng S.-C. Assessing the impact of determinants in e-magazines acceptance: An empirical study // *Computer Standards & Interfaces*. 2018. Vol. 57. P. 49–58. <https://doi.org/10.1016/j.csi.2017.11.004>
37. Wara L.S., Kalangi L., Gamaliel H. Testing the success model of the Delone and Mclean information systems on the examination application system (SIAP) at the supreme audit agency of the Republic of Indonesia representative for North Sulawesi province // *Jurnal Riset Akuntansi dan Auditing "GOODWILL"*. 2021. Vol. 12. No. 1. P. 1–15 (in Indonesian). <https://doi.org/10.35800/jrs.v12i1.31885>

Об авторах

Харданги Курниасари Нови

аспирант, кафедра бухгалтерского учета, Университет Бравиджайя, Индонезия, 65145, Маланг, улица Ветеранов;

E-mail: kurniasarinovih@gmail.com

ORCID: 0009-0004-9505-3224

Сутрисно Т.

доктор (PhD), профессор;

профессор, кафедра бухгалтерского учета, Университет Бравиджайя, Индонезия, 65145, Маланг, улица Ветеранов;

E-mail: sutrisno@ub.ac.id

ORCID: 0000-0002-0147-5070

Сарасвати Эрвин

доктор (PhD), профессор;

профессор, кафедра бухгалтерского учета, Университет Бравиджайя, Индонезия, 65145, Маланг, улица Ветеранов;

E-mail: erwin@ub.ac.id

ORCID: 0000-0002-8799-8605

Прастиви Арум

доктор (PhD);

Старший преподаватель, кафедра бухгалтерского учета, Университет Бравиджайя, Индонезия, 65145, Маланг, улица Ветеранов;

E-mail: arum@ub.ac.id

ORCID: 0000-0002-0993-8850

DOI: 10.17323/2587-814X.2024.1.79.88

Product information recognition in the retail domain as an MRC problem

Tho Chi Luong^a 

E-mail: tholc@vnu.edu.vn

Oanh Thi Tran^{b*} 

E-mail: oanhtt@gmail.com

^a Institut Francophone International, Vietnam National University, Hanoi
Address: E5, 144 Xuan Thuy St., Cau Giay Dist., Hanoi, Vietnam

^b International School, Vietnam National University, Hanoi
Address: G7, 144 Xuan Thuy St., Cau Giay Dist., Hanoi, Vietnam

Abstract

This paper presents the task of recognizing product information (PI) (i.e., product names, prices, materials, etc.) mentioned in customer statements. This is one of the key components in developing artificial intelligence products to enable businesses to listen to their customers, adapt to market dynamics, continuously improve their products and services, and improve customer engagement by enhancing effectiveness of a chatbot. To this end, natural language processing (NLP) tools are commonly used to formulate the task as a traditional sequence labeling problem. However, in this paper, we bring the power of machine reading comprehension (MRC) tasks to propose another, alternative approach. In this setting, determining product information types is the same as asking “Which PI types are referenced in the statement?” For example, extracting product names (which corresponds to the label PRO_NAME) is cast as retrieving answer spans to the question “Which instances of product names are mentioned here?” We perform extensive experiments on a Vietnamese public dataset. The experimental results show the robustness of the proposed alternative method. It boosts the performance of the recognition model over the two robust baselines, giving a significant improvement. We achieved 92.87% in the F1 score on recognizing product descriptions at Level 1. At Level 2, the model yielded 93.34% in the F1 score on recognizing each product information type.

Keywords: product information recognition, MRC framework, retail domain, large-language models, vBERT, vELECTRA

Citation: Luong T.C., Tran O.T. (2024) Product information recognition in the retail domain as an MRC problem. *Business Informatics*, vol. 18, no. 1, pp. 79–88. DOI: 10.17323/2587-814X.2024.1.79.88

* Corresponding Author

Introduction

Product Information (PI) is all the data about the products that a company sells. It includes a product's technical specifications, size, materials, prices, photos, schematics, etc. E-commerce requires companies to collect clear basic PI that consumers can actually understand and place orders. Without PI¹, the product could not be found and sold online at all.

Recognizing PI is crucial for widespread applications. For example, in the e-commerce field, it is vital to integrate this component to develop AI products like chatbots [1] to enhance customers' experiences. Chatbots significantly help reduce customer support costs, while increasing customer satisfaction with an AI chatbot that can recognize customer intents, instantly provide information, on any channel, and never take a day off. Identifying PI also helps to better analyze the sentiments [2] in comments/reviews of their customers. With PI, we can associate specific sentiments with different aspects of a product to analyze customer sentiments and opinions from reviews. This would improve the product and help make better marketing campaigns.

Conventionally, the task of PI recognition is formulated as a sequence labeling problem. It is a supervised learning problem that involves predicting an output sequence for a given input sequence. Most research in this field has proposed different machine learning approaches using handcrafted features or neural network approaches [3, 4] without using handcrafted features.

In this paper, we bring the power of machine reading comprehension (MRC) to this task. This idea is significantly inspired by a recent trend of transforming natural language processing (NLP) tasks to answering MRC questions. Specifically, Levy et al. [5] formulated the relation extraction task as a QA task. McCann et al. [6] transformed the tasks of summarization or sentiment analysis into question answering. For example, the task of summarization can be formalized as answering the question "*What is the summary?*" Li et al. [7] formalized the task of entity-relation extraction as a multi-turn question-answering problem.

