

№ 15(70) 2009
Выпуск 12/1

НАУЧНЫЙ РЕЦЕНЗИРУЕМЫЙ ЖУРНАЛ

Основан в 1995 г.

Журнал входит
в Перечень ведущих рецензируемых
научных журналов и изданий,
выпускаемых в Российской Федерации,
в которых рекомендуется публикация
основных результатов диссертаций
на соискание ученых степеней
доктора и кандидата наук

Учредитель:

Государственное образовательное учреждение
высшего профессионального образования
«Белгородский государственный университет»

Издатель:

Белгородский государственный
университет.
Издательство БелГУ

Журнал зарегистрирован
в Федеральной службе по надзору за соблюдением
законодательства
в сфере массовых коммуникаций
и охраны культурного наследия
Свидетельство о регистрации средства массовой
информации ПИ № ФС77-21121 от 19 мая 2005 г.

**РЕДАКЦИОННАЯ КОЛЛЕГИЯ
ЖУРНАЛА**

Главный редактор

Дятченко Л.Я.

ректор Белгородского государственного
университета, доктор социологических наук,
профессор

Зам. главного редактора

Давыденко Т.М.

проректор по научной работе Белгородского
государственного университета, доктор
педагогических наук, профессор

Ответственный секретарь

Московкин В.М.

заместитель по инновационной
деятельности проректора по научной
работе Белгородского государственного
университета, доктор географических наук,
профессор кафедры мировой экономики

**РЕДАКЦИОННАЯ КОЛЛЕГИЯ
СЕРИИ ЖУРНАЛА**

Председатель редколлегии

Дятченко Л.Я.

ректор Белгородского государственного
университета, доктор социологических наук,
профессор

Главный редактор

Шаповалов В.А.

доктор исторических наук, профессор
(Белгородский государственный университет)

Заместители главного редактора

Жильков Е.Г.

доктор технических наук, профессор
(Белгородский государственный университет)

Ломовцева О.А.

доктор экономических наук, профессор
(Белгородский государственный университет)

Шатохин И.Т.

кандидат исторических наук, доцент
(Белгородский государственный университет)

Шиллов В.Н.

доктор философских наук, профессор
(Белгородский государственный университет)

Ответственный секретарь

Василенко В.В.

кандидат исторических наук
(Белгородский государственный университет)

НАУЧНЫЕ ВЕДОМОСТИ

Белгородского государственного университета

История Политология Экономика
Информатика

Belgorod State University

Scientific Bulletin

History Political science Economics
Information technologies

СОДЕРЖАНИЕ

РЕГИОНАЛЬНАЯ И МУНИЦИПАЛЬНАЯ ЭКОНОМИКА

Концептуальные основы реализации кластерной политики
на территории российских регионов. **Н.А. Корчагина 5**
Ключевые положения антикризисного регулирования муницип
ального развития. **О.П. Овчинникова, Н.Н. Куриленко 14**
Проблемы методики оценки региональных инвестиционных
проектов. **Е.Н. Парфенова 22**
Внешнеэкономическая деятельность как средство гармонизации
развития информационного пространства. **Е.В. Попова 28**

ИНВЕСТИЦИИ И ИННОВАЦИИ

Влияние реструктуризации на доходность долгосрочной
кредитной операции
М.Ф. Тубольцев, В.И. Болтенков 34
Использование концессионных соглашений в управлении
инфраструктурными объектами экономики Российской
Федерации. **Д.И. Искренко 39**
Моделирование инвестиционного процесса в условиях
неопределенности. **В.А. Калугин, И.Н. Титова 46**

ОТРАСЛЕВЫЕ РЫНКИ И РЫНОЧНАЯ ИНФРАСТРУКТУРА

Модифицированные методы затратного подхода к оценке
товарных знаков российских промышленных предприятий.
Е.И. Каширина 55
Функциональные особенности и перспективы развития
оптовой торговли системы потребительской кооперации.
Р.П. Мешечкина, В.В. Лазарев 64

РЫНОК ТРУДА И ЭКОНОМИКА ОБРАЗОВАНИЯ

К вопросу о необходимости становления системы трансфера
знаний и технологий на основе инновационного потенциала
ведущих вузов России. **М.В. Владыка 70**

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

Исследование одного параметра булевых функций, близкого
к нелинейности. **В.Б. Алексеев, Р.Р. Омаров 81**
Новые архитектуры и алгоритмы обучения дискретных нейрон
ных сетей адаптивной резонансной теории. **В.Д. Дмитриен
ко, И.П. Хавина, А.Ю. Заковоротный 88**
О стойкой обфускации компьютерных программ. **Н.П. Варнов
ский, В.А. Захаров, Н.Н. Кузюрин, А.В. Шокуров 97**
Моделирование информационных процессов в вычислительной
подсистеме с применением агрегативных моделей.
А.Н. Привалов, В.Л. Кулешов 106

Члены редколлегии

Абрамзон М.Г., доктор исторических наук, профессор (Магнитогорский государственный университет)

Болгов Н.Н., доктор исторических наук, профессор (Белгородский государственный университет)

Глухова А.В., доктор политических наук, профессор (Воронежский государственный университет)

Дмитренко В.Д., доктор технических наук, профессор (Харьковский национальный технический университет «ХПИ»)

Илюхина Р.В., доктор экономических наук, профессор (Академия экономической безопасности МВД России)

Ишиakov О.В., заслуженный деятель науки РФ, доктор экономических наук, профессор (Волгоградский государственный университет)

Калузин В.А., доктор экономических наук, профессор (Белгородский государственный университет)

Капалин В.И., доктор технических наук, профессор (Московский государственный институт электроники и математики (технический университет))

Корсунов Н.И., заслуженный деятель науки РФ, доктор технических наук, профессор (Белгородский государственный университет)

Литовка О.П., доктор географических наук, профессор (Институт проблем региональной экономики РАН, г. Санкт-Петербург)

Лобанов К.Н., доктор политических наук, доцент (Белгородский юридический институт МВД России)

Маторин С.И., доктор технических наук, профессор (Белгородский государственный университет)

Молев Е.А., доктор исторических наук, профессор (Нижегородский государственный университет им. Н.И. Лобачевского)

Овчинникова О.П., доктор экономических наук, профессор (Орловская региональная академия государственной службы)

Посохов С.И., доктор исторических наук, профессор (Харьковский национальный университет им. В.Н. Каразина, Украина)

Пуцкарева И.М., доктор исторических наук, старший научный сотрудник (Институт российской истории Российской академии наук)

Рисин И.Е., доктор экономических наук, профессор (Воронежский государственный университет)

Рубанов В.Г., заслуженный деятель науки РФ, доктор технических наук, профессор (Белгородский государственный технологический университет им. В.Г. Шухова)

Шагин Э.М., доктор исторических наук, профессор (Московский государственный педагогический университет)

Оригинал-макет *В.В. Василенко, А.А. Махова*

E-mail: *vasilenko_v@bsu.edu.ru*

Подписано в печать 09.12.2009

Формат 60×84/8

Гарнитура Georgia, Impact

Усл. п. л. 24,87

Тираж 1000 экз.

Заказ 246

Подписные индексы в каталоге агентства

«Роспечать» – 18078,

в объединенном каталоге

«Пресса России» – 39723

Оригинал-макет тиражирован

в издательстве Белгородского государственного университета

Адрес: 308015, г. Белгород, ул. Победы, 85

Развитие методов быстрого вейвлет-преобразования с помощью фильтров Добеши. **Н.И. Червяков, Ю.В. Кондрашов 112**

Новый подход к автоматизированному выявлению шаблонов в телеметрических сигналах на основе декомпозиции на эмпирические моды. **Д.М. Клионский, Н.И. Орешко, В.В. Геппенер 118**

СИСТЕМНЫЙ АНАЛИЗ И УПРАВЛЕНИЕ

Некорректные задачи и непараметрическая идентификация систем управления. **В.И.Капалин 130**

Использование моделей Парето при формировании инновационной деятельности телекоммуникационной компании. **А.П. Игнатъев, З.В.Смирнова 138**

Основы и перспективы применения информационных интеллектуальных систем для процессов жизненного цикла наукоемких изделий. **А.Е. Бром, В.А. Шутеев 144**

Групповое многокритериальное принятие решений с несовпадающими предпочтениями. **А.Б. Петровский 151**

Синтез нечетких решающих правил для медико-экологических приложений на основе анализа структуры данных. **Н.А. Кореньевский, С.А. Филлист, Г.В. Чурсин 161**

Система поддержки принятия решения в медицине на основе синтеза структурированных моделей объектов диагностики. **А.И. Поворознюк 170**

Математическое моделирование финансовых процессов в условиях неопределенности. **М.Ф.Тубольцев, В.М.Михелев 177**

ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

Новые технологии передачи и обработки информации на основе хорошо локализованных сигнальных базисов. **В.П. Волчков 181**

Условия ортогональности обобщенных базисов Вейля-Гейзенберга для OFTDM сигналов. **В.П. Волчков, Д.А. Петров 190**

Об эффективности метода оптимальной фильтрации изображений. **Е.Г. Жиляков, А.А. Черноморец, В.А. Голощанова 200**

Сведения об авторах 206

Информация для авторов 210

№ 15(70) 2009
Issue 12/1

SCIENTIFIC PEER-REVIEWED JOURNAL

Founded in 1995

The Journal is included into the list of the leading peer-reviewed journals and publications coming out in the Russian Federation that are recommended for publishing key results of the theses for Doktor and Kandidat degree-seekers.

Founder:

State educational establishment of higher professional education
"Belgorod State University"

Publisher:

Belgorod State University
BSU Publishing house

The journal is registered in Federal service of control over law compliance in the sphere of mass media and protection of cultural heritage

Certificate of registration of mass media
ПИ № ФС 77-21121 May 19, 2005.

Editorial board of journal

Editor-in-chief

L.J. Djatchenko

Rector of Belgorod State University, doctor of sociological sciences, professor

Deputy editor-in-chief

T.M. Davydenko

Vice-rector for scientific research of Belgorod state university, doctor of pedagogical sciences, professor

Assistant Editor

V.M. Moskovkin

Doctor of geographical sciences, professor of world economy department

Editorial board of journal series

Chairman of editorial series

L.J. Djatchenko

Rector of Belgorod State University, doctor of sociological sciences, professor

Editor-in-chief

V.A. Shapovalov

Doctor of historical sciences, Professor (Belgorod State University)

Deputies of editor-in-chief

E.G. Zhilyakov

Doctor of technical sciences, Professor (Belgorod State University)

O.A. Lomovtseva

Doctor of economical sciences, Professor (Belgorod State University)

I. T. Shatohin

Candidate of historical sciences, Associate professor (Belgorod State University)

V.N. Shilov

Doctor of philosophical sciences, Professor (Belgorod State University)

Editorial assistant

V.V. Vasilenko

Candidate of historical sciences (Belgorod State University)

Belgorod State University
Scientific Bulletin
History Political science Economics
Information technologies

НАУЧНЫЕ ВЕДОМОСТИ БелГУ

История Политология Экономика
Информатика

CONTENTS

REGIONAL AND MUNICIPAL ECONOMY

Conceptual framework of cluster policy in the Russian regions.

N.A. Korchagina 5

Basic propositions of crisis regulation of municipal development.

O.P. Ovchinnikova, N.N. Kurilenko 14

Problems of the estimation mechanism of the regional investment projects. **E.N. Parfenova 22**

Foreign trade activities as a means of harmonization of development of the information sphere. **E.V. Popova 28**

INVESTMENT AND INNOVATIONS

Re- influence of re-structuring on profitability of long-term creditoperation.

M.F. Tuboltsev, V.I. Boltentsov 34

Operation of infrastructure under concession agreements

in the economy of the Russian Federation. **D.I. Iskrenko 39**

Modelling of the investment process in the conditions of uncertainty. **V.A. Kalugin, I.N. Titova 46**

SECTORAL MARKETS AND MARKET INFRASTRUCTURE

Modified methods of cost approach in the Russian industrial enterprises' trademarks valuation. **E.I. Kashirina 55**

Functional peculiarities and prospects of wholesale trade development of consumer cooperation system.

R.P. Meshechkina, V.V. Lazarev 64

LABOUR MARKET AND ECONOMICS OF EDUCATION

On the necessity of formation and development of system of knowledge and technologies transfer on the basis of innovative potential of higher education in Russia. **M.V. Vladyka 70**

COMPUTER SIMULATION HISTORY

Investigation of a parameter of boolean functions closed to nonlinearity.

V.B. Alekseev, R.R. Omarov 81

New architectures and training algorithms of discrete neural networks of adaptive resonance theory. **V.D. Dmitrienko, I.P. Khavina, A.Y. Zakovorotnyi 88**

On the secure obfuscation of computer programs. **N.P. Varnovsky, V.A. Zakharov, N.N. Kuzurin, A.V. Shokurov 97**

Performance evaluation of virtual training system with distributed data processing. **A.N. Privalov, V.L. Kuleshov 106**

Members of editorial board

M.G. Abramzon, Doctor of historical sciences, Professor (Magnitogorsk State University)

N.N. Bolgov, Doctor of historical sciences, Professor (Belgorod State University)

A.V. Glukhova, Doctor of political sciences, Professor (Voronezh State University)

V.D. Dmitrenko, Doctor of technical sciences, Professor (Kharkov National Technical University)

R.V. Ilyukhina, Doctor of economical sciences, Professor (Academy of Economic Security of Ministry of Internal Affairs of Russia)

O.V. Inshakov, Honoured Science Worker of Russian Federation, Doctor of economical sciences, Professor (Volgograd State University)

V.A. Kalugin, Doctor of economical sciences, Professor (Belgorod State University)

V.I. Kapalin, Doctor of technical sciences, Professor (Moscow State Institute of Electronics and Mathematics (technical university))

N.I. Korsunov, Honoured Science Worker of Russian Federation, Doctor of technical sciences, Professor (Belgorod State Technological University named after V.G. Shuhov)

O.P. Litovka, Doctor of geographical sciences, Professor (Institute of regional economy problems of Russian Academy of Sciences, Saint-Petersburg)

K.N. Lobanov, Doctor of political sciences, Associate professor (Belgorod Juridical Institute of Ministry of Home Affairs of Russian Federation)

S.I. Matorin, Doctor of technical sciences, Professor (Belgorod State University)

E.A. Molev, Doctor of historical sciences, Professor (Nizhniy Novgorod State University named after N.I. Lobachevskiy)

O.P. Ovchinnikova, Doctor of economical sciences, Professor (Orel Regional Academy of State Service)

S.I. Posokhov, Doctor of historical sciences, Professor (Kharkov National University named after V.N. Karazin, Ukraine)

I.M. Pushkareva, Doctor of historical sciences, Senior scientific worker (Institute of Russian History of Russian Academy of Sciences)

I.E. Risin, Doctor of economical sciences, Professor (Voronezh State University)

V.G. Rubanov, Honoured Science Worker of Russian Federation, Doctor of technical sciences, Professor (Belgorod State Technological University named after V.G. Shuhov)

E.M. Shagin, Doctor of historical sciences, Professor (Moscow State Pedagogical University)

Dummy layout by V.V. Vasilenko,
A.A. Mahova
e-mail: vasilenko_v@bsu.edu.ru

Passed for printing 09.12.2009
Format 60×84/8
Typeface Georgia, Impact
Printer's sheets 24,87
Circulation 1000 copies
Order 246

Subscription reference
in Rospechat' agency catalogue – 18078,
In joint catalogue Pressa Rossii – 39723

Dummy layout is replicated at Belgorod
State University Publishing House
Address: 85, Pobedy str., Belgorod, Russia, 308015

Development of methods fast wavelet – transformation by means of filters Dobeshi. **N.I. Chervykov, Y.V. Kondrashov 112**
New approach to automatic pattern extraction in telemetric signals on the basis of the empirical mode decomposition. **D.M. Klionskiy, N.I. Oreshko, V.V. Geppener 118**

SYSTEM ANALYSIS AND MANAGEMENT

Ill-posed problems and nonparametric identification of control systems.

V.I. Kapalin 130

Use of models of Pareto at formation of innovative activity of the telecommunication company. **A.P. Ignatev,**

Z.V. Smirnova 138

Basics and perspectives of using information intelligent systems to support life cycle processes of high technology products.

A.E. Brom, V.A. Shuteev 144

Group multicriterial decision making with inconsistent preferences.

A.B. Petrovsky 151

Synthesis of indistinct solving rules for mediko-ecological prilo-zheny on the basis of the analysis of structure of the data.

N.A. Korenevsky, S.A. Filist, G.V. Chursin 161

Decision support system in medicine on base of the syntheses

of the outline models object diagnosticses. **A.I. Povoroznyuk 170**

Mathematical modelling of financial processes in the conditions of uncertainty. **M.F. Tuboltsev, V.M. Mikhelev 177**

INFORMATION TECHNOLOGIES AND TELECOMMUNICATION

A new technology of transmitting and processing of information based on well-localized signal basis. **V.P. Volchkov 181**

Generalized Weyl-Heysenbeg bases orthogonality conditions

for OFTDM signals. **V.P. Volchkov, D.A. Petrov 190**

On efficiency of image optimal filtration method. **E.G. Zhilyakov,**

A.A. Chernomorets, V.A. Golaschapova 200

Information about Authors 206

Information for Authors 210

РЕГИОНАЛЬНАЯ И МУНИЦИПАЛЬНАЯ ЭКОНОМИКА

УДК 338.24

КОНЦЕПТУАЛЬНЫЕ ОСНОВЫ РЕАЛИЗАЦИИ КЛАСТЕРНОЙ ПОЛИТИКИ НА ТЕРРИТОРИИ РОССИЙСКИХ РЕГИОНОВ

Н. А. КОРЧАГИНА

*Астраханский
государственный
университет*

e-mail: lumlight@gmail.com

В рамках статьи обоснована необходимость использования кластерной политики как технологии повышения конкурентоспособности экономики Астраханского региона и рассмотрено основное содержание концептуальных направлений ее реализации. Кроме того, представлены возможные показатели оценки эффективности практической реализации предлагаемой концепции, а также результаты первоначальных шагов по ее внедрению.

Ключевые слова: кластер, кластерная политика, кластерные инициативы, кластерная концепция, региональная конкурентоспособность, региональное развитие, благосостояние российских регионов.