So far, most current work has focused on high-resource languages. Therefore, to narrow the gap between low and high-resource languages, this paper also targets the Vietnamese language. This paper proposed an alternative way to extract PI by modeling it as a MRC prob-

lem. We conduct many extensive experiments on a public dataset by Tran et al. [1] and the results demonstrate that this approach introduces a significant performance boost over robust existing systems. The main contribution of this paper can be highlighted as follows:

- ◆ We proposed an alternative method to recognize PI by tailoring the MRC framework to suit the specific requirement of the task.
- ◆ We have conducted extensive experiments to verify the effectiveness of the proposed approach on a public Vietnamese benchmark dataset².

The remainder of this paper is organized as follows. Related work is presented in Section 1. Section 2 shows how to formulate the task as an MRC problem and then describes the method for generating questions, as well as the model architecture. Section 3 describes the experimental setups, experimental results, and some discussions. Finally, we conclude the paper and figure out some future lines of work.

1. Related work

This section first presents the work on PI identification, and then describes related work about the machine reading comprehension (MRC) tasks.

1.1. Work on PI recognition

Information retrieval chatbots are widely applied as assistants, to support customers formulate their requirements about the products they want when placing an order online. In order to develop such chatbots, most current systems use information retrieval techniques [8, 9] or a concept-based knowledge model [10] to identify product information details mentioned by their customers. Towards building task-oriented chatbots, Yan et al. [11] presented a general solution for online shopping. To extract PI asked by customers, the system matched the question to basic PI using the DSSM model. Unfortunately, these studies do not support customers who are performing orders online, and some external data resources exploited in their research are intractable in many actual applications.

Most work has been done for rich-resource languages such as English and Chinese; work for poor-resource languages is much rarer. In Vietnam, there is only one work focusing on recognizing PI types in the retail domain. Specifically, Tran et al. [1] introduced a study on

¹ In this paper, we consider seven types of PI types including *categories*, *attributes*, *extra-attributes*, *brands*, *packsizes*, *numbers*, and *unit-of-measurements (uoms)*.

² https://github.com/oanh84/PI_dataset/tree/main

understanding what the users say in chatbot systems. They concentrated on recognizing PI types implied in users' statements. In that work, they modeled the task as a sequence labelling problem and then explored different deep neural networks such as CNNs and LSTMs to solve the task.

1.2. Work on MRC

MRC refers to the ability of a machine learning model to understand and extract relevant information from written texts. It is similar to how a human reader would do this and accurately answer questions related to the content of the texts. The power of the MRC model is evaluated by the ability to extract the correct answer to the user question.

Many published novel datasets inspired a large number of new neural MRC models. In the past several years, we have witnessed many neural network models created such as BERT [12, 16] RoBERTa [13] and XLNet [10]. Many large language models utilize transformers [14] to pre-train representations by considering both the left and right context across all layers. Due to their remarkable success, this approach has progressively evolved into a mainstream method, involving pre-training large language models on extensive corpora and subsequently fine-tuning them on datasets specific to the target domain. Deep learning neural networks, particularly those based on transfer learning, are widely employed to address diverse challenges in natural language processing (NLP). Transfer learning methods emphasize the retention of data and knowledge acquired during the exploration of one problem, then applying this acquired knowledge to address different yet related questions. The effectiveness of these cutting-edge neural network models is noteworthy. For example, Lithe SOTA neural network models by Therasa et al. [12] has already exceeded human performance over many related MRC benchmark datasets.

In this paper, we borrow the idea of MRC to propose another alternative approach to this task. To prove the effectiveness of the approach, we conduct extensive experiments on a public Vietnamese dataset released by Tran et al. [1]. The results showed a new SOTA result over the traditional existing techniques.

2. Recognizing PI as an MRC problem

In this section, we first formulate the task of recognizing PI as an MRC problem. Then, we show the method to generate questions/queries for finding the answers (which could be the product information instances) ap-

pearing in the users' input utterances. Finally, the model architecture is presented and explained in more details.

2.1. Problem formulation

Given a users' statement x including n syllables $\{x_1, x_2, \dots, x_n\}$, we need to build a model to identify every product information mentioned in x . For each instance of a product information type found in x we assign a label y to it. Here, y belongs to one of the pre-defined PI list including *product names*, *product size*, *product unit-of-measurements*, *product attribute*, *product brand*, *product number*, and *product extra attribute*.

To exploit the MRC approach, it is necessary to recast the task as an MRC problem. To this end, we construct triples of questions, answers, and contexts $\{q_{pi}, x_{start:end}, \dots, x\}$ for each label pi mentioned in x as follows:

- ◆ x : the user's statement.
- ◆ $x_{start:end}$: the product information mentioned in x . It is a sequence of syllables within x identified by the specified start and end indexes $\{x_{start}, \dots, x_{end}\}$, where the condition $start \leq end$ holds true. Expert knowledge is required to annotate this data.
- ◆ q_{pi} : the question to ask the model to find $x_{start:end}$ corresponding to the label pi . This is a natural question consisting of m syllables $\{q_1, q_2, \dots, q_m\}$. Various approaches will be investigated in order to generate such questions.