На протяжении последнего десятилетия в мире в качестве одной из наиболее эффективных технологий управления социально-экономическим развитием территорий рассматривается кластерная политика. При наличии все возрастающего числа фактов¹, подтверждающих, что кластеры² оказывают положительное воздействие на уровень регионального развития, растет интерес к разработке такого рода программ, которые позволяли бы развивать кластеры или увеличивать экономические выгоды от их существования.

В России также активно заговорили об использовании кластерной политики для повышения конкурентоспособности отраслей и регионов. На протяжении уже нескольких лет мы слышим это странное слово «кластер». При этом даже те, кто его использует, очень часто совершенно не понимают, о чем они говорят и что реально за ним стоит. В лаборатории конкурентоспособности и кластерных инициатив Астраханского государственного университета не так давно закончилось

¹ В настоящее время существует множество исследований, свидетельствующих о положительной корреляции между наличием кластеров и более высоким уровнем экономического развития территорий. Среди наиболее крупных проектов можно выделить Canadian Cluster Data, European Cluster Observatory, International Cluster Competitiveness Project, U. S. Cluster Mapping Project.

² Кластер - группа географически локализованных взаимосвязанных компаний – поставщиков оборудования, комплектующих, специализированных услуг, инфраструктуры, научно-исследовательских институтов, вузов и других организаций, дополняющих друг друга и, что самое главное, усиливающих свои конкурентные преимущества.

исследование стратегий развития всех российских регионов³. И практически во всех из них используются слова кластер, кластерная политика. Однако, как правило, преобладает формальный подход к решению многих проблем.

В большинстве российских регионов разработаны или в настоящий момент разрабатываются стратегии социально-экономического развития. Анализ сроков реализации стратегий выявил, что практически все регионы имеют долгосрочную стратегию развития (свыше 5 лет), что позволяет говорить о наличии предпосылок для разработки эффективной кластерной политики, поскольку ее проведение возможно только в долгосрочной перспективе.

Кроме того, в этих стратегиях в качестве основной задачи кластерной политики рассматривается создание кластеров. Однако, по мнению большинства исследователей, создание кластеров, наряду с таргетированием, свойственным промышленной политике, – очень опасный путь для экономики. Он приводит к концентрации усилий на тех категориях кластеров, которые рассматриваются в качестве стратегически важных для экономического развития. В таком случае многие районы, не обладающие соответствующими конкурентными преимуществами, будут бороться за одни и те же немногочисленные предприятия отрасли, пользующиеся поддержкой. И конкурировать они будут на основе материальных поощрений фирмам, а не на основе первоначальных условий, выгодных для осуществления их деятельности. Такой принцип определения размещения кластера приводит к централизации экономического развития, что очень затратно для государственного бюджета. Более того, он основывается на ограничении или искажении конкуренции, что может подорвать основы конкурентоспособности, поскольку высокий уровень местной конкуренции играет решающую роль в создании конкурентоспособных кластеров.

Анализ стратегий социально-экономического развития российских регионов также продемонстрировал, что в качестве основных фасилитаторов кластерной политики в большинстве регионов выступают региональные органы власти. При этом наблюдается низкий уровень взаимодействия структур, участие которых необходимо для эффективной реализации кластерной политики. Речь в данном случае идет о представителях бизнеса, учебных и научных заведений, общественных организаций и сетевых структур. Наличие такого взаимодействия выявлено лишь в некоторых регионах, что также свидетельствует о недостаточном уровне понимания условий эффективности для развития кластерной политики.

Полученные результаты количественного и качественного анализа показали, что уровень развития кластерной политики в российских регионах является низким, а ее содержание в значительной степени формализовано и не позволяет говорить о положительных эффектах региональных экономик.

Помимо изложенных выше фактов, свидетельствующих о слабой эффективности кластерной политики в российских регионах, анализ продемонстрировал следующее:

- выявлено наличие серьезных проблем с доступностью к текстам стратегий регионов в СМИ;
- большая часть стратегий написана «под копирку»;
- разработка стратегий и определение «приоритетных» кластеров не базируется на общепринятых в мировой практике методиках кластерного картирования, практически все кластеры определены «на глазок»;
- разработчиками стратегий являются привлеченные консалтинговые структуры, не имеющие реальной заинтересованности в глубоком изучении ситуации и реалистичности предлагаемых мер;

³ См. Корчагина Н.А., Ракиева О.В. Оценка уровня развития и содержания кластерной политики в российских регионах // Труды Вольного экономического общества России. Москва, 2009. Том 115. №4. С. 176 – 208.

- к процессу разработки стратегий и программ кластерного развития практически не привлекаются вузы, функционирующие на территории регионов;
- наблюдается практически полное отсутствие образовательных программ, нацеленных на повышение уровня знаний о кластерах и кластерной политике.

Несомненно, что успешная реализация кластерной политики⁴, подразумевающая решение проблем социально-экономического развития территории, повышение уровня благосостояния населения региона, обеспечение более высокой прибыльности компаний, появление новых малых и средних предприятий, повышение инновационной активности научных и учебных заведений и т.д., невозможна без глубокого осознания различными экономическими субъектами региона необходимости ее внедрения и полного осмысления ее содержания.

В этой связи представляется необходимым четкое видение концепции кластерной политики Астраханской области, а также осмысление первых шагов, которые уже были проделаны в направлении ее практической реализации.

Абсолютно очевидно, что разработка общей концепции и программы кластерной политики любого региона, а также программ развития каждого отдельного кластера, должна строиться на основе проведения детального анализа кластерного потенциала, определения уровня и этапа развития существующих кластеров. В мировой практике используются различные способы идентификации кластеров с последующим нанесением их на карту определенной территории. В целом выделяют два широких подхода, позволяющих выявлять кластеры: количественный, который опирается на экономическое моделирование и базируется на статистических методах, и качественный, в основе которого лежит сбор качественной информации о существующих кластерах. Наиболее распространенными количественными методами являются: метод, основанный на расчете коэффициента локализации, Рипли К-метод, метод изучения экспорта и таблиц «затраты-выпуск». Частным случаем качественного подхода может стать рассмотрение конкретных примеров (кейсов), опросы экспертов или метод интервьюирования, а также метод составления генеалогического древа отраслевых кластеров⁵. В настоящее время в Лаборатории конкурентоспособности и кластерных инициатив Астраханского государственного университета осуществляется исследование, нацеленное на выявление наиболее оптимального подхода к составлению кластерных портфелей российских регионов.

Для начала практической реализации кластерных инициатив на территории Астраханской области в настоящий период представляется возможным использование результатов, полученных в ходе совместного проекта компании «Бауман Инновейшн» и Института стратегии и конкурентоспособности Гарвардской

⁴ В данной работе под кластерной политикой понимаются совместные, целенаправленные действия представителей бизнеса, органов власти, образовательных и научно-исследовательских институтов, а также других элементов кластера по формированию благоприятных условий бизнес-среды для развития кластера и повышения эффективности деятельности всех его элементов.

⁵ Данная статья не имеет целью подробный анализ существующих методов кластерного анализа. Более подробная информация по данной проблематике может быть найдена в следующих источниках: M. Porter. The economic performance of regions, April. 2003. System requirements: Adobe Acrobat Reader. URL: www.regional-studies-assoc.ac.uk; Torsten Andersen, Markus Bjerre and Emily Wise Hansson. The Cluster Benchmarking Project: Pilot Project Report - Benchmarking clusters in the knowledge based economy, November. 2006. P. 14-23; Joseph Cortright. Making sense of clusters: regional competitiveness and economic development. The Brookings Institution Metropolitan Policy Program. Impresa, Inc. March. 2006. P. 28-37; Danny Quah, Helen Simpson. Spatial Cluster Empirics. LSE Economics Department and Institute for Fiscal Studies. June. 2003. P. 40; The concept of clusters and cluster policies and their role for competitiveness and innovation: main statistical results and lessons learned. PRO INNO Europe paper. 2008. No. 9. P.14-20; URL: <http://www.clusterobservatory.eu>

школы бизнеса по определению кластерного портфеля российских регионов (2007 г.), и обобщенных в базе данных «Кластеры в России»⁶.

Принимая во внимание положительный опыт проведения кластерной политики в различных регионах мира, в качестве следующего приоритетного шага предлагается широкомасштабное обучение кластерной методологии представителей органов власти, бизнеса, образовательных и других организаций, являющихся элементами выявленных кластеров. Формирование понимания менеджерами разных уровней необходимости внедрения новой технологии управления конкурентоспособностью региона и отдельных компаний в рамках кластерных инициатив является одним из основных факторов успеха проектов такого рода.

В процессе, а также по результатам, обучения должны быть сформированы инициативные кластерные группы, включающие, прежде всего, представителей бизнеса, имеющих «имя», мнение которых будет услышано и поддержано другими участниками кластера. Кроме того, каждая группа должна включать представителя органа власти, способного, с одной стороны, на начальном этапе разработки любых предложений со стороны кластерной группы четко обозначить правомерность и возможность дальнейшей реализации таких предложений, а с другой – имеющего достаточный вес для «продвижения» этих предложений в органах власти. Особое внимание при формировании таких кластерных групп должно уделяться персоналиям. В данном случае вопрос совместимости людей в этой группе, а также их креативность и вера в идею кластерного развития имеют самое принципиальное значение. Формирование кластерных инициативных групп приведет к объединению интеллектуального и организационного потенциала лидеров, а также к использованию индивидуальных ресурсов и возможностей, имеющихся у каждого отдельного представителя кластерной группы для решения проблем, стоящих перед кластерами. Образование кластерных групп на начальном этапе может основываться на подписании соглашений о намерениях с последующим их преобразованием в институциональные структуры различной формы собственности.

В ходе обучения и дальнейшей работы инициативных кластерных групп необходимо осуществить детальную формализацию выбранных кластеров, т.е. определение всех существующих и потенциально необходимых элементов кластера, а также оценку уровня взаимодействия существующих элементов. Детальная прорисовка кластерных схем позволяет четко увидеть, какого рода предприятий и организаций не хватает для эффективной деятельности. Соответственно, в последствии эта информация может предоставляться широкой общественности. Это позволит людям, желающим начать новый бизнес или диверсифицировать имеющийся, четко определять рыночные ниши, оценивать существующий спрос, а также видеть потенциальных потребителей или поставщиков. Кроме того, государственные органы, оказывающие финансовую и организационную поддержку, могут делать акценты именно на содействии возникновению и развитию таких предприятий.

Совместная работа представителей различных элементов кластера должна быть направлена, прежде всего, на выявление проблем, препятствующих развитию кластера, формирование перечня таких проблем, решение которых возможно в рамках работы кластерных групп (по приоритетам) и, наконец, определение конкретных путей их решения с формированием программы действий каждой кластерной группы с указанием участников реализации проектов, необходимого ресурсного обеспечения и сроков реализации. На данном этапе крайне важным представляется активный сбор информации по передовому опыту решения схожих проблем в других регионах и странах, а также организация совместных поездок для ознакомления с этим опытом и разработки бизнес-планов в рамках кластерных стратегий.

⁶ База данных «Кластеры в России», «Бауман Инновейшн» и Институт стратегии и конкурентоспособности Гарвардской школы бизнеса, кластерная классификация: версия 4.0 (2005/2007).

Кроме того, представляется крайне важным развитие сотрудничества с участниками других кластеров, а также координация инициатив с представителями связанных кластеров. При этом принципиальное значение имеет сопоставление результатов работы отдельных кластерных групп для выделения наиболее приоритетных проблем, являющихся общими для всех кластеров. Такого рода работа позволит сформировать перечень приоритетных проблем, сдерживающих развитие всех кластеров и региона в целом и определить совместную программу действий для всех кластерных групп. В рамках такой программы представляется целесообразным определить несколько групп проблем в зависимости от временных рамок их возможного решения. В целях формирования благоприятной основы для дальнейшей долгосрочной реализации кластерных инициатив и обеспечения уверенности в возможности совместных действий крайне необходимы плотная работа кластерных групп и правительства для решения сиюминутных проблем и немедленная реализация проектов, имеющих высокие шансы на успех.

Очевидно, что кластерно-ориентированный подход к управлению социально-экономическим развитием территории обеспечивает более рациональное финансирование решения общерегиональных проблем. Объясняется это, прежде всего, активным участием бизнеса и представителей других элементов кластера в выявлении существующих проблем, реальном понимании необходимости их решения, четком расчете бюджета под реализацию совместных проектов и, соответственно, в определении долей софинансирования со стороны бюджетов различных уровней, частного бизнеса, общественных организаций, научно-образовательных учреждений и т.д.

В целях обеспечения большей эффективности предпринимаемых усилий по реализации кластерной политики в качестве дополнительных мер возможно:

- выстраивание региональной экономической политики на основе кластерного подхода (например, предоставление финансовой поддержки проектам, оказывающим влияние на улучшение бизнес-среды и кластера в целом, а не отдельным предприятиям);
- организация информационной поддержки, включающей публикации в прессе, в Интернете, на телевидении, цель которой – максимально подробное доведение информации о понятиях кластера и кластерной политики, пояснение ее практической реализации на территории Астраханской области и эффекты для бизнеса и региона;
- организация обучения предпринимателей и лиц, желающих открыть собственное дело, для стимулирования развития малого и среднего бизнеса на территории Астраханской области;
- обеспечение послепроектной устойчивости посредством стимулирования кластерных групп к обеспечению непрерывности развития кластера по окончании реализации проекта кластерной инициативы, а также комплексной юридической оценки и формализации институциональной структуры кластера.

Необходимо четко представлять, что проведение кластерной политики является инструментом долгосрочного развития, и реальные результаты кластерных инициатив могут не проявляться на протяжении долгого времени. Однако, с момента разработки концепции и стратегии, а также в ходе всего процесса их реализации, необходим учет показателей ее эффективности. В первоначальном варианте показатели эффективности кластерной политики могут быть сгруппированы следующим образом:

- индикаторы процесса, которые используются для оценки результатов первого этапа развития инициативы, включающего определение кластерных лидеров

и формирования кластерной группы, объединяющей всех заинтересованных лиц. К таким индикаторам можно отнести количество проведенных семинаров, встреч и презентаций по проблемам конкурентоспособности, количество сформированных кластерных групп и подписанных соглашений о намерениях, количество разработанных конкретных планов действий и т.д.;

– индикаторы действий, которые используются для оценки кластерных анализов и разработки стратегий, нацеленных на повышение конкурентоспособности посредством развития кластеров. К ним можно отнести показатели, оценивающие объем технической помощи, предоставленной конкретным кластерам, завершение реализации стратегий развития каждого кластера, завершение планов действий с учетом степени участия отдельных представителей кластера;

– индикаторы инвестиций, которые используются для оценки уровня финансирования и софинансирования, осуществленных в процессе реализации кластерной инициативы. Софинансирование обычно имеет небольшие размеры в начале реализации инициативы, но увеличивается в процессе ее развития. Такие индикаторы могут включать объемы инвестиций в различные проекты, объем реинвестированной прибыли, полученной в ходе реализации кластерных проектов, объем венчурного капитала, инвестированного в такие сферы, как университетские исследования, частное финансирование и пр.;

– индикаторы достигнутых результатов, которые используются для измерения количественных результатов кластерных инициатив. К таким показателям можно отнести рост экспорта, доходов, занятости, заработных плат, доходности отрасли, а также количество созданных вновь предприятий⁷.

Данный перечень не является исчерпывающим и может и должен расширяться по мере практического воплощения и изучения опыта.

Что касается практической реализации концепции кластерной политики, необходимо отметить, что в настоящее время сотрудники Лаборатории конкурентоспособности и кластерных инициатив Астраханского государственного университета совместно с Министерством экономического развития Астраханской области уже предприняли определенные шаги для апробации предлагаемых мероприятий.

На основе кластерного портфеля, сформированного, как указано выше, в ходе совместного проекта компании «Бауман Инновейшн» и Института стратегии и конкурентоспособности Гарвардской школы бизнеса, был определен перечень кластеров, в рамках которых было принято решение о начале реализации кластерных инициатив. В дальнейшем были организованы обучающие семинары для руководителей и сотрудников органов государственной власти, представляющих различные министерства и ведомства с последующим их привлечением к схематичной прорисовке кластеров и составлению перечня участников кластеров.

В качестве первого пилотного проекта был определен кластер рыбодобычи и рыбопереработки. В ходе первоначального этапа работы с представителями различных элементов кластера была прорисована его детальная схема (см. рис.) и определены приоритетные задачи (см. табл.).

⁷ Clusters for Competitiveness. A Practical Guide and Policy Implications for Developing Cluster Initiatives, International Trade Department PREM the World Bank, February 2009. System requirements: Adobe Acrobat Reader. URL: http://siteresources.worldbank.org/INTEXP/COMNET/Resources/cluster_initiative_pub_web_ver.pdf



Рис 1. Кластер рыбодобычи и рыбопереработки Астраханской области

Кроме того, в процессе реализации пилотного проекта был выявлен ряд характерных особенностей, присущих представителям бизнеса, среди которых:

- низкий уровень доверия с их стороны по отношению к органам государственной власти и представителям образовательных учреждений;
- скепсис большинства бизнесменов по отношению к самой идее кластерного развития;
- отсутствие у большей части представителей бизнеса желания учиться;
- преобладание «потребительского» отношения: «вы нам расскажите, что и как мы должны сделать, чтобы мы могли работать более эффективно»;
- отсутствие доверия со стороны руководителей компаний друг к другу;
- низкий уровень информации о том, что происходит в регионе, какие предприятия есть, что они могут предложить;
- нежелание бизнесменов объединяться;
- нежелание представителей предприятий занять лидирующую позицию в реализации кластерных инициатив, упование на органы власти.

Однако подобное негативное впечатление является закономерным, поскольку внедрение инноваций – всегда сложный и болезненный процесс. Опыт большинства государств свидетельствует о том, что реализация кластерной политики не является исключением, особенно на первоначальных этапах.