This exactly establishes the triple (*Question, Answer, Context*) to be exploited in the proposed framework. And now, the task can be recast as an MRC problem as follows: Given a collection of k training examples $\{q^i, x_{start:end}^i, x^i\}$ (where $i = 1..k$). The purpose is to train a predictor which receives the statement x and the corresponding question q , and outputs the answer $x_{start:end}$. It is formulated as the following formula:

$$x_{start:end}^i = f(q^i, x^i).$$

2.2. Question generation

Each PI is associated with a specific question generated by combining the predefined templates and its training example values. It is a natural language question. In order to provide more prior knowledge about the label, we add some examples to the questions so that the model can recognize answers easier. These examples are randomly withdrawn from the training data set. Some typical generated questions for product information types are shown in *Table 1*.

Table 1.

Some questions generated for each PI types using templates

No.	Product information types	Generated questions
1	Product names	Which product names are mentioned in the text such as <i>smoothies</i> and <i>cakes</i> ?
2	Product sizes	Which product sizes are mentioned in the text such as <i>big</i> and <i>small</i> ?
3	Product colors	Which product colors are mentioned in the text such as <i>green</i> and <i>blue</i> ?
4	Product uoms	Which product uoms are mentioned in the text such as <i>cup</i> and <i>cm</i> ?
5	Product attributes	Which product attributes are mentioned in the text such as <i>extra ice</i> and <i>little sugar</i> ?
6	Product extra attributes	Which product extra attributes are mentioned in the text such as <i>strawberry flavor</i> and <i>orange flavor</i> ?
7	Product brand	Which brands are mentioned in the text such as <i>Samsung</i> and <i>Toyota</i> ?

Here we just provide some examples of each product information type to help the model find all of its instances appearing in the input statement.

2.3. Model architecture

Figure 1 shows the general architecture which includes several main components. The model in this framework is built with a pre-trained large language model (i.e., BERT encoder) and a network designed to produce candidate for *start* and *end* indexes, along with their associated confidence scores indicating the likelihood of being product information.

Given the question q_{pi} , the purpose is to find the text span $x_{start:end}$ categorized as the product information type pi . In the first step, q_{pi} and x are concatenated to establish the string $\{[CLS], q_1, q_2, \dots, q_m; [SEP], [x_1, x_2, \dots, x_n]\}$, where $[CLS]$ and $[SEP]$ are special tokens employed in the conventional pre-trained LLMs. Then, the string is inputted into BERT to generate a contextual representation matrix $E \in R^{n \times d}$, (here d indicates the vector dimension of the final layer). Here, we do not make any prediction for the question, so its final vector representation is ignored.

2.4. Producing the indexes of start:end

To this end, we follow the method proposed by Li et al. [7] to build two corresponding binary classifiers. These two classifiers estimate the probability of each token to be a start or an end index using a softmax function. Specifically, $p^{start}, p^{end} \in R^n$ indicate the vectors that show the likelihood in probability of each token being the start index and end index, respectively:

$$[p^{start}, p^{end}] = \text{softmax}(EW + B),$$

where both W and $B \in R^{n^2}$ are trainable parameters.

Then, a ranked list of potential Product Information (PI) along with corresponding confidence scores is produced by the model. These scores are computed as the sum of the probabilities associated with their *start* and *end* tokens.

In training, the overall objective is to minimize the global loss of three types which are losses for start index, end index and start-end index matching. These losses are simultaneously trained in an end-to-end framework. We use [15] to optimize the loss.

3. Experiments

This section first shows the general information about the public benchmark dataset used for experiments. Then, it tells us about the setups of experiments. Finally, the experimental results and discussion are shown.

3.1. Dataset

In this paper, we used the dataset released by Tran and Luong [1] to perform comparative experiments. This data was collected from a history log of a retail restaurant, some forums and social websites. It was annotated with seven main types of PI which are *product category*, *product attribute*, *product extra-attribute*, *product brand*, *product packsizes*, *product number*, and *product uoms*. Two levels of annotation were provided. At the first level, descriptions of products (Level 1) are extracted. Then, these product descriptions are further decomposed into some detailed PI types (Level 2). An example is given in Table 2.