Таблица 1

**Проблемы кластера рыбодобычи и рыбопереработки Астраханской области
и возможные пути их решения**

Элемент кластера	Проблема	Возможные пути решения
1	2	3
Ключевые предприятия	Отсутствие оборотных средств (сложная система подготовки документации, задержка субсидий, низкая залоговая база, отсутствие кредитования проектов в полном объеме). Слабая материально-техническая база	Совершенствование системы кредитования (создание нескольких кредитных продуктов) с целью оптимизации приобретения наиболее конкурентоспособного оборудования
	Отсутствие системы оптовой реализации выращенной продукции. Отсутствие емкостей для передержки рыбы	Строительство живорыбных баз и зимовальных прудов, обеспечивающих длительную передержку живой рыбы. Создание производственного кооператива – возможность предоставления заемных средств на кооператив. Строительство базы за счет государства, затем их аренда
	Высокие цены на энергоносители для предприятий аквакультуры и, как результат, потеря объемов выращенной рыбы	Субсидирование затрат
	Низкая рентабельность производства вследствие экстенсивных технологий производства в прудовых хозяйствах	Организация образовательных программ по новым технологиям производства. Организация зарубежных стажировок для ознакомления с передовым опытом
	Требования к постоянным поставкам определенных объемов рыбной продукции для реализации при сезонности производства рыбы	Координирование сроков поставки между всеми поставщиками рынка Астраханской области
	Нехватка высококвалифицированных кадров	Повышение имиджа специалистов рыбной отрасли. Повышение заработной платы и создание комфортных условий для привлечения молодых специалистов. Повышение роли производственной практики студентов. Внедрение механизма аутстаффинга (лизинга персонала)
Родственные и поддерживающие отрасли	Отсутствие местного кормопроизводства (специализированных рыбных кормов)	Развитие инфраструктуры отечественного кормопроизводства. Строительство завода по производству кормов
	Отсутствие дистрибьюторов (оптового торгового звена)	Формирование оптовой товарной биржи
	Отсутствие специализированных транспортных компаний	Строительство транспорта для перевозки живой рыбы
	Высокий износ производственных фондов	Повышение доступности кредитных ресурсов
	Отсутствие в регионе необходимой пленочной этикетки	Установление соответствующего оборудования в типографиях города
	Отсутствие единой торговой политики, утрата позиций бренда астраханской рыбы	Маркетинг региона, формирование спроса за счет бренда
	Отсутствие стратегии развития туризма на Каспии	Создание туристической ассоциации стран каспийского региона
	Административные барьеры	Проведение образовательных семинаров по нововведениям в законодательство
	Слабый обмен информацией между всеми элементами кластера	Разработка информационного портала и справочника



Завершая исследование вопроса о возможных направлениях реализации кластерной политики на территории регионов РФ, представляется целесообразным сделать следующие выводы:

1. В большинстве российских регионов кластерная политика уже рассматривается в качестве одного из наиболее эффективных инструментов повышения их конкурентоспособности, однако уровень ее развития является низким, а содержание – в значительной степени формализовано и не позволяет рассчитывать на положительные эффекты для региональных экономик. Это говорит о необходимости четкого понимания как самой сути кластерной политики, так шагов по ее реализации.

2. Основные этапы внедрения кластерной политики предполагают:

– проведение детального анализа кластерного потенциала, определение уровня и этапов развития существующих кластеров;

– широкомасштабное обучение кластерной методологии представителей всех элементов выявленных кластеров;

– формирование инициативных кластерных групп;

– детальная прорисовка кластерных схем в целях определения всех существующих и потенциально необходимых элементов кластера, а также оценка уровня их взаимодействия;

– выявление проблем, препятствующих развитию кластера и определение конкретных путей их решения;

– координация инициатив с представителями связанных кластеров;

– выстраивание системы оценки эффективности кластерной политики.

3. Апробация предлагаемой концепции реализации кластерной политики на территории Астраханской области в рамках пилотного проекта позволила выявить определенные проблемы и вопросы, а также обозначить перспективы для ее дальнейшего совершенствования.

CONCEPTUAL FRAMEWORK OF CLUSTER POLICY IN THE RUSSIAN REGIONS

N. A. KORCHAGINA

Astrakhan State University

e-mail: lumlight@gmail.com

This article underlines the necessity of cluster policy implementation within Astrakhan region as it is regarded as an effective tool for building up its long-term competitiveness. The article also describes the conceptual basis for this kind of policy. Besides some possible measures of the concept practical implementation are provided as well as the first impressions of the initial steps are described.

Key words: cluster, cluster policy, cluster initiatives, cluster concept, regional competitiveness, regional development, prosperity of Russian regions.

КЛЮЧЕВЫЕ ПОЛОЖЕНИЯ АНТИКРИЗИСНОГО РЕГУЛИРОВАНИЯ МУНИЦИПАЛЬНОГО РАЗВИТИЯ

О. П. ОВЧИННИКОВА¹⁾
Н. Н. КУРИЛЕНКО²⁾

¹⁾ Орловская региональная академия государственной службы

²⁾ Белгородский государственный университет

*e-mail:
kurilenko@bsu.edu.ru*

Изложены базовые понятия и механизм антикризисного регулирования муниципального развития. Обоснована система прямых и косвенных методов антикризисного регулирования, предложены мероприятия по выводу социальной сферы муниципальных образований из кризисного состояния, ее стабилизации и дальнейшему развитию. Исходные данные, определяющие формирование антикризисного регулирования муниципального развития, представлены моделью, в которой выделены основные параметры, влияющие на процесс стабилизации хозяйственной деятельности местных предприятий.

Ключевые слова: муниципальное развитие, антикризисное регулирование, методы и механизмы стабилизации экономического развития, модель антикризисного регулирования муниципального развития.

Проблемы большинства муниципальных образований центрально-европейской части Российской Федерации идентичны: упадок сельскохозяйственной деятельности, неразвитость инфраструктуры, депопуляция, плохое качество услуг сельских территорий по сравнению с городской зоной, сильная зависимость от сельского хозяйства и отсутствие возможностей для несельскохозяйственных видов деятельности.

Финансовый кризис нанес серьезный удар по сельским территориям. Многие из предприятий, работающих здесь, перешли на сокращенный режим работы. До кризиса программы развития сельской местности включали поддержку фермерских хозяйств и развитие альтернативных видов деятельности (туризм, лесное хозяйство, промышленность), которые позволяют обеспечивать большее количество возможностей занятости и увеличения доходов¹, а в «Программе антикризисных мер Правительства Российской Федерации на 2009 год»² уже ничего не сказано о поддержке сельскохозяйственных территорий. В основном, меры Правительства направлены на поддержку крупных предприятий и финансовой системы. Косвенно муниципальные образования в программе могут рассчитывать на поддержку малого бизнеса и обеспечение общественных работ для местных жителей.

Следует отметить, что значительная часть трудоспособного населения сельских территорий уезжала «на заработки» в крупные города и районы Крайнего Севера. Однако с консервацией значительной доли строительного комплекса и ряда предприятий потребность в кадрах снизилась, следовательно, эти люди возвращаются в свои города, создавая дополнительную нагрузку на местный рынок труда.

В таких условиях возникает необходимость в проведении антикризисных мероприятий на местном уровне.

Организационно-экономическое содержание антикризисного регулирования муниципального развития можно рассматривать с нескольких теоретико-методических позиций:

¹ Суглобов А.Е. Перспективы социального развития сельских регионов и использование результатов экономического анализа для управленческих решений // Региональная экономика. 2006. №11. С. 35.

² Программа антикризисных мер Правительства Российской Федерации на 2009 год [Электронный ресурс] / Интернет-портал Правительства Российской Федерации.



а) как экономическую категорию, которая обладает собственным содержанием и понятийным аппаратом;

б) как самостоятельный управленческий процесс и целостную систему воздействий на социально-экономические процессы и явления.

Целью антикризисного регулирования муниципального развития является разработка и реализация мер, направленных на нейтрализацию наиболее опасных (интенсивно влияющих на завершающее явление) факторов, приводящих к кризисному состоянию. Цель антикризисного управления предполагает стремление к оптимальной степени стабилизации в отдельных сферах муниципального хозяйства посредством эффективной системы воздействий на «кризисные точки» муниципальной экономики (предприятия, рынки и т.п.).

Конкретные цели антикризисного регулирования могут относиться к определенной сфере хозяйствования, отрасли производства товаров и услуг, тому или другому виду продукции, предприятию, конкретным рынкам и т.п.

Тогда, *антикризисное регулирование муниципального развития* – это регламентированная нормами права деятельность органов местного самоуправления, базирующаяся на специфических принципах экономического и административного воздействия на кризисные ситуации в муниципальной социально-экономической системе. С точки зрения практической деятельности, антикризисное регулирование представляет собой совокупность действий и операций по наблюдению за финансовыми сторонами деятельности как органов местного самоуправления, так и хозяйствующих субъектов на территории муниципального образования и воздействия на них с применением специфических административных и финансовых форм и методов.

Таким образом, антикризисное регулирование муниципального развития – это реализация права органов местного самоуправления и государственного управления законодательным путем обеспечивать стабильное функционирование муниципальной экономики, осуществляемая в рамках проводимой государственной, региональной и местной политики через систему законодательных, организационных, административных, правовых и др. мер.

Задачей антикризисного регулирования муниципального развития выступает контролируемость состояния муниципальной экономики в соответствии с заданными целями и нормами функционирования.

К общим принципам антикризисного регулирования могут быть отнесены: нейтральность, экономичность, непрерывность, действенность, своевременность и др. Они концентрируют и фиксируют объективные закономерности функционирования общества и государства, выражают решающие условия целенаправленного развития системы государственного регулирования, реализацию принципов свободы, демократизма, законности и др., определяют объективную необходимость выполнения учетно-аналитической, статистической, контрольной деятельности в работе органов местного самоуправления.

Основные принципы антикризисного регулирования в совокупности с задачами, объектами и субъектами регулирования, используемыми приемами и методами составляют целостную категориальную систему теоретико-методологического обоснования антикризисной деятельности органов местного самоуправления.

В системе антикризисного регулирования взаимодействуют органы местного самоуправления и субъекты хозяйствования. К последней группе относятся экономические агенты: предпринимательские структуры, включая физических лиц; хозяйствующие субъекты с государственным участием, в том числе государственные и муниципальные предприятия; некоммерческие организации; прочие субъекты хозяйственной деятельности.

Объектом антикризисного регулирования муниципального развития выступают процессы и явления, на которые направлено воздействие мер антикризисного управления всех уровней управления. Иными словами, в качестве объекта регулирования выступает муниципальная социально-экономическая система, основные характеристики ее состояния и поведения: количественные, качественные, структурные, пространственные, временные.

Стратегически антикризисное регулирование на местном уровне включает анализ социально-экономических тенденций в динамике, проводимый на основе мониторинга социально-экономических процессов в муниципальном образовании, и выработку комплекса мер по предотвращению и/или недопущению кризисных ситуаций в местном хозяйстве.

Это подразумевает выработку концепции стабильного экономического развития, формирование и поддержание на должном уровне стратегических направлений экономического развития, способности обеспечивать в течение длительного периода конкурентное преимущество местных товаропроизводителей как на внутреннем, так и на внешнем рынках.

Задачи местных органов власти по проведению антикризисных мероприятий видятся в следующем:

- 1) формирование методик диагностики кризисного состояния на основе мониторинга социально-экономического развития муниципальной экономики;
- 2) разработка комплексных целевых программ, определяющих меры воздействия на муниципальную социально-экономическую систему для преодоления кризисных явлений, а также формирующих условия для предотвращения их появления;
- 3) определение единой системы показателей кризисного состояния муниципальной экономики;
- 4) обеспечение единства и достоверности формирования информационной базы антикризисной деятельности;
- 5) разработка методики прогнозирования и планирования поступлений в бюджет с учетом возможных кризисных состояний муниципальной экономики или отдельных ее элементов;
- 6) формирование эффективных механизмов взаимодействия органов местного самоуправления и предпринимательских структур по преодолению кризисов и т.п.

Методами антикризисного регулирования со стороны местных органов власти можно считать:

- мониторинг и диагностику кризисных процессов в муниципальной социально-экономической системе;
- формирование прогнозов и планов с учетом вероятности возникновения кризисных явлений;
- применение налоговых, бюджетных, кредитных и прочих мер воздействия на социально-экономическую систему с целью нейтрализации кризисных воздействий (налоговые кредиты, субсидирование процентной ставки и т.п.).

Следует отметить, что, с одной стороны, стратегические решения по предотвращению кризиса должны быть приняты и реализованы на ранних стадиях антикризисного управления, когда процесс движения к кризису еще не приобрел необратимого характера. С другой стороны, принимаемые на ранних стадиях кризисных процессов решения могут быть неадекватны возможной угрозе формирования кризисной ситуации и направлены на решение частных хозяйственных проблем, скрывающих глубинные кризисные явления.

Возможно выделение двух видов реакции на кризисное или предкризисное состояние:



а) «защитная» реакция, предусматривающая сокращение расходов, продажу части имущества и т.п.;

б) «наступательная», предполагающая активные действия по поиску решений выхода из кризиса, реализацию прогрессивной стратегической концепции, отвечающей изменившимся условиям хозяйствования.

Однако, как правило, антикризисные меры носят чрезвычайный характер, реализуются в кратчайший период времени, когда кризис уже наступил. В этой связи наиболее перспективными видятся превентивные меры антикризисного воздействия на развитие муниципальной экономики.

Тактика и стратегия антикризисного регулирования ориентированы на достижение определенных результатов на каждом этапе кризисной ситуации. Исходя из установленных целей антикризисного регулирования, основными результатами мероприятий по его осуществлению являются:

1) прогнозирование развития кризисных явлений, оценка неблагоприятных тенденций в муниципальной экономике;

2) определение эффективных методов и механизмов проведения государственной и муниципальной антикризисной политики;

3) моделирование поведения хозяйствующих субъектов на рынке и, соответственно, разработка сценариев развития социально-экономической ситуации;

4) выбор направлений стабилизации муниципальной социально-экономической системы;

5) определение механизмов послекризисного развития муниципальной экономики.

Функции антикризисного регулирования:

– разработка механизма проведения мониторинга социально-экономического состояния развития муниципальной экономики с целью своевременного обнаружения кризисных явлений;

– диагностика кризисных процессов, предполагающая не только фиксирование отклонений, но и анализ причин их появления, выявление возможных тенденций развития;

– прогнозирование развития кризисных ситуаций, обобщение и оценка информации по факторам, вызвавшим кризисные явления, определение логики и закономерностей самого кризиса, что позволяет рассматривать различные прогнозы и сценарии развития кризисных ситуаций;

– разработка и реализация антикризисных мер, формирование программ, направленных на противодействие кризисным явлениям.

Для вывода социальной сферы из кризисного состояния, ее стабилизации и дальнейшего развития можно предложить следующие меры:

1) стабилизировать уровень жизни сельского населения путем повышения мотивации сельскохозяйственного труда на основе протекционистской, целевой кредитной, бюджетно-налоговой и инвестиционной политики на селе;

2) обеспечить в сельскохозяйственных отраслях переход на низкочастотные, ресурсосберегающие технологии, системы и приемы ведения хозяйства, позволяющие производить рентабельную, конкурентоспособную продукцию, от реализации которой можно будет получать доходы, необходимые для решения социальных проблем;

3) диверсифицировать производственную деятельность в направлении развития экономически эффективных несельскохозяйственных отраслей и производств (прежде всего, по переработке сельскохозяйственной продукции, строительной индустрии, народных промыслов), позволяющих повысить уровень занятости и рост доходов на селе, обеспечить действенный контроль за переселением мигрантов и вынужденных переселенцев в регионы, нуждающиеся в дополнительной рабочей силе;

4) обеспечить развитие и эффективную деятельность сельских кредитных кооперативов;

5) создать целевые фонды для аккумуляции средств, направляемых на развитие социальной сферы.

Исходные данные, определяющие антикризисное регулирование муниципального развития, можно представить моделью, в которой выделены основные параметры, влияющие на процесс стабилизации хозяйственной деятельности:

$$Y(t_k) \leftarrow [\{a_j\}, \{q_i(g_v)\}, \{k_s\}, \{x_h\}] \quad (1)$$

где:

$Y(t_k)$ – ситуация принятия управленческого решения по реализации антикризисных мер как инвариантное (обеспечивающее устойчивость и управляемость) сочетание факторов в определенный временной интервал;

a_j – одна из альтернатив преодоления кризисного процесса ($j=1, 2, \dots, N$);

t_k – условная временная координата, соответствующая периодичности измерения фактических показателей кризисного процесса ($k=1, 2, \dots, W$);

q_i – фактические показатели действенности антикризисных мероприятий (оценка финансовых показателей, уровень занятости и т.п.) ($i=1, 2, \dots, M$), определяемые g_v – совокупностью внешних условий окружающей среды ($v=1, 2, \dots, U$);

k_s – управленческие воздействия по изменению показателей антикризисного процесса ($s=1, 2, \dots, R$);

x_h – последующее состояние антикризисного процесса ($h=1, 2, \dots, H$).

Антикризисное регулирование органами местного самоуправления $Y(t_k)$ рассматривается как ситуативно актуальная совокупность управленческих воздействий $\{k_s\}$ в момент t_k , которые могут перевести некоторый кризисный процесс, реализующий определенный вариант развития кризисной ситуации $\{a_j\}$, от текущего состояния $\{x_h\}$ к новому завершеному этапу $\{x_{h'}\}$. Ход антикризисных процессов сопровождается изменением внешних условий $\{g_v\}$, которые вместе с действием внутренних факторов выражаются в фактических показателях осуществления антикризисной деятельности $\{q_j\}$.

Вероятность перехода к новому состоянию зависит от конкретных приемов управленческих воздействий: $p(x_h \rightarrow x_{h'})_{k_s}$. Если обозначить через e_s меру эффективности k_s -го варианта управленческого воздействия при a_j -й стратегии, то задачу нахождения оптимальной совокупности управленческих воздействий (то есть формирования механизма антикризисного регулирования) можно сформулировать следующим образом: найти такую оптимальную совокупность управленческих воздействий k_s^* , чтобы мера эффективности была максимальной. В этом случае мера эффективности определяется вероятностью достижения следующего состояния кризисного процесса:

$$e_s(x_h \rightarrow x_{h'})_{k_s} = p(a_j)_{x_h} \times p(x_h \rightarrow x_{h'})_{k_s}, \quad (2)$$

где:

$e_s(x_h \rightarrow x_{h'})_{k_s}$ – эффективность управленческих воздействий;

$p(a_j)_{x_h}$ – вероятность перехода в состоянии x_h для направления антикризисных мероприятий a_j ;

$p(x_h \rightarrow x_{h'})_{k_s}$ – вероятность перехода в новое состояние при определенном управленческом воздействии.