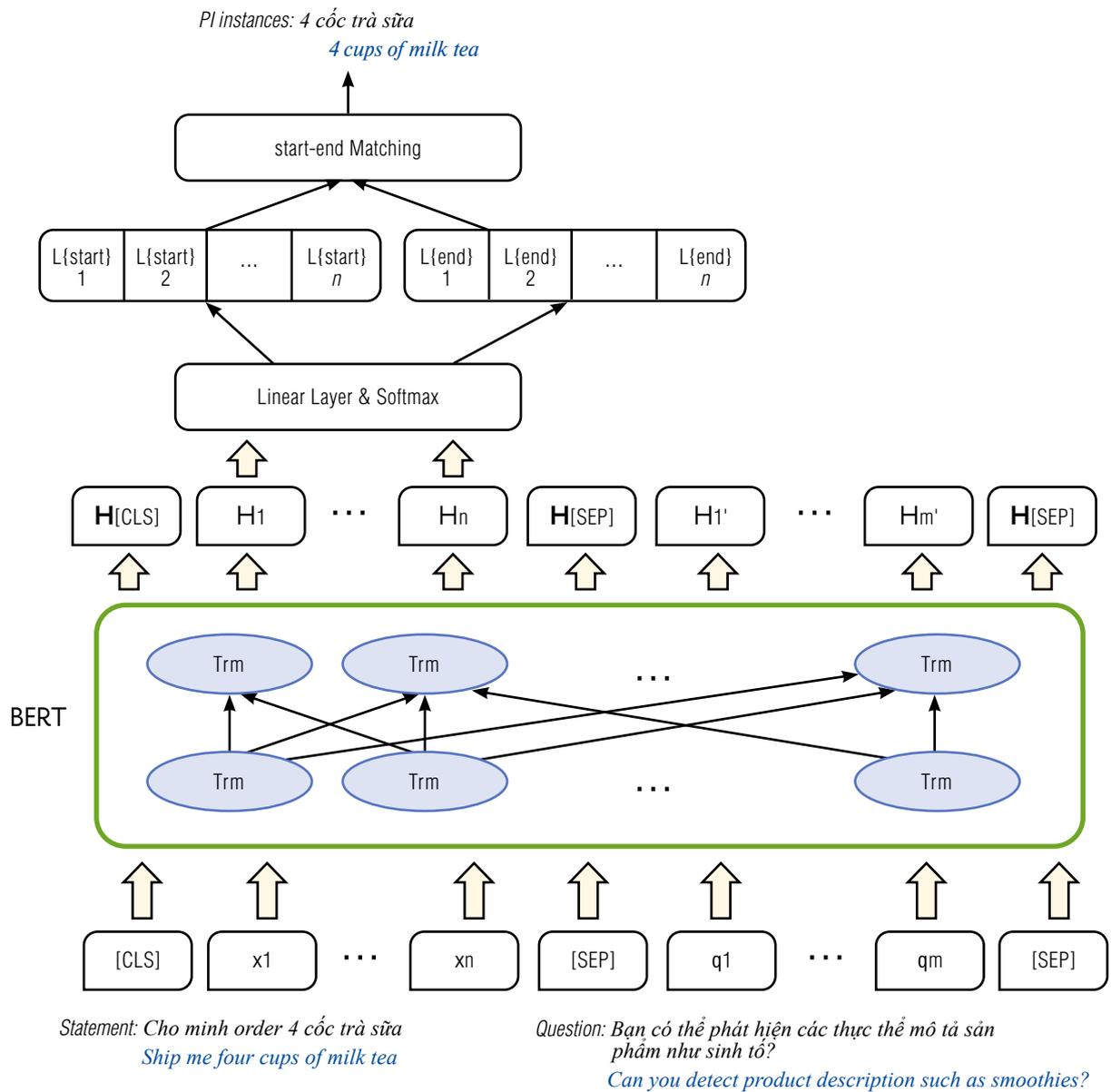


Fig. 1. An architecture using BERT to solve the PI recognition task as an MRC problem. (English translation is given right below the Vietnamese texts).

Table 2.

One example of a user's statement annotated at two levels.

(English translation is provided right after the Vietnamese statement at the first row)

Utterances	<i>Cho em đặt</i>	1	<i>hộp</i>	<i>bánh kem</i>	<i>vị xoài</i>	<i>Cỡ lớn</i>
	<i>Let me order</i>	1	<i>pack</i>	<i>cream cake</i>	<i>mango flavor</i>	<i>big size</i>
Level 1	other	Product description				
Level 2	other	number	uom	category	attribute	size

3.2. Experimental setups

The models are evaluated using popular metrics such as precision, recall, and F1 scores [17]. The best parameters were fine-tuned on development sets. The best values for parameters and hyper-parameters are listed as follows:

- ◆ Train sequence length: 768
- ◆ Number of epochs: 300
- ◆ Batch size: 8
- ◆ Learning rate: $3e-5$
- ◆ Adam epsilon: $1e-8$
- ◆ Max gradient norm: 1.0
- ◆ Bert embedding: 768 dimensions.

We adapted the MRC framework³ to this task and exploited the viBERT⁴ and vELECTRA⁵, a pre-trained large language model optimized for Vietnamese, to build the PI recognition model. In case the pre-trained models are not available to optimize for a specific language, it is also feasible to use a multilingual pre-trained model, such as mBERT (a.k.a multi-lingual BERT) in order to get the vector representations for its sentences. We trained the model on the GPU Tesla V100 SXM2 32GB.

3.3. Experimental results

Tables 3 and 4 show the experimental results of the proposed model in comparison to the two baselines which are BiLSTM-CRF and CNN-CRF [1].

Table 3.

Experimental results of the models at Level 1 – Product descriptions

	Precision	Recall	F1-scores
biLSTM-CRF	89.71	91.35	90.52
CNN-CRF	90.6	91.24	90.91
MRC-viBERT	94.1	91.68	92.87
MRC-vELECTRA	94.5	92.18	93.33

At Level 1, we can see that the MRC approach boosted the performance by a large margin on all evaluation metrics. In comparison to the best baseline CNN-CRF, it enhanced F1 score by nearly 2% in the case of using viBERT and 2.4% in the case of using vELECTRA. This suggested that the MRC approach is very prom-

ising and yields a better performance than other traditional approaches.