Максимальное значение меры эффективности или оптимальное антикризисное регулирование достигается определением набора максимальных состояний e_s^* :

$$e_s^* = \max e_s(x_h \rightarrow x_{h'})_{k_s}. \quad (3)$$



Если предположить, что антикризисное регулирование Y является аддитивной функцией, которая определяется эффективностью управленческих воздействий, предпринимаемых в рамках реализации определенной стратегии регулирования e_j , то для определения оптимальной совокупности управленческих воздействий можно использовать принцип Беллмана оптимальности многошагового процесса принятия решения.

При этом оптимальная совокупность управленческих воздействий обладает тем свойством, что каково бы ни было первоначально используемое управленческое воздействие на состояние регулируемого процесса, последующее воздействие должно быть оптимальным относительно первоначально установленного воздействия:

$$e_j = \max_{s=1}^R [e_{sj} (x_{hj} \rightarrow x_{h'})_{ks} + e_{sj} (x_{h'} \rightarrow x_{h''})_{ks}] \quad (4)$$

Схематично процесс формирования и реализации антикризисного регулирования муниципального развития на основе описанной модели может быть представлен в виде блок-схемы отбора управленческих решений (рис. 2). Реализация представленной схемы антикризисного регулирования органами местного самоуправления может быть осуществлена на базе использования экспертных технологий, основывающейся на результатах мониторинга при диагностике степени кризисных явлений в различных секторах муниципальной экономики.

Обобщая вышесказанное, следует отметить, что состав мер антикризисного регулирования регионального и местного уровня³ может быть определен следующим образом:

1. Регулярное наблюдение за финансово-экономическим состоянием региона и муниципальных образований с целью раннего обнаружения признаков его кризисного развития. Для этого необходима реализация проекта «Мониторинг социально-экономического развития муниципального образования», который входит составной частью в подобный проект, реализуемый на региональном уровне.

2. Определение масштабов кризисного состояния. При обнаружении существенных отклонений от нормального хода финансовой деятельности значительного объема предприятий, определяемого системой плановых и нормативных финансовых показателей, выявляются масштабы кризисного состояния, т.е. распространения угрозы банкротства в массовом масштабе. Такая идентификация кризисного состояния позволяет осуществлять соответствующий селективный подход к выбору системы механизмов защиты от возможных негативных последствий кризиса.

3. Диагностика кризиса. Она предполагает изучение основных факторов, обусловивших развитие негативных тенденций в экономике, и включает работу по следующим направлениям:

- предварительная группировка факторов по основным определяющим признакам;
- исследование степени влияния отдельных факторов на формы и масштабы кризисного развития;
- прогнозирование развития факторов, оказывающих кризисное воздействие.

³ Следует отметить, что говорить только о местном антикризисном регулировании сложно, т.к. на местном уровне недостаточно ресурсов для проведения полномасштабных антикризисных мероприятий. Поэтому необходимы меры антикризисного регулирования со стороны региональных властей.

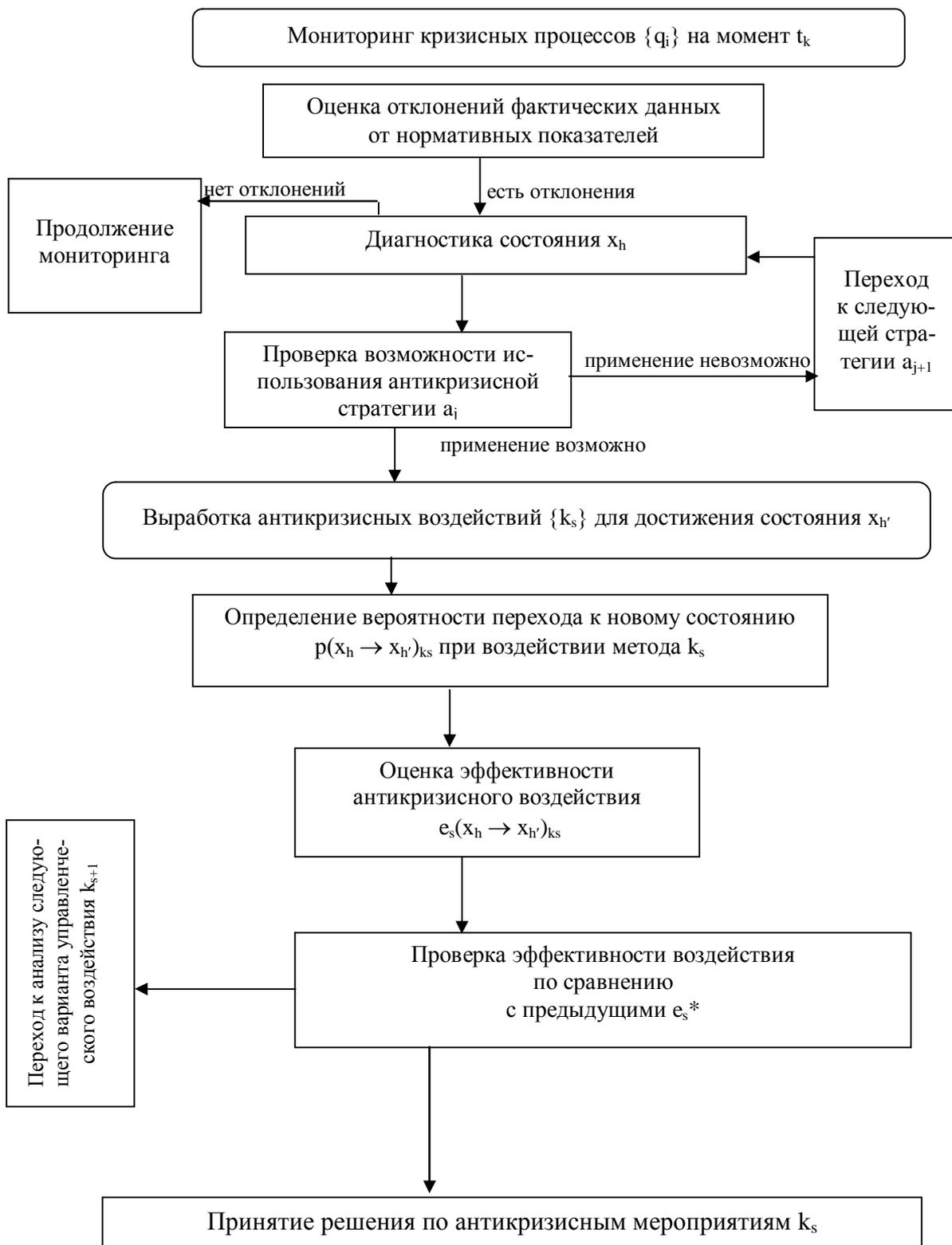


Рис. 2. Модель антикризисного регулирования муниципального развития

Диагностика кризиса представляет собой синтез результатов изучения вопросов: 1) исследование состояния муниципальной и региональной экономики с



целью раннего обнаружения признаков его кризисного развития; 2) определение масштабов кризисного состояния предприятий региона и муниципального образования; 3) выявление основных факторов, обуславливающих кризисное развитие экономики.

4. Формирование целей и выбор основных механизмов антикризисного регулирования. Цели и механизмы антикризисного регулирования должны соответствовать масштабам кризисного состояния и учитывать прогноз развития основных факторов, определяющих угрозу осуществления кризиса.

5. Внедрение механизмов финансовой стабилизации. Механизмы финансовой стабилизации должны обеспечить реализацию срочных мер по возобновлению платежеспособности и восстановлению финансовой устойчивости местных предприятий за счет внутренних резервов и внешней (при необходимости) финансовой помощи. Эти механизмы основаны на последовательном использовании определенных моделей управленческих решений, выбираемых в соответствии со спецификой хозяйственной деятельности предприятий и масштабами кризисных явлений в экономике муниципального образования.

Если масштабы кризисного финансового состояния предприятий не позволяют выйти из него за счет реализации внутренних резервов, предприятия вынуждены прибегнуть к внешней помощи, которая обычно принимает форму его санации. В процессе санации необходимо обосновать выбор наиболее эффективных ее форм (включая формы, связанные с реорганизацией предприятия) с тем, чтобы в возможно более короткие сроки достичь финансового оздоровления и не допустить объявления банкротства предприятия.

Реализация представленных этапов антикризисной деятельности со стороны местных органов власти требует четкого определения нормативно-правовых основ проведения подобной работы.

Следует отметить, что органы местного самоуправления не могут самостоятельно проводить комплекс антикризисных мер, необходимо участие государства в лице федеральных и региональных органов государственной власти. Только комплекс антикризисных мероприятий может помочь местной экономике пережить сложное время и выйти на траекторию устойчивого роста.

BASIC PROPOSITIONS OF CRISIS REGULATION OF MUNICIPAL DEVELOPMENT

O. P. OVCHINNIKOVA¹⁾
N. N. KURILENKO²⁾

¹⁾ *Orel Regional Academy
of Government Service*

²⁾ *Belgorod State University*

e-mail: kurilenko@bsu.edu.ru

This article is devoted to fundamental definitions and mechanism of crisis regulation of municipal development. It deals with the system of direct and indirect methods of crisis regulation. The author suggests measures for with drawl of social sphere of municipal organizations from the recession, its stabilization and further development. The source data defined the formation of crisis regulation of municipal development are introduced by such model where basic operation factors are marked. These basic operation factors influence on stabilization of local enterprises' economical activity.

Key words: municipal development, crisis regulation, methods and mechanisms of economic development stabilization, the model of crisis regulation of municipal development.

ПРОБЛЕМЫ МЕТОДИКИ ОЦЕНКИ РЕГИОНАЛЬНЫХ ИНВЕСТИЦИОННЫХ ПРОЕКТОВ**Е. Н. ПАРФЕНОВА***Белгородский
государственный
университет**e-mail:
parfenova@bsu.edu.ru*

В статье предложено создание оптимальных критериев анализа региональных инвестиционных проектов, которые позволят учитывать индивидуальные предпочтения и приоритеты каждого субъекта Российской Федерации, что даст возможность выбрать наилучший проект, направленный на укрепление и рост региональной экономики, ориентированной на повышение уровня и качества жизни населения.

Ключевые слова: региональный инвестиционный проект, критерии оценки, показатели общественной эффективности, индивидуальные предпочтения и приоритеты региона.

На современном этапе российских преобразований основной целью управления региональной экономикой в долгосрочной перспективе является создание условий для устойчивого роста и развития территориальной экономической системы, т.е. создание в субъекте Российской Федерации режима функционирования, ориентированного на позитивную динамику параметров уровня и качества жизни населения, обеспеченную устойчивым, сбалансированным и многофакторным воспроизводством социального, хозяйственного, ресурсного и экологического потенциалов территории. Достижение этой цели невозможно без привлечения инвестиций в региональную экономику¹. Инструментом их реализации является инвестиционный проект, с помощью которого выясняется и обосновывается возможность и экономическая целесообразность создания объекта избранной целевой направленности в решении социально-экономических проблем субъекта Российской Федерации. При этом критериями оценки экономической эффективности инвестиционного проекта являются как количественные, так и качественные. Количественные критерии имеют денежную оценку, качественные не поддаются стоимостному выражению, их нельзя суммировать с денежной оценкой, но они должны учитываться в рамках инвестиционных расчетов. В таких условиях актуальной остается разработка возможных критериев анализа общественной эффективности региональных инвестиционных проектов, с помощью которых будут более точно учтены и оценены все аспекты экономической эффективности, что позволит выявить наиболее значимые и приоритетные проекты для субъекта Российской Федерации.

Оценка эффективности региональных проектов должна играть определяющую роль при выборе возможных вариантов вложения бюджетных денежных средств в реализацию проектных решений, что позволит исполнительной власти судить о том, какой ценой достигаются поставленные цели².

Одна из распространенных методик определения эффективности инвестиционных проектов включает следующую систему показателей:

- показатели общественной (социально-экономической) эффективности, учитывающие затраты и результаты, связанные с реализацией этого проекта с точки зрения интересов всего народного хозяйства региона. К ним относятся: доходы и занятость населения, подготовка кадров, условия труда работающего населения, экологическая обстановка в регионе и другие социальные результаты, характеризующие качество жизни местного населения;
- показатели коммерческой эффективности, учитывающие финансовые последствия его осуществления для региона. Основными показателями являются: чис-

¹ Пчелинцев О. С. Региональная экономика в системе устойчивого развития. М.: Наука, 2004. С. 117.

² Мухетдинова Н. Проблемы инвестиционного роста российских регионов // Финансы, Деньги, Инвестиции. 2008. № 1. С. 15.

тый дисконтированный доход, внутренняя норма доходности, индексы доходности затрат и инвестиций и срок окупаемости проекта;

- показатели бюджетной эффективности, учитывающие влияние результатов осуществления проекта на расходы и доходы регионального бюджета, главным из которых является бюджетный эффект, определяющийся как разница между притоками и оттоками бюджетных средств с учетом дисконтирования³ (см. рис. 1).



Рис. 1. Система оценочных показателей эффективности регионального инвестиционного проекта

При оценке и выборе наилучшего из вариантов инвестиционных проектов власти региона основываются на определенной системе критериев. Критерии коммерческой эффективности основываются на минимизации затрат по их разработке и осуществлению: абсолютные (разность между суммой капиталовложений и денежной оценкой его результатов), относительные (отношение денежной оценки результатов и совокупных затрат), временные (период возврата инвестиций). Критерий оценки бюджетной эффективности характеризуется положительным сальдо регионального бюджета. Критерии коммерческой и бюджетной эффективности отражают экономические интересы властей региона по поводу соотношения затрат и результатов в связи с вложением бюджетных средств в реализацию инвестиционных решений и выражаются в стоимостной оценке. Однако они позволяют определить лишь средства, необходимые для выполнения намеченных целей. Критерии оценки показателей общественной эффективности дают представление о качественной стороне дос-

³ Бобылев Г., Кузнецов А., Горбачева Н. Условия и факторы реализации инновационного потенциала региона // Регион: экономика и социология. 2008. № 1. С. 117.

тигнутых результатов, ради которых региональный инвестиционный проект разрабатывался и осуществлялся: увеличение доходов населения, улучшение условий труда, увеличение количества рабочих мест в регионе и другие социальные результаты, которые в конечном итоге повысят качество жизни людей. Эти критерии очень важны для органов исполнительной власти, и их необходимо учитывать.

При этом следует иметь в виду, что методически сложно объединить эффективность финансовых результатов властей региона и социальных результатов общества для того, чтобы принять окончательное решение о целесообразности осуществления конкретного инвестиционного проекта. Такая методика используется многими субъектами Российской Федерации для того, чтобы рассмотреть и одобрить инвестиционные проекты, планируемые к реализации в своем регионе⁴.

Оценка эффективности инвестиционных проектов в Белгородской области осуществляется Инвестиционным советом при губернаторе области в соответствии с вышеизложенной методикой. При этом Правительство Белгородской области в качестве приоритетных рассматривает развитие в следующих направлениях:

- горнодобывающая и металлургическая отрасли;
- агропромышленный комплекс;
- малое и среднее предпринимательство.

Российский внутренний рынок в настоящее время потребляет всего около 50 % произведенного проката, но имеет значительный потенциал роста при изменении сортаментной структуры, отвечающей возрастающим запросам металлопотребляющих отраслей. Учитывая потребность в дальнейшем развитии горно-металлургической отрасли, инвестиционные проекты ведущих предприятий (проект ОАО «Оскольский электрометаллургический комбинат» стоимостью более 450 млн. долларов США по строительству мелкосортно-среднесортного стана мощностью 1 млн. т проката в год; проект ОАО «Лебединский горно-обогатительный комбинат» стоимостью около 800 млн. долларов США по строительству первой и второй очереди цеха горячебрикетированного железа) направлены на создание производств с использованием новейших технологий по добыче и переработке сырья. Реализация этих проектов позволит обновить ассортимент выпускаемой продукции и увеличить ее качество и конкурентоспособность на внутреннем и внешнем рынках.

Инвестиционные проекты в агропромышленном комплексе направлены на формирование эффективного сельскохозяйственного производства, выступающего конкурентоспособным участником внутрироссийского и мирового рынков и обеспечивающего потребности населения области и перерабатывающей промышленности в основных видах сельскохозяйственной продукции и сырья. В последние годы в Белгородской области реализуются такие инвестиционные проекты, как «Развитие свиноводства в Белгородской области на 2005-2010 годы», «Развитие птицеводства в Белгородской области на период с 2005 по 2010 годы» и «О мерах по увеличению молока в Белгородской области». В ходе реализации этих инвестиционных проектов в области планируется повысить производство мяса в 1,8 раза, а молока – в 2,2 раза. В целом агропромышленный сектор Белгородской области должен обеспечить мясом и молоком как минимум 10% потребительского рынка России.

Малый бизнес в регионе представляет собой эффективно развивающийся сектор экономики. В последние годы в Белгородской области реализуется инвестиционный проект «Реализация иных мероприятий субъектов Российской Федерации по поддержке и развитию малого предпринимательства», направленный на создание условий для самореализации экономически активных слоев населения и развитие конкурентоспособных малых предприятий в производственной и инновационной сферах.

⁴ Илларионов А. Проблемы социально-экономического развития муниципальных образований // Рынок ценных бумаг. 2007. № 1. С. 58.

Оценивая значимость инвестиционных проектов в приоритетных сферах экономики, критериями отбора у властей Белгородской области стали результаты общественной, коммерческой и бюджетной эффективности.

Так, критериями коммерческих результатов явились положительные значения показателей чистого дисконтированного дохода и рентабельности. Для инвестиционных проектов горно-металлургического сектора они составили, соответственно, около 2,8 млрд. рублей и 30%, для агропромышленного сектора – 4,6 млрд. рублей и 37% и для малого бизнеса – 1,02 млрд. рублей и 28%.

Бюджетная эффективность также показала положительное влияние реализации инвестиционных проектов в этих отраслях экономики на бюджет области. Критерием отбора этих инвестиционных проектов стало ежегодное пополнение доходов областного бюджета от их реализации:

- в горно-металлургической отрасли – на сумму свыше 600 млн. рублей;
- в агропромышленном секторе – на сумму около 800 млн. рублей;
- в малом бизнесе – на сумму около 1,48 млрд. рублей.