At Level 2, in comparison to biLSTM-CRF, it significantly outperformed this baseline in all product information types. The MRC-viBERT approach also slightly increased the F1 score by about 0.3% in comparison the best baseline of CNN-CRF method. Among seven PI types, it achieved a significant improvement over the two baselines by a large margin on three PI types (i.e. *product branch*, *product category*, and *product extra_attribute*). For the type of attribute, the proposed approach got the competitive results. It surpassed biLSTM-CRF, but could not overcome CNN-CRF on the remaining three PI types (i.e., *product packsize*, *product sys_number*, and *product uom*).

Among two types of MRC basing on viBERT and vELECTRA as backbone, we witnessed that MRC-vELECTRA performed slightly better than MRC-viBERT. It increased the performance on four PI types (i.e. *product attribute*, *product branch*, *product category*, and *product extra_attribute*). However, similar to MRC-viBERT, the MRC-vELECTRA also could not surpass CNN-CRF on the remaining three PI types. Overall, in comparison to the best baseline – CNN-CRF, the MRC-vELECTRA increased the F1 score by 0.85%. This result is quite promising.

3.4. Discussion

Looking at the results shown in Table 3 and Table 4, we acknowledge that using the MRC approach yielded higher F1 scores at both levels. This is because the queries/questions generated provide more prior knowledge to guide the identification process of product information.

It can be also seen that the proposed method yielded better performance on recognizing long PI types (such as *product attributes*, *product description*, *product extra attribute*) in comparison to the best baseline – CNN-CRF. This can be explained as follows: the MRC approach captures the sequence information better than CNN. CNN only leverages the local contexts based on n-gram characters and word embeddings. So, it does not have the power of capturing long PI types as compared to the MRC approach. Among two types of word embeddings, the MRC-vELECTRA was slightly better than MRC-viBERT on both two PI levels.

³ <https://github.com/CongSun-dlut/BioBERT-MRC>

⁴ <https://github.com/fpt-corp/viBERT>

⁵ <https://github.com/fpt-corp/vELECTRA>

Table 4.

Experimental results of the models at Level 2 – Product Information Types

PI types	biLSTM-CRF			CNN-CRF			MRC-viBERT			MRC-vELECTRA		
	Pre	Rec	F1									
attribute	93.69	95.63	94.63	95.9	97.24	95.8	95.82	95.25	95.53	96.02	95.71	95.86
brand	82.44	83.24	82.77	89.38	88.64	88.98	92.04	89.90	90.90	92.79	90.65	91.71
category	86.24	88.45	87.32	91.44	91.90	91.67	93.57	93.88	93.72	94.17	93.98	94.07
extra attribute	87.89	86.76	87.26	88.83	86.24	87.39	94.03	88.76	91.31	95.01	89.04	91.93
packsize	85.03	86.82	85.84	91.62	93.14	92.36	92.23	88.77	90.41	93.04	89.21	91.08
sys number	95.24	95.35	95.28	95.88	95.92	95.89	95.29	92.04	93.62	96.12	92.57	94.31
uom	88.80	91.73	90.16	92.12	92.33	92.19	89.16	93.07	91.05	90.01	93.11	91.53
Total	89.39	90.86	90.11	92.95	93.21	93.08	93.69	93.01	93.34	94.11	93.76	93.93

This proposed approach can be generalized to any language. In case BERT is not available to a specific language, we can instead use the mBERT (multi-language BERT) as the backbone.

Conclusion

This paper described the task of identifying product information mentioned by customers' statements in a retail domain. This is a vital step in developing many artificial intelligence commercial products. In contrast to many previous studies, we did not formulate the task as a conventional sequence labeling problem. Instead, we make use of the robustness of MRC tasks to propose an alternative approach. The proposed MRC architecture also leverages the knowledge gained during pre-training a large language model and then applies it to a new, related task – MRC. We performed experiments on a Vietnamese public benchmark dataset to verify the effectiveness of the proposed method. We achieved a new SOTA

result by boosting the recognition performance over the two strong baselines. Specifically, we achieved 93.33% in the F1 score on recognizing product descriptions at Level 1 (upgraded by 2.4%). At Level 2, the model slightly improved the performance and yielded 93.93% in the F1 score on recognizing each product information type by using MRC-vELECTRA. The results also suggested that this approach is more effective in predicting long PI types with high precision.

In the future, we will continue exploring different kinds of generating questions by providing more clues to help find the product information. Furthermore, we will explore alternative robust pre-trained language models to improve the predictive model. ■

Acknowledgements

This paper was funded by the International School, Vietnam National University Hanoi under the project CS.NNC/2021-07.