При оценке общественной эффективности этих инвестиционных проектов критериями отбора стали улучшение социальных аспектов, сравниваемых с достигнутым ранее уровнем по таким показателям, как количество рабочих мест, изменение экологической ситуации, уровень безопасности и условия труда, жизненный уровень населения региона и др.

Государственная поддержка инвестиционных проектов по этим направлениям включает использование следующих механизмов:

- заключение соглашений о социально-экономическом сотрудничестве между правительством области и инвестором, сопровождение реализации проектов; содействие в выделении земельных участков под промышленные площадки;
- предоставление гарантий Белгородской области и государственного имущества области для обеспечения привлекаемых инвестором кредитных ресурсов;
- участие средств областного бюджета в формировании уставных капиталов создаваемых предприятий; субсидирование процентной ставки по привлекаемым банковским кредитам;
- обустройство строящихся и реконструируемых животноводческих комплексов инженерной инфраструктурой: подведение инженерных сетей за счет средств газо- и энергоснабжающих организаций, строительство подъездных дорог за счет средств областного бюджета.

Однако, используемая методика оценки эффективности инвестиционных проектов не позволяет в полной мере властям Белгородской области сопоставить свои финансовые затраты и социальные результаты для населения и выявить наиболее значимые и приоритетные проекты, поэтому размер их государственной поддержки в Белгородской области одинаков.

Данное обстоятельство побуждает для определения результативности бюджетных вложений и социальных последствий при реализации инвестиционных проектов в субъектах Российской Федерации разработать другие, оптимальные критерии оценки общественной эффективности, которые позволят оценить социальный эффект в стоимостном выражении, как и экономическую эффективность. Определение реальности достижения именно таких результатов от реализации регионального инвестиционного проекта является ключевой задачей оценки их эффективности⁵.

Для объективного анализа инвестиционных проектов субъекта Российской Федерации необходимо ввести шкалу критериев для системы показателей общественной эффективности, которая может изменяться в зависимости от особенностей конкретного региона и приоритетов его развития. Относительная значимость каждого показателя в сравнении с другими из системы показателей общественной эффек-

⁵ Господарчук Г.Г. Развитие регионов на основе финансовой интеграции. М.: Финансы и статистика, 2006. С.54.

тивности может отражаться весовыми коэффициентами, учитывая развитие экономики и особенности каждого конкретного субъекта Российской Федерации. Вычисление этих весовых коэффициентов осуществляется экспертной комиссией, созданной властями региона с учетом целей, задач, ситуации и прочих внутренних и внешних факторов, формирующих приоритеты для области (см. табл. 1).

Таблица 1

Предлагаемая шкала оценки показателей общественной эффективности

Показатель	Критерий оценки показателя	Вес показателя
1. Качество жизни населения региона	1- падение более чем на 10% 10 – рост более чем на 10%	В зависимости от уровня и приоритета каждого региона
2. Оплата труда работников	1- падение более чем на 10% 10 – рост более чем на 10%	"
3. Номинальные и реальные доходы населения региона	1- падение более чем на 10% 10 – рост более чем на 10%	"
4. Условия труда и техники безопасности	1- падение более чем на 10% 10 – рост более чем на 10%	"
5. Число рабочих мест, занятость населения	1- падение более чем на 10% 10 – рост более чем на 10%	"
6. Безработица	1- рост более чем на 10% 10 – падение более чем на 10%	"
7. Ситуация здоровья и демографии в регионе	1- падение более чем на 10% 10 – рост более чем на 10%	"
8. Заболеваемость	1- рост более чем на 10% 10 – падение более чем на 10%	"
9. Экологическая обстановка	1- падение более чем на 10% 10 – рост более чем на 10%	"
10. Численность работников, которые проходят переподготовку и получают новую профессию	1- падение более чем на 10% 10 – рост более чем на 10%	"
11. Число рабочих мест с тяжелыми, вредными и опасными условиями труда	1- рост более чем на 10% 10 – падение более чем на 10%	"
12. Количество профессиональных заболеваний и производственного травматизма	1- рост более чем на 10% 10 – падение более чем на 10%	"

Для того, чтобы сравнить показатели, характеризующие общественную значимость для региона, предлагается десятибалльная шкала измерения. Так, критерий оценки показателя, равный единице, будет характеризовать ухудшение достигнутого раннее уровня по этому показателю более чем на 10%, а критерий оценки показателя, равный десяти – рост уровня этого показателя более чем на 10%. Значения весовых коэффициентов также могут варьироваться в зависимости от уровня развития, приоритетов и особенностей каждого конкретного региона. Таким образом, данная система критериев адаптирована и максимально приближена к интересам каждого субъекта Российской Федерации.

После определения критерия оценки каждого показателя определяется комплексное значение критерия оценки регионального инвестиционного проекта по формуле:

$$K_{интегр} = \frac{\sum_{i=1}^n k_i g_i}{\sum_{i=1}^n g_i}, \quad (1)$$

где

Кинтегр – интегральный критерий оценки системы показателей общественной эффективности;

n – число показателей;

k_i – критерий оценки каждого показателя;

g_j – вес каждого показателя.

В зависимости от полученного значения критерия оценки системы показателей общественной эффективности будет выбираться наилучший вариант из альтернативы рассматриваемых инвестиционных региональных проектов.

Таким образом, можно сделать вывод о том, что предложенная интегральная оценка общественной эффективности позволит более точно оценить приоритетные социальные аспекты и индивидуальные предпочтения каждого субъекта Российской Федерации, вычислить их в стоимостном выражении и рассмотреть совместно с количественными показателями коммерческой и бюджетной эффективности. Это поможет властям региона учесть свои интересы и особенности, а также результаты с точки зрения народнохозяйственного эффекта и принять окончательное решение о целесообразности осуществления конкретного инвестиционного проекта.

PROBLEMS OF THE ESTIMATION MECHANISM OF THE REGIONAL INVESTMENT PROJECTS

E. N. PARFENOVA

Belgorod State University

e-mail:

parfenova@bsu.edu.ru

The author suggested the most suitable criteria of the analysis of the regional investment projects which could help to account the individual preferences and priorities of every subject of the Russian Federation. These estimation criteria will give an opportunity to choose the best project, the main idea of which is concerned with strengthening and growth of the regional economy connected with the increasing standard of life of the population.

Key words: regional investment project, estimation criteria, social effectiveness indices, individual preferences and priorities of the region.

ВНЕШНЕЭКОНОМИЧЕСКАЯ ДЕЯТЕЛЬНОСТЬ КАК СРЕДСТВО ГАРМОНИЗАЦИИ РАЗВИТИЯ ИНФОРМАЦИОННОГО ПРОСТРАНСТВА

Е. В. ПОПОВА

*Магнитогорский
государственный
университет*

*e-mail:
science_masu@mail.ru*

Рассматривается внешнеэкономическая деятельность как средство гармонизации развития информационного пространства России, одной из особенностей которого является неоднородность, обусловленная отличиями в социально-экономических условиях развития регионов. Приводятся структура информационного пространства, этапы и направления гармонизации его развития. Определяется роль информационного бизнеса во внешнеэкономической деятельности.

Ключевые слова: информационное пространство, внешнеэкономическая деятельность, гармонизация.

В настоящее время формирование информационного пространства России является одним из приоритетных направлений государственной политики. Так, в Концепции формирования и развития единого информационного пространства России и соответствующих государственных информационных ресурсов¹ указано, что «формирование информационного пространства России в интересах органов государственной власти должно быть направлено на объединение и развитие существующих информационно-аналитических ресурсов, предназначенных для обеспечения их эффективной управленческой деятельности». Это предусматривает, в первую очередь, обеспечение оперативного доступа к имеющимся информационным ресурсам и проведение работ по их включению в единое информационное пространство. Вновь формируемые информационные ресурсы должны быть на законном основании доступны органам управления государственной власти, хозяйствующим субъектам и гражданам. Таким образом, целью формирования и развития информационного пространства России является упрощение процесса получения информации, а также максимально полное удовлетворение потребностей информационных субъектов в получении информации, а именно:

- открытый доступ к правовым базам данных;
- равнодоступность и открытость государственной информации;
- доступ к информации о властных структурах, как центральных, так региональных и местных;
- свободный доступ к научной и технической информации;
- свободный доступ к информации по интересам.

В настоящее время доминирующим становится представление об информационном пространстве как о комплексе компонентов различной природы, в соответствии с которым информационное пространство в наиболее широком смысле – это все то, что связано с хранением и целенаправленной обработкой информации². В более узком смысле информационное пространство понимается как совокупность: 1) банков и баз данных, 2) технологий их сопровождения и использования; 3) инфор-

¹ Концепция формирования и развития единого информационного пространства России и соответствующих государственных информационных ресурсов (одобрена решением Президента РФ от 23.11.1995 N Пр-1694). Режим доступа: <http://expert.consultant.ru/doc240699.html>.

² Райков А.Н. Развитие России и единое информационное пространство // Вестник РФФИ: Наука и информационное общество (специальный выпуск) [Электрон, ресурс]. 1999. №3. Режим доступа: http://www.rfbr.ru/default.asp?doc_id=5212.

мационных телекоммуникационных систем, функционирующих на основе общих принципов и обеспечивающих информационное взаимодействие организаций и граждан, удовлетворение их информационных потребностей³.

В рамках данного подхода в качестве основных подсистем информационного пространства следует рассматривать:

- 1) информационные ресурсы, содержащие сведения и знания, данные, зафиксированные на соответствующих носителях информации;
- 2) организационные структуры, обеспечивающие функционирование и развитие единого информационного пространства, в частности, создание, сбор, обработку, хранение, распространение, доступ к информации;
- 3) средства информационного взаимодействия граждан и организаций, обеспечивающие им доступ к информационным ресурсам на основе соответствующих информационных технологий, включающие программно-технические средства и организационно-нормативные документы.

Организационные структуры и средства информационного взаимодействия образуют информационную инфраструктуру. И.Ф. Максимычев и В.Г. Машлыкин отмечают, что информационные ресурсы как компонент информационного пространства должны быть представлены в виде автоматизированных баз данных. Кроме того, данные авторы указывают также и на необходимость выделения в качестве компонента информационного пространства прикладных лингвистических средств — классификаторов и тезаурусов⁴. По мнению С.В. Шаповой, информационное пространство следует рассматривать как человеко-машино-информационную систему: «человека — как с позиции его профессиональной подготовки, так и его физических возможностей по восприятию информации; машину — как совокупность, технического и программного обеспечения; информацию — как организацию информационных потоков, каждый из которых характеризуется рядом показателей, таких как направление потока, количественный и качественный его состав»⁵. Таким образом, в данном определении к числу компонентов информационного пространства (наряду с уже перечисленными выше) дополнительно относится человек.



Рис. 1. Основные объекты единого информационного пространства

³ Информационная политика: Учебник / Под общ. ред. В.Д. Попова. М.: Издательство Российской академии государственной службы при Президенте Российской Федерации, 2003. 463 с.

⁴ Максимычев И.Ф., Машлыкин В.Г. Европейское информационное пространство // Современная Европа (общественно-политический ежеквартальный журнал) [Электрон, ресурс] / Институт Европы РАН. Апрель-июнь 2000 г. Вып. 2. Режим доступа: <http://www.ieras.ru/journal/journal2.2000/4.htm>.

⁵ Шапова С.В. Организационно-технологическое обеспечение системы поддержки принятия решений при формировании единого информационного пространства предприятия (на примере гальванического производства). Автореф. дисс... канд. технич. наук. Кострома, 1999. 18 с.

На региональном уровне целью формирования и развития информационного пространства является создание и поддержание необходимого для устойчивого развития хозяйственных систем уровня информационного потенциала на основе предметной ориентации информационных ресурсов, запасов и резервов, что позволит обеспечить согласованность решений, принимаемых органами государственной власти различных уровней⁶.

Следует отметить, что информационное пространство регионов в настоящий момент развивается неравномерно. Это обусловливается неравенством их возможностей и перспектив, нелинейным характером социальных и политических изменений современного общества. Информационное пространство охватывает все общности, слои и социальные коммуникации, а также и экономическое, социальное, политическое пространства, участвующие в процессе накопления и передачи информации. В условиях глобализации экономики в качестве инструмента гармонизации информационного пространства региона может выступить его внешнеэкономическая деятельность.

Внешекономическая деятельность региона, осуществляясь по нескольким направлениям (внешняя торговля товарами, технико-экономическое и научно-техническое сотрудничество) обладает большим потенциалом по следующим направлениям:

- развитие информационной составляющей деятельности за счёт увеличения её нематериальных и высокотехнических объектов;
- информатизация практической работы по обеспечению эффективной и устойчивой внешнеэкономической деятельности субъектов экономических отношений;
- подготовка и повышение квалификации специалистов в области информатизации внешнеэкономической деятельности, работающих на всех уровнях хозяйственно-финансовой деятельности региона.

Как правило, под информатизацией внешнеэкономической деятельности подразумевается создание комплексной информационной системы внешнеэкономической деятельности и мониторинга экспортно-импортных операций; формирование единой информационной базы данных по тендерам; введение международных стандартов бухгалтерского и статистического учета; подключение субъектов внешнеэкономической деятельности к международным информационным сетям. При этом к основным концептуальным положениям в области внешнеэкономической деятельности относят:

- развитие новых информационных технологий, систем и средств коммуникаций, формирование банков данных, позволяющих вести общение на международном уровне с использованием Интернет;
- создание конкурентоспособной на зарубежных рынках научно-технической продукции;
- овладение навыками маркетинговых исследований, позволяющих объективно оценить потребность в новых изделиях, соотнести их с мировым уровнем и сохранить на этой основе «ноу-хау» или реализовывать их по лицензиям;
- совместное выполнение международных проектов и программ;
- разработка информационных технологий поддержки производственной, экономической и коммерческой деятельности хозяйствующих субъектов (предоставление информации по рынку труда, товарному рынку, фондовому рынку, законодательным актам и т.д.);
- информационные технологии схем формирования развития и размещения комплексов.

Обмен информацией и другие виды информационной активности, являются основой реализации экономических процессов. Более совершенная среда информа-

⁶ Информационная политика: Учебник / Под общ. ред. В.Д. Попова. М.: Издательство Российской академии государственной службы при Президенте Российской Федерации, 2003. 463 с.



ционной активности, снижая неопределенность хозяйственной деятельности субъектов, позволяет им достигать более высокого уровня возможной эффективности использования ресурсов. Развитие Internet позволяет вовлекать в обмен все больше информационных ресурсов, использование которых обеспечивает количественное и качественное изменение структуры затрат и предъявляет новые требования к технологической составляющей экономического прогресса. Необходимость сокращения трансформационных и трансакционных издержек, связанных с использованием информации, обуславливает развитие информационного пространства.

Информация оказывает существенное влияние на все отрасли производства и инфраструктуры; в действии рыночного механизма информационного общества также появляются новые черты, отличающие его от рынка индустриальной эпохи. Развитие технологий обработки информации привело к появлению новой разновидности экономической деятельности – информационному бизнесу, который охватывает деловые отношения, складывающиеся в сферах производства, сбыта и приобретения информации.

Информационный бизнес имеет две формы существования. С одной стороны, он является неперенным атрибутом деятельности любого субъекта бизнеса вне зависимости от того, к какому именно направлению принадлежит эта деятельность. С другой стороны, он выступает как самостоятельное направление бизнеса, занятие которым предполагает известную специализацию работающих в информационном секторе экономики предпринимательских фирм. Информационный бизнес выстраивается вокруг специфических объектов деловых отношений и деловой деятельности, а именно – информационных ресурсов, творческих ресурсов, информационных продуктов и информационных услуг⁷.

Переходя к инфраструктуре информационной отрасли, необходимо сразу отметить, что сфера информационного бизнеса очень обширна и включает различные виды деятельности, которые прямо или косвенно связаны с информацией. Стремительное развитие в последние десятилетия получили предпринимательские фирмы, специализирующиеся на бизнесе в области информационных технологий, пакетов прикладных программ и телекоммуникаций. Бизнес в области изготовления программных продуктов и проектирования локальных информационных систем превращается в чрезвычайно значимое явление на фоне бурного развития информационной среды деловых отношений, а Интернет и другие глобальные сети становятся каналом доступа к огромным массивам сторонней информации и вместе с тем предоставляют субъектам бизнеса возможность размещения собственной информации, которая становится доступной миллионам людей одновременно.

Специфической разновидностью информационного бизнеса является образовательный бизнес⁸.

В области информатизации практической работы по обеспечению эффективной и устойчивой внешнеэкономической деятельности субъектов экономических отношений целесообразно обеспечить широкий доступ жителей региона, органов государственного управления и субъектов хозяйственной деятельности к национальным информационным ресурсам в согласованных сферах и мировым информационным фондам. При этом особенно важно усилить взаимодействие информационных систем, обеспечивающих постоянное развитие межгосударственных информационных обменов по согласованным сферам деятельности.

Региональная специфика информационного пространства определяется природными условиями хозяйственной системы, ее масштабом, объектами, субъектами, сферами деятельности, уровнем развития технологий, нормами и правилами, количеством и структурой организаций, способами информационного представления хозяйственной деятельности, что находит отражение в соответствующей информаци-

⁷ Рубин Ю.Б. Основы бизнеса. 8-е изд., перераб. и доп. М.: Маркет ДС. 2008. 320 с.

⁸ Там же.

онной системе. Архитектура этого информационного пространства включает информационные ресурсы и транзакционную инфраструктуру.

Развитие информационного пространства региона, таким образом, предполагает три этапа.

Этап 1. Элементарный информационный обмен, при котором определяются первоначальный перечень информационных ресурсов, состав и структура их информационных показателей, основные критерии и принципы гармонизации информационных ресурсов, начинается разработка регламента обмена информационными ресурсами и доступа к ним. Информация на этом этапе может пересылаться по электронной почте и публиковаться на Web-серверах для совместного доступа.

Этап 2. Активный информационный обмен, характеризующийся последовательным сближением нормативно-правового обеспечения. На этом этапе разрабатываются унифицированные протоколы обмена информационными ресурсами, утверждается регистр информационных ресурсов и пользователей, вырабатываются общие требования к автоматической системе обмена информационными ресурсами, проводится унификация методов и средств обмена информационными ресурсами, создается система обработки запросов и предоставления информационных сервисов.