References

1. Tran O.T., Luong T.C. (2020) Understanding what the users say in chatbots: A case study for the Vietnamese language. *Engineering Applications of Artificial Intelligence*, vol. 87, 103322. <https://doi.org/10.1016/j.engappai.2019.103322>
2. Tran O.T., Bui V.T. (2020) A BERT-based hierarchical model for Vietnamese aspect based sentiment analysis. *Proceedings of the 12th International Conference on Knowledge and System Engineering (KSE), Can Tho, Vietnam, 2–14 November 2020*, pp. 269–274. <https://doi.org/10.1109/KSE50997.2020.9287650>
3. Bui V.T., Tran O.T., Le H.P. (2020) Improving sequence tagging for Vietnamese text using transformer-based neural models. *arXiv:2006.15994*. <https://doi.org/10.48550/arXiv.2006.15994>

4. Lample G., Ballesteros M., Subramanian S., Kawakami K., Dyer C. (2016) Neural architectures for named entity recognition. Proceedings of the *2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, California, June 2016*, pp. 260–270. <https://doi.org/10.18653/v1/N16-1030>
5. Levy O., Seo M., Choi E., Zettlemoyer L. (2017) Zero-shot relation extraction via reading comprehension. *arXiv:1706.04115*. <https://doi.org/10.48550/arXiv.1706.04115>
6. McCann B., Keskar N.S., Xiong C., Socher R. (2018) The natural language decathlon: Multitask learning as question answering. *arXiv:1806.08730*. <https://doi.org/10.48550/arXiv.1806.08730>
7. Li X., Yin F., Sun Z., Li X., Yuan A., Chai D., Zhou M., Li J. (2019) Entity-relation extraction as multi-turn question answering. Proceedings of the *57th Conference of the Association for Computational Linguistics, Florence, Italy, July 2019*, pp. 1340–1350. <https://doi.org/10.18653/v1/P19-1129>
8. Ji Z., Lu Z., Li H. (2014) An information retrieval approach to short text conversation. *arXiv:1408.6988*. <https://doi.org/10.48550/arXiv.1408.6988>
9. Qiu M., Li F., Wang S., Gao X., Chen Y., Zhao W., Chen H., Huang J., Chu W. (2017) AliMe chat: A sequence to sequence and rerank based chatbot engine. Proceedings of the *55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Vancouver, Canada, July 2017*, pp. 498–503. <https://doi.org/10.18653/v1/P17-2079>
10. Goncharova E., Ilvovsky D.I., Galitsky B. (2021) Concept-based chatbot for interactive query refinement in product search. Proceedings of the *9th International Workshop "What can FCA do for Artificial Intelligence?" (FCA4AI 2021)*, vol. 2972, CEUR-WS, pp. 51–58. Available at: <http://ceur-ws.org/Vol-2972/paper5.pdf> (accessed 15 February 2024).
11. Yan Z., Duan N., Chen P., Zhou M., Zhou J., Li Z. (2017) Building task-oriented dialogue systems for online shopping. Proceedings of the *AAAI Conference on Artificial Intelligence*, vol. 31, no. 1. <https://doi.org/10.1609/aaai.v31i1.11182>
12. Liu Y., Ott M., Goyal N., Du J., Joshi M., Chen D., Levy O., Lewis M., Zettlemoyer L., Stoyanov V. (2019) RoBERTa: A robustly optimized BERT pretraining approach. *arXiv:1907.11692*. <https://doi.org/10.48550/arXiv.1907.11692>
13. Yang Z., Dai Z., Yang Y., Carbonell J., Salakhutdinov R.R., Le Q.V. (2019) XLNet: Generalized autoregressive pretraining for language understanding. *33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada, arXiv:1906.08237*. <https://doi.org/10.48550/arXiv.1906.08237>
14. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser Ł., Polosukhin I. (2017) Attention is all you need. *31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, arXiv:1706.03762*. <https://doi.org/10.48550/arXiv.1706.03762>
15. Kingma J., Ba J. (2015) Adam: A method for stochastic optimization. *arXiv:1412.6980*. <https://doi.org/10.48550/arXiv.1412.6980>
16. Devlin J., Chang M.W., Lee K., Toutanova K. (2018) Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv:1810.04805*. <https://doi.org/10.48550/arXiv.1810.04805>
17. Therasa M., Mathivanan G. (2022) Survey of machine reading comprehension models and its evaluation metrics. Proceedings of the *6th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India*, pp. 1006–1013. <https://doi.org/10.1109/ICCMC53470.2022.9754070>

About the authors

Tho Chi Luong

Researcher, Institut Francophone International, Vietnam National University, Hanoi, E5, 144 Xuan Thuy St., Cau Giay Dist., Hanoi, Vietnam;

E-mail: tholc@vnu.edu.vn

ORCID: 0000-0002-7664-705X

Oanh Thi Tran

Associate Professor, PhD;

Lecturer, International School, Vietnam National University, Hanoi, G7, 144 Xuan Thuy St., Cau Giay Dist., Hanoi, Vietnam;

E-mail: oanhtt@gmail.com

ORCID: 0000-0002-3286-3623

Распознавание информации о продуктах в розничной торговле как задача MRC

Луонг Т.Ч.^a

E-mail: tholc@vnu.edu.vn

Тран О.Т.^{b*}

E-mail: oanhtt@gmail.com

^a Международный институт франкоязычных стран, Вьетнамский национальный университет в Ханое
Адрес: Вьетнам, Ханой, Каузяй, Суан Туи, 144, E5

^b Международная школа Вьетнамского национального университета в Ханое
Адрес: Вьетнам, Ханой, Каузяй, Суан Туи, 144, G7

Аннотация

В статье рассматривается задача распознавания информации о продуктах (названий, цен, материалов и т.д.), упомянутой в комментариях клиентов. Эта задача является одной из ключевых при разработке продуктов с помощью искусственного интеллекта. Ее решение позволяет компаниям прислушиваться к своим клиентам, адаптироваться к динамике рынка, постоянно совершенствовать свои продукты и услуги, а также улучшать взаимодействие с клиентами за счет повышения эффективности чат-бота. С этой целью инструменты обработки естественного языка обычно используются для формулирования традиционной задачи о маркировке последовательностей. Однако в настоящей статье мы предлагаем другой, альтернативный подход, основанный на использовании возможностей модели машинного обучения MRC (machine reading comprehension, машинное чтение и понимание текста). В данной постановке определение типов информации о продукте аналогично заданию вопроса «Какая информация о продукте упоминается пользователями?». Например, извлечение названий продуктов (которое соответствует метке PRO_NAME) выполняется как извлечение интервалов ответов на вопрос «Какие примеры названий продуктов упоминаются?». Нами проведены обширные эксперименты с общедоступным набором данных, имеющихся во Вьетнаме. Результаты экспериментов показывают надежность предложенного альтернативного метода: он повышает производительность модели распознавания по сравнению с двумя базовыми показателями, обеспечивая их значительное улучшение. В частности, мы достигли уровня 92,87% по шкале F1 при распознавании описаний продуктов на уровне 1. На уровне 2 модель показала результат 93,34% по шкале F1 при распознавании каждого типа информации о продукте.

Ключевые слова: распознавание информации о продукте, фреймворк MRC, розничная торговля, большие языковые модели, viBERT, vELECTRA

Цитирование: Luong T.C., Tran O.T. Product information recognition in the retail domain as an MRC problem // Business Informatics. 2024. Vol. 18. No. 1, P. 79–88. DOI: 10.17323/2587-814X.2024.1.79.88

Литература

1. Tran O.T., Luong T.C. Understanding what the users say in chatbots: A case study for the Vietnamese language // Engineering Applications of Artificial Intelligence. 2020. Vol. 87. 103322. <https://doi.org/10.1016/j.engappai.2019.103322>

* Автор, ответственный за переписку

2. Tran O.T., Bui V.T. A BERT-based hierarchical model for Vietnamese aspect based sentiment analysis // Proceedings of the 12th International Conference on Knowledge and System Engineering (KSE), Can Tho, Vietnam, 2–14 November 2020. P. 269–274. <https://doi.org/10.1109/KSE50997.2020.9287650>
3. Bui V.T., Tran O.T., Le H.P. Improving sequence tagging for Vietnamese text using transformer-based neural models // arXiv:2006.15994. 2020. <https://doi.org/10.48550/arXiv.2006.15994>
4. Lample G., Ballesteros M., Subramanian S., Kawakami K., Dyer C. Neural architectures for named entity recognition // Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, California, June 2016. P. 260–270. <https://doi.org/10.18653/v1/N16-1030>
5. Levy O., Seo M., Choi E., Zettlemoyer L. Zero-shot relation extraction via reading comprehension // arXiv:1706.04115. 2017. <https://doi.org/10.48550/arXiv.1706.04115>
6. McCann B., Keskar N.S., Xiong C., Socher R. The natural language decathlon: Multitask learning as question answering // arXiv:1806.08730. 2018. <https://doi.org/10.48550/arXiv.1806.08730>
7. Li X., Yin F., Sun Z., Li X., Yuan A., Chai D., Zhou M., Li J. Entity-relation extraction as multi-turn question answering // Proceedings of the 57th Conference of the Association for Computational Linguistics, Florence, Italy, July 2019. P. 1340–1350. <https://doi.org/10.18653/v1/P19-1129>
8. Ji Z., Lu Z., Li H. An information retrieval approach to short text conversation // arXiv:1408.6988. 2014. <https://doi.org/10.48550/arXiv.1408.6988>
9. Qiu M., Li F., Wang S., Gao X., Chen Y., Zhao W., Chen H., Huang J., Chu W. AliMe chat: A sequence to sequence and rerank based chatbot engine // Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Vancouver, Canada, July 2017. P. 498–503. <https://doi.org/10.18653/v1/P17-2079>
10. Goncharova E., Ilvovsky D.I., Galitsky B. Concept-based chatbot for interactive query refinement in product search // Proceedings of the 9th International Workshop “What can FCA do for Artificial Intelligence?” (FCA4AI). 2021. Vol. 2972. CEUR-WS. P. 51–58. Available at: <http://ceur-ws.org/Vol-2972/paper5.pdf> (accessed 15 February 2024).
11. Yan Z., Duan N., Chen P., Zhou M., Zhou J., Li Z. Building task-oriented dialogue systems for online shopping // Proceedings of the AAAI Conference on Artificial Intelligence. 2017. Vol. 31. No. 1. <https://doi.org/10.1609/aaai.v31i1.11182>
12. Liu Y., Ott M., Goyal N., Du J., Joshi M., Chen D., Levy O., Lewis M., Zettlemoyer L., Stoyanov V. RoBERTa: A robustly optimized BERT pretraining approach // arXiv:1907.11692. 2019. <https://doi.org/10.48550/arXiv.1907.11692>
13. Yang Z., Dai Z., Yang Y., Carbonell J., Salakhutdinov R.R., Le Q.V. XLNet: Generalized autoregressive pretraining for language understanding // 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada. arXiv:1906.08237. 2019. <https://doi.org/10.48550/arXiv.1906.08237>
14. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser Ł., Polosukhin I. Attention is all you need // 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. arXiv:1706.03762. 2017. <https://doi.org/10.48550/arXiv.1706.03762>
15. Kingma J., Ba J. Adam: A method for stochastic optimization // arXiv:1412.6980. 2015. <https://doi.org/10.48550/arXiv.1412.6980>
16. Devlin J., Chang M.W., Lee K., Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding // arXiv:1810.04805. 2018. <https://doi.org/10.48550/arXiv.1810.04805>
17. Therasa M., Mathivanan G. Survey of machine reading comprehension models and its evaluation metrics // Proceedings of the 6th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2022. P. 1006–1013. <https://doi.org/10.1109/ICCMC53470.2022.9754070>