Этап 3. Полномасштабное информационное взаимодействие по всем направлениям внешнеэкономической деятельности, подразумевающее постоянный обмен данными с использованием единых организационно-технологических схем. Этап является перспективным, однако уже на первом этапе необходимо заложить такие проектные решения, которые позволят успешно перейти от элементарного информационного обмена к полномасштабному информационному взаимодействию⁹.

При этом основные задачи формирования и развития информационного пространства региона можно сформулировать как:

- создание и поддержка необходимого для устойчивого развития хозяйственных систем уровня информационного потенциала на основе предметной ориентации информационных ресурсов, запасов и резервов, что позволит обеспечить согласованность решений, принимаемых органами государственной власти различных уровней;
- создание правовых, политических, экономических и технических условий для межгосударственной кооперации в сфере информации и информатизации, согласование механизмов осуществления межгосударственного информационного обмена.

Конкретное содержание совместных межгосударственных мероприятий по формированию информационного пространства региона должно определяться национальными интересами¹⁰.

Следует отметить, что важной частью информатизации внешнеэкономической деятельности выступает подготовка и повышение квалификации в этой области специалистов, работающих на всех уровнях хозяйственно-финансовой деятельности региона. Проблема информационного неравенства чрезвычайно многоаспектна, но экономические, технические, содержательные проблемы напрямую связаны с недостатком образовательного потенциала людей в сфере информационных технологий.

В процессе формирования информационного пространства становится необходимо наличие организационных структур, интегрирующих все информационно-образовательные ресурсы, стимулирующих информационные взаимосвязи участников экономической деятельности региона, осуществляющих информационный обмен между образовательными, информационными и административными структурами. Между образованием людей и их информированием как двумя видами деловой деятельности существует следующая взаимная связь. Осваивая информационные

⁹ Лаврухин С. Создание единого информационного пространства таможенных служб стран-участников СНГ. Режим доступа: http://www.rdtex.ru/win/root/article_ceip.html.

¹⁰ Информационная политика: Учебник / Под общ. ред. В.Д. Попова. М.: Издательство Российской академии государственной службы при Президенте Российской Федерации, 2003. 463 с.



продукты, пользуясь информационными услугами, привлекая информационные ресурсы, потребители в любом случае приобретают новые знания, становятся более подготовленными к продолжению и развитию своей деятельности.

Интернет как средство группового взаимодействия, с его массовой интерактивностью, общедоступностью содержания коммуникации, возможностями сохранения и использования результатов коммуникации, интегрируемостью, оперативностью позволяет объединить информационные ресурсы образовательных учреждений, сделав медиатеку мощным образовательным ресурсом Сети¹¹.

Таким образом, внешнеэкономическая деятельность выступает инструментом гармонизации развития информационного пространства региона. Предприятия, выступая в качестве партнёров в совместных проектах, вынуждены постоянно развивать свои информационные ресурсы, ориентируясь на уровень развития информационных технологий в мире. Кроме того, между участниками внешнеэкономической деятельности происходит постоянный обмен технологиями, информацией, который позволяет сгладить неравенство ИТ-развития.

FOREIGN TRADE ACTIVITIES AS A MEANS OF HARMONIZATION OF DEVELOPMENT OF THE INFORMATION SPHERE

E. V. POPOVA

*Magnitogorsk State
University*

*e-mail:
science_masu@mail.ru*

Foreign trade activities as means of harmonisation of development of an information field of Russia are considered, one which their features the heterogeneity caused by differences in social and economic conditions of development of regions is. The information field structure, stages and directions of harmonisation of its development are resulted. The role of information business in foreign trade activities is advanced.

Key words: information field, foreign trade activities, harmonization.

¹¹ Кузнецова Н.М., Гусакова Т.М. Развитие регионального информационного образовательного пространства. Режим доступа: <http://www.ito.su/2002/IV/IV-0-711.html>.

ИНВЕСТИЦИИ И ИННОВАЦИИ

УДК 303.732.4

ВЛИЯНИЕ РЕСТРУКТУРИЗАЦИИ НА ДОХОДНОСТЬ ДОЛГОСРОЧНОЙ КРЕДИТНОЙ ОПЕРАЦИИ

М. Ф. ТУБОЛЬЦЕВ
В. И. БОЛТЕНКОВ

*Белгородский
государственный
университет,*

e-mail: tuboltsev@bsu.edu.ru

e-mail: boltenkov@bsu.edu.ru

Рассмотрены вопросы изменения доходности долгосрочной кредитной операции при ее реструктуризации. Показано, что только при использовании постоянной сложной процентной ставки реструктуризация не меняет доходности. Использование номинальных процентных ставок и принципа эквивалентности платежей при реструктуризации приводят к существенному изменению доходности кредитования. На изменение доходности при использовании номинальных ставок оказывают влияние практически все параметры реструктуризации, что делает затруднительным полный контроль над доходностью. При определенных условиях доходность операции долгосрочного кредитования может достигать эффективного процента, вычисленного по номинальной ставке.

Ключевые слова: долгосрочный кредит, финансовая операция, финансовые системы, методика агрегирования финансовых операций, реструктуризация финансовых операций, принцип эквивалентного изменения платежей, принцип эквивалентного изменения финансовых систем, реструктуризация финансовых систем.

Проблема выплаты долгосрочного кредита многогранна и может рассматриваться с разных сторон. В условиях финансово-экономического кризиса вопросы реструктуризации долгосрочного кредита приобретают большое значение. Часть этих вопросов (применительно к ипотечным кредитам) затронута в работе авторов¹. Однако, вопрос влияния реструктуризации на доходность (для кредитора) долгосрочного кредита остался не до конца выясненным.

В данной работе рассматриваются долгосрочные кредитные операции общего вида, а не только ипотечные операции. Следует отметить, что ипотечные кредитные операции, хотя формально относятся к типу долгосрочных кредитных операций, но обладают рядом специфических особенностей, затрудняющих точный анализ доходности. В первую очередь это:

- невозможность абстрагирования от комиссионных выплат;
- дополнительные затраты заемщика по страхованию.

¹ Тубольцев М.Ф., В.И. Болтенков. Реструктуризация выплат по ипотечному кредиту // Научные ведомости Белгородского государственного университета. №7 (62). Выпуск 10/1. 2009. С. 31-37.



Первое обстоятельство существенно в ипотечных операциях и операциях потребительского кредитования физических лиц, поскольку здесь комиссионные выплаты соизмеримы с размером кредита. Для корпоративных долгосрочных кредитов эти суммы несопоставимы, и комиссионными выплатами можно пренебречь при анализе доходности. Второе обстоятельство на корпоративном уровне проявляется (возможно) только в лизинговых операциях. Поэтому, при рассмотрении общего случая, от этого обстоятельства также абстрагируемся. Последнее упрощение связано с тем, что рассматривается только одна кредитная операция, требующая реструктуризации.

Следует отметить, что если применяются не номинальные процентные ставки, а постоянные сложные процентные ставки, то доходность является инвариантом структурных преобразований². Это означает, что при использовании постоянных сложных процентных ставок никакие изменения в графике платежей по кредиту (в том числе и изменение размера платежей) не могут изменить доходность операции долгосрочного кредитования. При этом комиссионные и страховые выплаты должны либо отсутствовать, либо быть пренебрежимо малыми по сравнению с выплатами по кредиту.

В тех случаях, когда применяются схемы погашения долгосрочного кредита с номинальными процентными ставками, реструктуризация может существенно изменить доходность. Эта доходность всегда больше номинальной процентной ставки и максимально может достигать уровня эффективного процента.

Прежде всего, разберемся, почему использование номинальных ставок может влиять на доходность кредитной операции, в то время как при использовании сложной процентной ставки доходность постоянна. Следующий пример показывает, что доходность кредитной операции при расчетах по номинальной ставке существенным образом зависит от срока кредитования. Пусть кредит в размере 1000000 рублей выдан по номинальной ставке 12% годовых. Требуется определить доходность кредитования при сроках кредита 3, 6, 9 и 12 месяцев, используя известную формулу для эффективного процента:

$$r_{ef} = \left(\frac{S}{K} \right)^{\frac{1}{T}} - 1. \quad (1)$$

Получаем, что эффективный процент равен 12,8997%, 12,5465%, 12,2725% и 12%, соответственно. Очевидно, что доходность уменьшается с увеличением срока кредита, что является прямым следствием увеличения срока докапитализации процентов. При сроке кредита 3 месяца капитализация процентов происходит 4 раза в году, а при сроке кредита 12 месяцев – только один раз. Чем чаще капитализация, тем выше доходность. Следует обратить также внимание на то, что если капитализация осуществляется один раз в конце года, то различия между номинальной процентной ставкой и сложной процентной ставкой отсутствуют. При сложной процентной ставке капитализация происходит непрерывно, но рост процентов на малых интервалах времени (менее года) происходит медленнее, чем при использовании номинальных (т.е. простых) ставок. Использование номинальных ставок дает кредиторам дополнительную (по сравнению с использованием в расчетах сложных процентных ставок) маржу, если капитализация осуществляется чаще, чем один раз в год. Этим, вероятно, можно объяснить продолжающееся использование номинальных ставок банками, несмотря на недостатки последних.

В рассмотренном примере кредит был краткосрочным, но и в долгосрочном кредитовании принято использовать номинальные ставки для определения величины выплат по кредиту. Сами платежи, при этом, образуют (m, p) – ренту, где m – число капитализаций в году, а p – число выплат. При долгосрочном кредитовании

² Тубольцев М.Ф. Использование инвариантов при моделировании финансовых потоков // Научные ведомости Белгородского государственного университета. №7 (62). Выпуск 10/1. 2009. С. 146-152.

обычно реализуется частный случай общей (m, p) – ренты, когда $m=p$ (заемщику нет смысла платить по кредиту чаще, чем начисляются проценты).

Наиболее часто на практике используются две схемы погашения долгосрочного кредита: выплата долга равными частями и выплата долга равными срочными платежами (в случае ипотечного кредитования эти схемы называются дифференциальной и аннуитетной). В первом случае общие выплаты (часть долга + проценты по остатку долга) уменьшаются, во втором случае они постоянны (увеличивающиеся выплаты по основному долгу уравниваются уменьшением процентных выплат). Рассмотрим вторую схему погашения долгосрочного кредита, поскольку она является для кредитора более выгодной.

Пусть срочные уплаты постоянны (в размере R/p) в течение n лет, выплачиваются p раз в году и рассчитываются по номинальной ставке j с капитализацией процентов m раз в году. Тогда, используя формулы для современной и наращенной сумм (m, p) – ренты³, получаем значение доходности:

$$r_{ef} = \left(\frac{Rs_{n,m,j}^{(p)}}{Ra_{n,m,j}^{(p)}} \right)^{\frac{1}{n}} - 1 = \left(\frac{\left(1 + \frac{j}{m}\right)^{mn} - 1}{1 - \left(1 + \frac{j}{m}\right)^{-mn}} \right)^{\frac{1}{n}} - 1 = \left(1 + \frac{j}{m}\right)^m - 1. \quad (2)$$

Таким образом, если выплаты по долгосрочному кредиту осуществляются целое число лет, то доходность равна эффективному проценту и не зависит от размера кредита и срока, на который он выдан. Например, если номинальная ставка по кредиту размером 100 млн. рублей равна 12%, а капитализация осуществляется ежемесячно, то при сроке кредита $n=3$ года выплаты составят 3,32 млн. рублей, при $n=4$ года выплаты составят 2,63 млн. рублей, при $n=5$ лет выплаты составят 2,22 млн. рублей, а при $n=10$ лет выплаты составят 1,43 млн. рублей. Но во всех случаях доходность будет равна 12,664105%, что с учетом погрешности вычислений соответствует значению, которое дает формула (2). Формула (2) показывает и расчеты подтверждают это, что доходность полностью определяется двумя параметрами: номинальной процентной ставкой и периодом капитализации. Если погашается кредит размером 35 млн. рублей за 5 лет с тем же периодом капитализации и номинальной ставкой, то выплаты составят 0,78 млн. рублей, а доходность не изменится.

Если выплаты осуществляются не целое число лет, а целое число периодов капитализации (например, при капитализации раз в полгода кредит выплачивается три с половиной года, но не три с четвертью), то доходность также вычисляется по формуле (2). Таким образом, если при реструктуризации сохраняется схема погашения и не меняется период капитализации, то доходность сохраняется и равна эффективному проценту. Сложнее доказывается, но все же верно и в схеме выплаты долгосрочного кредита равными частями. Если, например, кредит в размере 20 млн. рублей выплачивается равными частями ежемесячно в течение 6 лет по номинальной ставке 12%, то выплаты меняются от начального значения 0,48 млн. рублей до конечного значения 0,28 млн. рублей в арифметической прогрессии, а доходность, как и в схеме равных срочных уплат, равняется 12,664105%.

Если же реструктуризация не сводится только к пролонгации выплат, а меняет также период капитализации и другие параметры, то меняется и доходность кредитования. При этом при увеличении периода капитализации доходность падает. Поскольку капитализация реже, чем один раз в год, не производится, то упасть доходность ниже номинальной ставки не может.

³ Тубольцев М.Ф., Болтенков В.И. Введение в финансовую математику: Учеб. пособие. Белгород: Изд-во БелГУ, 2005. 108 с.



Есть основания считать, что от применения номинальных ставок следовало бы отказаться в пользу сложных процентных ставок. Во-первых, реструктуризация схем выплат с номинальными ставками не меняет доходность только в специальных случаях. Во-вторых, номинальные ставки, будучи по своей природе простыми процентными ставками, не дают множителю наращивания или дисконтирования быть (в терминах статистики) цепными индексами, что приводит к ряду проблем теоретического и практического плана. В-третьих, применение номинальных ставок затрудняет работу с хронологическими датами, что особенно важно в условиях повсеместного применения вычислительной техники. Если установить, что капитализация осуществляется 4 раза в году, то компьютер «не поймет», что это квартал, имеющий переменное число дней, а не 0,25 года. Поэтому вся теория (m,p) -рент фактически представляет собой достаточно грубое приближение к реальности, уместное только в докомпьютерную эпоху.

Желательно было бы сохранить возможность применения рент, поскольку это единственный вид финансовых потоков, допускающих аналитическое представление для современных и наращенных сумм. Сделать это можно, изменив базовый период и перейдя к непрерывной капитализации процентов, т.е. к сложным процентным ставкам.

Изменение базового периода означает приведение всех параметров в компьютерных расчетах к суткам, поскольку сутки – это наибольший период времени, которому кратны все остальные периоды. Возможно, что для людей такое представление процентных ставок и множителей дисконтирования было бы непривычным и малоинформативным (возможно только на первых порах), но тогда ничто не мешает приводить показатели на презентациях для привычного базового периода.

Применение сложных (а не номинальных) процентных ставок снимает практически все проблемы реструктуризации, поскольку при неизменной сложной процентной ставке доходность является инвариантом структурных преобразований.

Теория рент, при этом, значительно упрощается, поскольку m -ренты и (m,p) -ренты становятся ненужными из-за непрерывной капитализации. Обычные годовые ренты и p -ренты заменяются T -рентами, где T – период ренты в днях.

Пусть D – размер рентных выплат, t_0 – момент начала ренты (момент первого рентного платежа), T – период ренты в днях, r – сложная процентная ставка, отнесенная к базовому периоду одни сутки, $R = 1+r$ – базовый множитель наращивания, n – количество элементов ренты. Тогда легко получается формула для наращенной суммы ренты в любой момент времени $t > t_0 + nT$:

$$S(t) = DR^{t-t_0} \frac{1 - R^{-nT}}{1 - R^{-T}}. \quad (3)$$

Аналогично для современной суммы ренты можно получить формулу для $t < t_0$:

$$A(t) = DV^{t_0-t} \frac{1 - V^{nT}}{1 - V^T}. \quad (4)$$

В формуле (4) V есть базовый множитель дисконтирования. Сохраняются все обычные соотношения между параметрами ренты, например:

$$S = AR^{(n-1)T}, \quad (5),$$

где S – наращенная сумма ренты на момент последнего платежа, а A – на момент первого платежа, $(n-1)T$ – длительность ренты.

Можно сделать вывод, что в компьютерных расчетах Т-ренды удобнее обычных рент еще и потому, что период Т не обязательно считать постоянным, поэтому их можно было бы назвать обобщенными рентами. Все современные системы программирования поддерживают тип хронологических дат. Поэтому легко можно запрограммировать прибавление к заданной дате любого количества месяцев. Несмотря на то, что период такой ренты не постоянен, обобщенная рента более адекватна реальным потокам платежей. Например, заработная плата выплачивается сотрудникам по фиксированным датам каждого месяца, а не через равные промежутки времени. В этом случае нужно использовать общие формулы для современных и наращенных сумм финансовых потоков, а не формулы (3) и (4), в формуле (5) также нужно указать точное число дней.

RE- INFLUENCE OF RE-STRUCTURING ON PROFITABLENESS OF LONG-TERM CREDIT OPERATION

M. F. TUBOLTSEV

V. I. BOLTENKOV

Belgorod State University

e-mail: tuboltsev@bsu.edu.ru

e-mail: boltenkov@bsu.edu.ru

Questions of change of profitableness of long-term credit operation are considered at its re-structuring. It is shown, that only at use of the constant difficult interest rate re-structuring does not change profitableness. Use of nominal interest rates and a principle of equivalence of payments at re-structuring lead to essential change of profitableness of crediting. Profitableness change at use of nominal rates is influenced practically by all parameters of re-structuring that does inconvenient the full control over profitableness. Under certain conditions profitableness of operation of long-term crediting can reach the effective percent calculated under the nominal rate.

Key words: the long-term credit, financial operation, financial systems, a technique of aggregation of financial operations, re-structuring of financial operations, a principle of equivalent change of payments, a principle of equivalent change of financial systems, re-structuring of financial systems.