Об авторах

Тхо Чи Луонг

исследователь, Международный институт франкоязычных стран, Вьетнамский национальный университет в Ханое, Вьетнам, Ханой, Каузяй, Суан Туи, 144, Е5

E-mail: tholc@vnu.edu.vn

ORCID: 0000-0002-7664-705X

Оан Тхи Тран

доцент, доктор (PhD);

преподаватель, Международная школа Вьетнамского национального университета в Ханое, Вьетнам, Ханой, Каузяй, Суан Туи, 144, здание G7;

E-mail: oanhtt@gmail.com

ORCID: 0000-0002-3286-3623



XXV ЯСИНСКАЯ (АПРЕЛЬСКАЯ) МЕЖДУНАРОДНАЯ НАУЧНАЯ КОНФЕРЕНЦИЯ ПО ПРОБЛЕМАМ РАЗВИТИЯ ЭКОНОМИКИ И ОБЩЕСТВА

Уважаемые коллеги!



Национальный исследовательский университет
«Высшая школа экономики» *открывает прием заявок
на участие в XXV Ясинской (Апрельской)
международной научной конференции по проблемам
развития экономики и общества (XXV ЯМНК).*

**Основные мероприятия XXV ЯМНК
состоятся в Москве с 23 по 26 апреля 2024 года.**

**Подать заявку на участие в качестве слушателя
можно до 19 апреля 2024 г.**

Более подробная информация размещена в разделе [«Участникам»](#).

В рамках тематических направлений XXV ЯМНК будут представлены и обсуждены доклады о результатах новых научных исследований, отобранные на основе рассмотрения заявок.

Наряду с этим конференция будет, по сложившейся традиции, включать обсуждения наиболее актуальных проблем экономической, социальной, внутренней и внешней политики в рамках круглых столов и ассоциированных мероприятий.

Мероприятия конференции проводятся на русском или английском языке, в отдельных случаях на двух языках с синхронным переводом.

В интересах привлечения участников из различных регионов России и мира, а также с учетом возможного сохранения некоторых ограничений эпидемиологического характера XXV ЯМНК будет проведена в смешанном формате.

В рамках XXV ЯМНК, как и в предыдущие годы, будет проведен конкурс заявок на поддержку участия в конференции молодых исследователей из российских регионов.



XXV ЯСИНСКАЯ (АПРЕЛЬСКАЯ) МЕЖДУНАРОДНАЯ НАУЧНАЯ КОНФЕРЕНЦИЯ ПО ПРОБЛЕМАМ РАЗВИТИЯ ЭКОНОМИКИ И ОБЩЕСТВА

**Заявки на участие с докладом будут приниматься
по следующим тематическим направлениям:**

1. Востоковедение
2. География и геоинформационные технологии
3. Демография и рынки труда
4. Инструментальные методы в экономических и социальных исследованиях
5. Корпоративные финансы в условиях глобальных вызовов
6. Макроэкономика и макроэкономическая политика
7. Международные отношения
8. Методология экономической науки
9. Мировая экономика
10. Наука и инновации
11. Образование
12. Политические процессы
13. Право в цифровую эпоху
14. Психология
15. Развитие здравоохранения
16. Региональное и городское развитие
17. Менеджмент
18. Социальная политика
19. Социокультурные процессы
20. Социология
21. Теоретическая экономика
22. Умный город
23. Финансовые институты, рынки и платежные системы
24. Фирмы и рынки
25. Экономика данных

Как подать заявку на участие в качестве слушателя?

**Заявка подается онлайн в системе конференции НИУ ВШЭ
до 19 апреля 2024 года.**

Оплата регистрационного взноса:

Для слушателей конференции (участников без доклада) сумма регистрационного взноса составляет **3 000 рублей**.

Подробнее об оплате регистрационного взноса в разделе **«Оплата участия»**.

Все остальные участники освобождаются от уплаты регистрационного взноса, в т. ч.:

- ♦ студенты/аспиранты всех вузов (при предъявлении студенческого билета);
- ♦ сотрудники НИУ ВШЭ (при предоставлении карточки сотрудника);
- ♦ участники, приглашенные Программным и Организационным комитетами конференции.