ИСПОЛЬЗОВАНИЕ КОНЦЕССИОННЫХ СОГЛАШЕНИЙ В УПРАВЛЕНИИ ИНФРАСТРУКТУРНЫМИ ОБЪЕКТАМИ ЭКОНОМИКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Д. И. ИСКРЕНКО

Филиал ОАО
«Национальный» банк
«Траст», г. Вологоград

e-mail:
Dmitry.iskrenko@raiffeisen.ru

Обозначена задача привлечения частного капитала и управленческого опыта в инфраструктурные отрасли экономики РФ в связи с недостатком инвестиционных ресурсов государства и необходимостью поддержания деятельности стратегически важных объектов, чей переход в частную собственность невозможен или нежелателен. Выделены основные формы государственно-частного партнерства. Показано, что концессия является наиболее эффективной из них и позволяет достичь долгосрочные социально-экономические цели концедента и концессионера. Определены приоритетные инвестиционные проекты, которые могут быть эффективно реализованы на основе концессий при строительстве или реконструкции инфраструктурных объектов.

Ключевые слова: инфраструктурные отрасли экономики, государственно-частное партнерство, распределение рисков, повышение эффективности управления.

Одной из функций общественного воспроизводства, возложенных на государство, является обеспечение оптимальной деятельности предприятий и отраслей хозяйства, являющихся малорентабельными, традиционно не представляющими интереса для частного капитала, но при этом определяющих общие условия воспроизводства. Фактически государство берет на себя основную долю затрат по стимулированию важнейших интенсивных факторов развития производства и накопления капитала, по проведению НИОКР. В результате возникают отрасли общественной деятельности, освоением которых частный капитал не занимается, хотя и заинтересован в их развитии.

Экономическая инфраструктура обеспечивает создание общих условий производства, непосредственно, технологически связывающая производство продукта и его производительное потребление. Недостаточное развитие отраслей инфраструктуры отрицательно сказывается на всей экономической системе, сдерживает возможности интенсификации производства, рост его эффективности.

К отраслям инфраструктуры относят: транспорт; связь; энергетическое хозяйство; строительство дорог, каналов, портов, мостов, аэродромов, складов; водоснабжение; канализацию; системы наружного освещения городов и поселков; благоустройство населенных пунктов; индустрию досуга и туризма; образование; здравоохранение; оборону, охрану границ, системы безопасности и правосудия и т.д.¹

В современных российских условиях развитие инфраструктурных отраслей приобретает стратегический характер, поскольку обеспечивает целостность и эффективность всей экономической системы. Однако недостаток бюджетных средств не позволяет осуществлять управление объектами инфраструктуры должным образом. В связи с этим, у государства возникает необходимость поиска дополнительных форм сотрудничества с частным бизнесом в виде новых инструментов привлечения капитала в инфраструктурные отрасли.

Следует учитывать, что для объектов инфраструктуры, имеющих стратегическое значение с экономической, социальной или военной точек зрения, приватиза-

¹ Рогинко С. Концессия как форма привлечения инвестиций в развивающиеся страны // Концессия. Об эффективном договоре между государством и бизнесом. Сб. научных материалов. М.: Аналитический центр «Эксперт», 2002. С. 79.

ция которых в силу этого невозможна, использование государственно-частного партнерства (ГЧП) в эксплуатации и управлении безусловно, предпочтительно.

ГЧП представляет собой экономическое, организационное и институциональное взаимодействие государства и бизнеса, основной целью которого является мобилизация финансовых ресурсов для реализации общественно значимых инвестиционных проектов в различных секторах экономики.

Современная практика свидетельствует о многообразии форм участия частного бизнеса в управлении государственной собственностью:

- интеграция государственного и частного капитала в рамках отдельной совместной компании. Формирование смешанных предприятий осуществляется путем покупки государством акций частных фирм или путем продажи части принадлежащих ему акций частным инвесторам;

- передача объектов госсобственности в трастовое (доверительное) управление частным предпринимателям. В соответствии с договором, учредитель траста передает управляющему на определенный срок часть своего имущества, которое используется последним в интересах учредителя, либо указанного им третьего лица без перехода права собственности;

- передача государственного имущества в аренду. По договору аренды (имущественного найма) арендодатель берет на себя обязательство предоставить арендатору имущество за плату во временное владение и (или) пользование;

- использование контрактов функционирования и управления. Первые представляют собой соглашения между администрацией предприятия и государственным органом по достижению согласованных параметров развития. Вторые представляют собой соглашения между правительством и частным лицом по осуществлению управления государственным предприятием за вознаграждение;

- применение договоров подряда, при которых заказы на выполнение услуг размещаются среди частных компаний на конкурсной основе;

- государственные контракты на выполнение работ по финансированию, проектированию и строительству, на оказание общественных услуг, поставку продукции для государственных нужд, на оказание технической помощи. Их отличительной чертой является административная форма, а также то, что право собственности на предмет контрактных отношений не передается предпринимателю. Вся деятельность в рамках контракта осуществляется, как правило, на средства государства;

- концессионные соглашения, предоставляющие инвесторам исключительное право осуществления хозяйственной деятельности на объектах госсобственности.

С точки зрения эффективного управления государственной собственностью, особый интерес представляют инфраструктурные концессии как механизмы передачи частным инвесторам объектов инфраструктуры, не предназначенных для приватизации (изъятых из гражданского оборота) с проведением их реконструкции, развития с целью получения прибыли. В современных условиях использование механизма концессии позволяет преодолеть противоречия по поводу передачи государством (центральными органами власти, муниципалитетами и региональными властями) в управление частному бизнесу объектов экономической и социальной инфраструктуры на основе распределения рисков.

В экономической литературе существуют два основных значения концессии:

1) уступка права пользования государственной собственностью²;

2) договор о передаче в эксплуатацию на определенных условиях и определенный срок природных недр, предприятий и других хозяйственных объектов, находящихся в собственности государства или муниципалитетов для их постройки, реконструкции, управления, модернизации³.

² Большая советская энциклопедия. Т. 34 / Гл. ред. О.Ю. Шмидт. М.: Гос. Ин-т «Сов. Энциклопедия», 1937. С. 182.

³ Большой экономический словарь / Под ред. А.Н. Азриляна. М.: Институт новой экономики, 1999. С. 177.



Интегральной целью концессионного механизма является повышение эффективности удовлетворения общественных потребностей в инфраструктурном обеспечении. Дифференциальные цели зависят от субъектов концессионных отношений и включают экономические и социальные (см. табл. 1).

Таблица 1

Основные цели субъектов концессионного механизма

Цели Субъект	Экономические	Социальные
Концедент	Привлечение инвестиций для поддержания существующих материально-технических объектов, создания новых (нерентабельных при обычной системе налогообложения)	Создание новых рабочих мест и снижение безработицы
	Реорганизация и развитие предприятий общественного сектора экономики (общественных служб), реализующих услуги населению	Сокращение дифференциации в доходах
	Создание импортозамещающих производств на основе новых технологий производства и управления	Повышение уровня образования и квалификации персонала
	Улучшение счета операций с капиталом и финансовыми инструментами платежного баланса	Вовлечение в хозяйственный оборот невостребованных ресурсов
	Стимулирование НИОКР на концессионных объектах	Удовлетворение общественных потребностей за счет оптимального сочетания качества и цены товаров и услуг
	Обновление основных фондов предприятий	Снижение социальной напряженности
Концессионер	Повышение доходности на вложенный капитал за счет ограниченной конкуренции, более высокой рентабельности инвестиционных проектов, действующих на концессионной основе в льготных налоговых условиях	Выплата налогов в местный и федеральный бюджеты
	Расширение рынков сбыта продукции	Создание хорошо оплачиваемых рабочих мест, безопасной рабочей среды за счет использования современных технологий
	Перемещение производств в страны/регионы, характеризующиеся более низкими издержками (дешевые источники сырья, квалифицированные трудовые ресурсы), а также налоговой нагрузкой (за счет налоговых или таможенных льгот на время окупаемости проекта)	Предоставление потребителям качественных товаров и услуг. Обеспечение комфортных условий жизни
	Диверсификация бизнеса посредством включения в производственно-технологическую цепочку новых производств	Охрана здоровья и развитие персонала Формирование лояльности трудового коллектива
	Создание замкнутого цикла производства (при вертикальной интеграции в рамках холдинговых структур)	Экологическая безопасность

Концессия является формой привлечения в инфраструктурные отрасли предпринимательского капитала, при которой происходит уступка концедентом (государственным или муниципальным органом) концессионеру (частному инвестору) права владения и пользования концессионным объектом на определенных договором условиях и период времени в обмен на концессионные платежи. Субъектами концессии являются концедент в лице органов государственной власти и концессионер в лице инвесторов (иностраных и внутренних).

В качестве объектов важно выделить две группы:

1) материально-технические комплексы, производственные предприятия, транспортную инфраструктуру (автомобильные дороги, мосты, железные дороги, аэропорты, морские порты, нефте- и газопроводы, линии электропередач), объекты коммунального хозяйства;

2) осуществление определенной деятельности (тепло-, водо-, электро-, газо-снабжение, обслуживание объектов ЖКХ, управление предприятиями).

Экономическая сущность концессии заключается во взаимном возникновении обязательств сторон соглашения, создающем возможность осуществления инвестиций в объекты государственной собственности на совершенно ином уровне, чем в обычных экономико-правовых условиях.

Основной задачей инфраструктурных концессий является привлечение частного капитала и управленческого опыта, накопленного в негосударственном секторе экономики, к созданию и управлению экономической инфраструктурой государства⁴.

Главным отличием концессий от других форм участия частного капитала в управлении объектами государственной собственности в категориях права является то, что они функционируют в рамках не только гражданско-правового поля, но и публично-правовых норм. Часто для концессий создается специальное законодательство⁵.

Использование концессий имеет место как в промышленно наиболее развитых странах, так и в государствах с переходной экономикой, а также в развивающихся странах. Примерами такого использования являются: строительство аэропортов в Египте; строительство и эксплуатация магистральных железных дорог, портов и портовых сооружений, телекоммуникационной инфраструктуры в Мексике; управление водоснабжением и железными дорогами в Аргентине; строительство электростанций в Индии, Турции; строительство платных дорог в Канаде, Великобритании.

Схемы финансирования инфраструктурных концессий разнообразны. Наиболее существенно различаются две из них:

1) все издержки концессионера покрываются за счет взимания платы за услуги с населения;

2) все издержки концессионера полностью покрываются за счет государственных субсидий. Однако на практике оба метода финансирования, как правило, сочетаются в той или иной пропорции.

Возмещение затрат концессионера на содержание концессионного объекта и доходы концессионера осуществляются за счет третьих лиц – заказчиков, потребителей, услугополучателей и других клиентов публичной службы. Среди них могут быть и государственные, и муниципальные потребители (предприятия, учреждения, службы), выступающие в роли клиентуры концессионера (но не в роли концедента). В зависимости от условий соглашения, концессионер может быть либо розничным поставщиком услуг конечным потребителям, либо поставлять их оптом специализированной государственной организации, которая затем предоставляет их потребителям в розницу. В этом последнем случае оплату услуг концессионера принимает на себя государственная организация, взимающая абонентскую (или иную) плату с конечных получателей услуг.

Характеристики концессионных отношений представлены в табл. 2.

⁴ Некрасов Д. Международный опыт законодательного регулирования концессионных отношений // Концессия. Об эффективном договоре между государством и бизнесом. Сб. научных материалов. М.: Аналитический центр «Эксперт», 2002. С.71.

⁵ Варнавский В.Г. Государственно-частное партнерство в Европе // Современная Европа. 2005. № 2. С. 65.



Таблица 2

Основные характеристики концессионных отношений

Характеристика	Содержание
Субъекты	Концедент – государство (как агент общества), представленное органами власти различного уровня; концессионер – иностранное или отечественное физическое или юридическое лицо
Объекты	Объекты государственной или муниципальной собственности; виды деятельности
Принципы реализации	Платность (концессионные платежи); срочность (долгосрочный характер); возвратность объекта концессии концессионеру по истечении срока договора; целевой характер использования (взаимные обязательства сторон)
Правовые нормы	Концессионный договор (соглашение); сервисное соглашение; риск-сервисное соглашение; учредительный договор; лицензия; лицензия с гражданско-правовыми оговорками

Принятие в 2005 г. Федерального закона «О концессионных соглашениях» с поправками от 30.06.2008 г⁶. и учреждение правительством РФ Инвестиционного фонда, ориентированного на реализацию проектов на основе ГЧП, способствовало формированию необходимой базы, создающей новые институциональные и организационные возможности повышения эффективности функционирования инфраструктурных отраслей. В условиях приостановки процессов приватизации и увеличения количества объектов инфраструктуры, переход которых в частную собственность невозможен, концессии выступают альтернативой традиционным механизмам трансформации отношений собственности и финансирования.

Концессионный механизм применяется на постсоветском пространстве в странах, законодательство которых предусматривает возможность заключения концессионных соглашений (см. табл. 3).

Таблица 3

Концессионное законодательство в странах СНГ

Государство	Год принятия закона	Название основного и сопутствующих нормативно-правовых актов
Казахстан	1991	«О концессия в Республике Казахстан»
Кыргызстан	1992	«О концессиях и иностранных концессионных предприятиях» «Об иностранных инвестициях», «О собственности», «Об аренде и арендных отношениях», «Об общих началах разгосударствления, приватизации и предпринимательства»
Украина	1999	«О концессиях»
Узбекистан	1998	«Об иностранных инвестициях»
Молдова	1995	«О концессиях»
Беларусь	2001	«Инвестиционный кодекс Республики Беларусь»
Таджикистан	1997	«О концессиях»
Россия	2005	«О концессионных соглашениях»

Стратегическими приоритетами развития инфраструктурных отраслей экономики РФ на принципах концессии являются: жилищно-коммунальное хозяйство, где износ основных фондов достиг критических величин при отсутствии необходимых

⁶ Федеральный закон «О концессионных соглашениях». Гарант: Информационная программа официальных документов. 2008: <http://garant.ru>.

объемов финансирования; модернизация гидроэлектростанций; строительство нефтепроводов; создание необходимой инфраструктуры при освоении и обустройстве территорий; формирование эффективной инфраструктуры национальной инновационной системы, способствующей развитию высокотехнологичных отраслей, научно-исследовательских и образовательных структур, обеспечивающих рост объемов товаров и услуг, поставляемых на внешний и внутренние рынки.

Инвестиционными проектами в регионах России, которые могут быть реализованы на концессионной основе в ближайшее время, являются:

- в Волгоградской области – строительство моста через р. Волга, объездной автомобильной дороги вокруг областного центра, санатория Эльтон на озере, постройка Волжского бизнес-центра, а также завода биоэтанола;

- освоение Нижнего Приангарья. Инициаторами проекта выступила администрация Красноярского Края совместно с крупными российскими компаниями «Русский алюминий» и «Базовый элемент»;

- развитие инфраструктуры нефтехимического комплекса в г. Нижнекамске нефтяной компанией «Татнефть»;

- модернизация национальной транспортной инфраструктуры на основе долгосрочной стратегии развития транспорта Российской Федерации⁷.

Несмотря на наличие нормативной базы и готовности к сотрудничеству государственных органов власти и частных предпринимателей инвестиционные проекты на основе концессии все еще остаются предметом теоретических изысканий, а не практической реализации. Тем не менее, особая актуальность взаимодействия государства и частного сектора в настоящее время вызвана также кризисными явлениями в мировой экономике и, особенно, в финансовом секторе, что предопределяет новые подходы к управлению рисками финансирования инвестиционных проектов, принципы допуска частных инвесторов, возможности оперативного реагирования законодательных властей на изменяющиеся условия внешней бизнес-среды, позволяющие привлекать долгосрочных инвесторов со стабильными источниками фондирования и современными управленческими технологиями.

Среди мер, способствующих выстраиванию долгосрочных отношений между концедентом и концессионером и требующих особого внимания в отечественной практике, следует выделить:

- предоставление правительственных гарантий концессионерам;

- распределение рисков и способы урегулирования конфликтных процедур между субъектами концессионных отношений;

- оценка инвестиционных проектов с учетом реализации социальных целей;

- обеспечение мониторинга и контроля устойчивости выполнения договорных обязательств сторон;

- безопасность использования государственных инвестиций, предполагающая наличие транспарентной финансовой отчетности.

Таким образом, инфраструктурные концессии являются механизмами передачи объектов инфраструктуры концедентом концессионеру на срок, закрепленный в договоре. Существуют различные варианты такой передачи, каждый из которых определяет специфику перехода права пользования и управления определенными объектами. На практике возмещение издержек концессионера осуществляется как за счет потребителей, так и за счет государственных субсидий. Преимуществом для концедента является возможность переноса расходов и рисков на концессионера. Концессионер получает возможность работы с гарантированной, хотя и невысокой прибылью, используя таможенные и налоговые льготы.

⁷ Федеральная целевая программа «Модернизация транспортной системы России (2002 – 2010 годы)». Официальный сайт Министерства транспорта Российской Федерации. <http://www.mintrans.ru/>



OPERATION OF INFRASTRUCTURE UNDER CONSESSION AGREEMENTS IN THE ECONOMY OF THE RUSSIAN FEDERATION

D. I. ISKRENKO

**Volgograd Branch
OAO «National Bank "Trast"»**

**e-mail:
Dmitry.iskrenko@raiffeisen.ru**

The objective of private capital and managerial experience attraction in infrastructure of The Russian Federation is highlighted. This issue is very acute since the federal budget does not have enough financial resources to operate properly infrastructural objects which are strategically important and are unwanted or may not be privatized. The principal forms of the public-private partnership are excluded. Concession being a form of it is marked as the most effective one, as it enables concessionaire and concedent to attain long-term social and economic objectives. The prior investment projects which can be implemented under concession agreements are defined in construction or rehabilitation of infrastructure.

Key words: infrastructure, public-private partnership, risk allocation, operating efficiency increase.

МОДЕЛИРОВАНИЕ ИНВЕСТИЦИОННОГО ПРОЦЕССА В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

В. А. КАЛУГИН
И. Н. ТИТОВА

*Белгородский
государственный
университет*

e-mail: Kalugin@bsu.edu.ru

e-mail: Titova@bsu.edu.ru

Показано, что полезный эффект от инвестиций в условиях неопределенности на каждом шаге расчетного периода может быть адекватно описан только совокупностью как количественных (денежные потоки), так и качественных показателей, характеризующих прибыль от реализации проекта. Представлена структура инновационного процесса и иерархическая модель процесса принятия инвестиционных решений. Предложен способ моделирования в условиях неопределенности полезного эффекта от инвестиций и способ построения единой количественной оценки инвестиционного проекта на каждом шаге расчетного периода.

Ключевые слова: инвестиции, инвестиционный процесс, условия неопределенности и риска, полезный эффект от инвестиций, субъективная шкала качественных градаций, методология анализа иерархических структур, шкала отношений, матрица парных сравнений, главный собственный вектор матрицы.

В соответствии с Федеральным законом РФ «Об инвестиционной деятельности в Российской Федерации, осуществляемой в форме капитальных вложений» №39-ФЗ от 25 февраля 1999 года, под инвестициями будем понимать «... денежные средства, ценное, иное имущество, в том числе имущественные права, иные имеющие денежную оценку, вкладываемые в объекты предпринимательской и (или) иной деятельности в целях получения прибыли и (или) достижения иного полезного эффекта».

Деятельность, направленную на реализацию инвестиций, принято называть инвестиционной деятельностью, которая как любой вид деятельности включает цель, средства, результат и сам процесс — инвестиционный, представляющий совокупность последовательных действий субъекта инвестиционной деятельности для достижения определенного результата (прибыли и/или полезного эффекта).

Субъектами инвестиционной деятельности могут быть физические и юридические лица, в том числе иностранные, а также государства и международные организации. Основной субъект инвестиционной деятельности — инвестор, осуществляющий вложение собственных, заемных или привлеченных средств в форме инвестиций и обеспечивающий их целевое использование.

Инвестиционный процесс протекает в рамках конкретного инвестиционного проекта (ИП), который представляет собой «... обоснование экономической целесообразности, объема и сроков осуществления капитальных вложений, в том числе необходимая проектно–сметная документация, разработанная в соответствии с законодательством РФ и утвержденными в установленном порядке стандартами (нормами и правилами), а также описание практических действий по осуществлению инвестиций (бизнес-план)»¹.

При осуществлении инвестиционной деятельности инвесторы непрерывно сталкиваются с проблемой принятия решения о том, стоит ли расходование ресурсов — времени или денег — ожидаемых денежных доходов. Их беспокоит вопрос, каким образом распределять имеющиеся в его распоряжении ресурсы между множеством возможных вариантов инвестиций. Следует ли строить завод? Производить ли замену оборудования? Осваивать новые виды продукции?

¹ Методические рекомендации по оценке эффективности инвестиционных проектов, № ВК 477 от 21.06.99 г., утверждено Минэкономки РФ, Минфином РФ, Госкомитетом РФ по строительной, архитектурной и жилищной политике.



Все это можно рассматривать как *проблемы принятия инвестиционных решений*.

Повышение качества принимаемых инвестиционных решений было и остается важнейшей проблемой не только на уровне предприятий, но и рассматривается как панацея от многих трудностей на макроэкономическом уровне. От того, насколько дальновидны, рациональны, экономически обоснованы и эффективны инвестиционные решения, во многом зависит не только будущее конкретных предприятий и организаций, но и будущее экономики страны в целом, поскольку их функционирование влияет на поведение других субъектов экономического сообщества, социально-экономическую и природную среду.

В немалой степени успешному решению проблемы принятия инвестиционных решений должны способствовать используемые при этом методы оценки инвестиций. В настоящее время наиболее адекватными признаются динамические методы, основанные на концепции временной стоимости денег. Такие методы получили название дисконтированных методов оценки инвестиций.

В отечественной и зарубежной практике принятия инвестиционных решений получили распространение следующие, ставшие классическими, дисконтированные методы оценки эффективности инвестиционных проектов, основанные на расчете показателей:

- чистая текущая (приведенная) стоимость инвестиций (net present value – NPV, чистый дисконтированный доход – ЧДД, или интегральный эффект);
- индекс рентабельности инвестиций (profitability index – PI, индекс доходности дисконтированных инвестиций – ИДД);
- внутренняя норма доходности (рентабельности) инвестиций (internal rate of return – IRR, внутренняя норма доходности – ВНД);
- модифицированная внутренняя норма доходности (modified internal rate of return – MIRR).

Однако надежное использование этих показателей в качестве инструментов принятия инвестиционных решений в определяющей степени опирается на способность экономистов и финансовых аналитиков выносить обоснованные суждения относительно ожидаемой величины денежных потоков анализируемых ИП на каждом шаге расчетного периода.

Основу этих суждений составляют исходные предположения относительно внешних факторов, имеющих к инвестиционному процессу непосредственное отношение, поскольку напрямую воздействуют на показатели его эффективности. К внешним факторам относятся:

- предполагаемый в будущем уровень деловой активности;
- конъюнктура на рынке капитала, определяющая как условия получения займа и ставку кредитного процента, так и доходность доступных инвестору альтернативных вложений его капитала с гарантированными и рисковыми вложениями;
- конъюнктура на рынке сбыта, характеризующая будущее соотношение спроса и предложения, а, значит, и ожидаемый уровень цен и объемы продаж, и, в конечном счете, денежные поступления от инвестиции;
- конъюнктура на рынке труда и материальных ресурсов, обуславливающая как расходы инвестора на приобретение инвестиционного объекта, так и текущие расходы на его эксплуатацию;
- налоги и налоговая политика федеральных и муниципальных органов;
- достижения техники и технологии;
- политика и действия конкурентов и т. п.

Так как каждый из этих факторов характеризуется высокой степенью неопределенности, рассчитанные показатели в лучшем случае представляют лишь наброски будущих результатов деятельности по осуществлению ИП.

Причем иногда очень сильно переоценивается способность экономистов и финансовых аналитиков предвидеть изменения в экономике страны или в отдельной отрасли, на конкретном предприятии, организации или для отдельного продукта. На самом деле, как справедливо отмечается многими авторами, их прогнозы и планы постоянно демонстрируют свое несовершенство и неполноту.

Не следует также забывать, что принятие инвестиционных решений — это целенаправленная человеческая деятельность, а не объективный механистический процесс. В рамках этой деятельности значительное место занимают субъективные оценки и суждения. Именно этот факт оказался упущенным в традиционной теории инвестиций.

Следует заметить, что вообще любые суждения относительно будущих событий — в данном случае относительно будущих денежных потоков ИП, — выносимые в виде абсолютных значений, почти всегда ненадежны. Вместе с тем в теории инвестиций незримо присутствует образ экономически целесообразного принимающего решения человека, который ориентирован на максимизацию прибыли, обладает сверхинтеллектуальными способностями и совершенными познаниями в области финансов, способен безошибочно прогнозировать будущие события и выносить точные оценки денежных потоков ИП. Это обнадеживающая, но нереальная ситуация.

Повышению эффективности принимаемых инвестиционных решений, на наш взгляд, в значительной степени мог бы способствовать такой методологический инструментарий, который позволял бы работать с суждениями, выносимыми не в виде абсолютных оценок денежных потоков ИП, а в виде относительных, показывающих степень превосходства одного оцениваемого денежного потока над другим. Действительно, менеджер предприятия вполне определенно может судить о том, что денежный поток одного ИП на t -ом шаге расчетного периода превосходит денежный поток второго ИП. Однако сказать, каковы будут точно или с определенной вероятностью конкретные размеры денежных потоков ИП намного сложнее. Подробнее об оценке денежных потоков ИП на каждом шаге расчетного периода в девятибалльной шкале отношений можно ознакомиться по работам²³.

Следует заметить, что последствия принятия инвестиционных решений не всегда удается полностью выразить в виде денежных потоков. Рассмотрим в этой связи структуру инвестиционного процесса.

Структура инвестиционного процесса. Особенность построения модели инвестиционного процесса заключается в том, что она должна охватывать два подпроцесса — описание потока расходования денежных средств (инвестиционные затраты, оттоки) и описание полезного эффекта (отдача от инвестиций, притоки), который следует ожидать при осуществлении ИП.

Основная трудность при моделировании инвестиционного процесса заключается в описании полезного эффекта от инвестиций. Первый шаг в этом направлении заключается в разработке его структуры — определение расчетного периода, расчленение его на этапы (шаги), различающиеся своим содержанием, размерами денежных доходов и расходов, а также качественными показателями эффекта.

В том случае, когда полезный эффект можно измерить в деньгах, различают следующие условия моделирования инвестиционного процесса (условия принятия решений об инвестициях):

- определенность, если каждому ИП ставится в соответствие одна и только одна оценка денежного потока на каждом шаге расчетного периода;

² Калугин В.А. Критериально-экспертная оценка инвестиционных проектов // Проблемы теории и практики управления». 2006. №7. С. 84 – 92.

³ Калугин В. А. Многокритериальные методы принятия инвестиционных решений. Спб.: Химиздат, 2004. 211 с.



- вероятностная неопределенность (риск), если каждому ИП на каждом шаге расчетного периода ставится в соответствие множество оценок, каждая из которых имеет определенную вероятность осуществления;

- интервальная неопределенность (неопределенность), если каждому ИП на каждом шаге расчетного периода ставится в соответствие множество (или числовой интервал) оценок и при этом ничего неизвестно относительно вероятности осуществления каждой.

Однако полезный эффект не всегда удастся полностью выразить в деньгах, имеются важные для инвестора качественные показатели полезного эффекта — качественные аналоги прибыли. В общем случае полезный эффект может характеризоваться различными видами качественной «прибыли»: V_1, \dots, V_m . Причем виды «прибыли» могут быть выражены в различных единицах измерения, поэтому их нельзя объединить в одну общую меру прибыли. Например, V_1 — экономия человеко-часов, V_2 — архитектурное качество инвестиционного объекта и т.д.

Кроме того, ИП на некоторых шагах расчетного периода могут обладать так называемыми управленческими возможностями (управленческими опционами):

- увеличение масштабов проекта в случае, если он успешен;
- продажа проекта, если он потерпит неудачу;
- развитие смежных видов деятельности, пользуясь приобретенным на первом проекте опытом;
- разработка новых продуктов в русле начатого проекта;
- расширение рынков сбыта продукции;
- расширение или переоснащение производства;
- прекращение проекта и др.

Традиционно учет управленческих возможностей осуществляется путем их денежной оценки и корректировки показателя NPV, рассчитанного так: реальный NPV = традиционный NPV + стоимость управленческих опционов.

Таким образом, полезные последствия принятия инвестиционных решений (полезный эффект) в условиях неопределенности на каждом шаге расчетного периода могут быть адекватно описаны только совокупностью как количественных (денежных потоков), так и качественных показателей, характеризующих «прибыль» от реализации проекта (табл. 1).

Таблица 1

Описание полезного эффекта от реализации ИП на каждом шаге расчетного периода

№ шага расчетного периода	«прибыль-1»	«прибыль-2»	...	«прибыль-м»
1	b_{11}	b_{12}	...	b_{1m}
...
i	b_{i1}	b_{i2}	...	b_{im}
...
T	b_{T1}	b_{T2}	...	b_{Tm}

Проблема состоит в агрегировании на каждом i -ом шаге расчетного периода различных видов «прибыли» от реализации конкретного ИП в одну составную величину прибыли V_i , которая представляла бы оценку этого проекта на соответствующем шаге.

Одним из обычных приемов такого объединения является введение ряда весовых коэффициентов w_1, w_2, \dots, w_m и определение составной прибыли в виде: $V_i = w_1 b_{i1} + w_2 b_{i2} + \dots + w_m b_{im}$.

При этом, естественно, единицы измерения w_i должны быть таковы, чтобы все отдельные члены $w_1b_{11}, w_2b_{12}, \dots, w_mb_{1m}$ были выражены в соизмеримых единицах. На практике трудность состоит в том, чтобы найти подходящие весовые коэффициенты. Часто это делается либо при помощи объективного «рыночного» механизма, либо на субъективной основе — каждой рассматриваемой величине ставятся в соответствие определенные суммы.

Однако непосредственно объединить оценки по качественным и количественным критериям таким способом не представляется возможным. Обычной практикой в этом случае является преобразование качественных оценок в количественные, денежные оценки и дальнейшая их агрегация с использованием весовых коэффициентов. В том случае, если критерием является такой, как «экономия человеко-часов», это легко можно осуществить⁴. Однако если критерием является такой, как «архитектурное качество», возникают серьезные трудноразрешимые проблемы.

Другим, альтернативным может быть подход, предусматривающий преобразование количественных оценок в качественные с дальнейшей агрегацией последних в единую качественную оценку ИП с позиции каждого шага расчетного периода. Строго говоря, проблема оценки преимуществ того или иного подхода представляет самостоятельный научный интерес и в данной статье не рассматривается.

Предлагаемый подход. Нами предлагается метод моделирования инвестиционного процесса в условиях неопределенности, предусматривающий, во-первых, оценку полезных последствий от реализации ИП на каждом шаге расчетного периода в субъективной порядковой шкале качественных градаций, во-вторых, преобразование качественных оценок в количественные.

Модель процесса принятия инвестиционного решения будем представлять в виде следующей иерархии (рис .1).

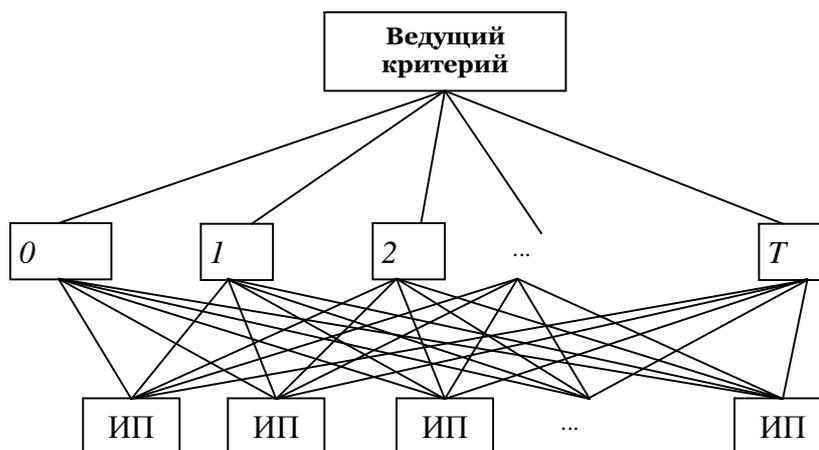


Рис. 1. Модель процесса принятия инвестиционных решений

Поскольку речь идет об условиях неопределенности, предполагается, что каждый ИП на каждом шаге расчетного периода описывается совокупностью количественных и качественных критериев.

Для того, чтобы получить количественную оценку ИП с позиции некоторого шага расчетного периода, предлагаем вначале дать соответствующую оценку ИП в

⁴ Фокс Д. Дж., Грегори К. Ричард. Конкурентные преимущества в денежном выражении / Пер. с англ. Изд. 3-е. М.: Альпина Бизнес Букс, 2008. 304 с.

субъективной порядковой шкале качественных градаций. Такая количественная оценка необходима для последующей иерархической композиции, результатом которой является интегральная количественная оценка ИП с позиции ведущего критерия, позволяющая, например, ранжировать ИП.

Первый этап. Полезный эффект на каждом шаге расчетного периода предлагается оценивать с позиции субъективной порядковой шкалы качественных градаций (см. ниже: *Пример*).

Процесс получения оценки ИП с позиции этой шкалы протекает в три шага.

Первый шаг. Множество денежных потоков, имеющих место в условиях неопределенности на каждом шаге расчетного периода, преобразуется в единый денежный поток. Для этого можно, например, воспользоваться критерием Гурвица:

$$CF_{ож} = \lambda \cdot CF_{max} + (1-\lambda) \cdot CF_{min} \quad (1)$$

где

$CF_{ож}$ — ожидаемый денежный поток (*Cash Flow*);

CF_{max}, CF_{min} — соответственно, максимальное и минимальное значение, взятые из множества денежных потоков;

λ — некоторое число, заключенное в интервале от 0 до 1 и определяемое субъективно лицом, принимающим решение (ЛПР).

Полученная оценка единого денежного потока преобразуется в качественную оценку путем ее соотнесения с качественными градациями. Например, ЛПР считает, что полученное значение $CF_{ож}$ представляет собой «низкий уровень» денежного потока.

Второй шаг. Качественные виды «прибыли» от реализации конкретного ИП на рассматриваемом шаге расчетного периода непосредственно оцениваются в шкале качественных градаций.

Третий шаг. ЛПР объединяет совокупность полученных оценок в единую качественную оценку ИП с позиции конкретного шага расчетного периода.

Следует заметить, что авторы отдают себе отчет в том, что описанные шаги нуждаются в более детальном изложении, однако, на наш взгляд, это выходит за рамки настоящей статьи.

Второй этап. Преобразование качественных оценок ИП в количественные. Для этого воспользуемся методологией анализа иерархических структур (МАИ)⁵⁶.

Пусть субъективная порядковая шкала качественных градаций имеет следующий вид (рис. 2).

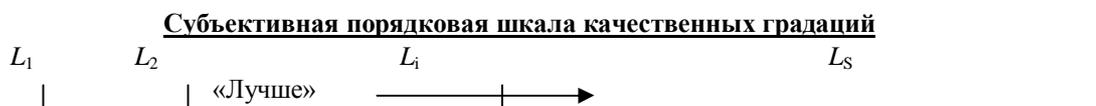


Рис. 2. Субъективная порядковая шкала

Предположим, что ИП получили в соответствии с этой шкалой следующие оценки (табл. 2.)

⁵ Саати Т. Принятие решений. Метод анализа иерархий; пер. с англ. М.: Радио и связь, 1989. 316 с.

⁶ Андрейчиков А.В., Андрейчикова О.Н. Анализ, синтез, планирование решений в экономике. М.: Финансы и статистика, 2000. 368 с.

Таблица 2

Оценки ИП по качественным градациям субъективной шкалы

ИП	ИП-1	...	ИП- <i>i</i>	...	ИП- <i>K</i>
Значение	L_{s1}	...	L_{si}	...	L_{sk}

Проведем парное сравнение качественных градаций по девятибалльной шкале отношений с позиции того, насколько одна качественная градация превосходит другую. Результат поместим в обратносимметричную матрицу парных сравнений. Пусть она получила следующий вид (табл. 3).

Таблица 3

Матрица парных сравнений качественных градаций

L	L_1	...	L_i	...	L_s
L_1	1	...	L_{1i}	...	L_{1s}
...
L_i	l_{i1}	...	1	...	l_{is}
...
L_s	l_{s1}	...	l_{si}	...	1

Найдем главный собственный вектор этой матрицы. Его компоненты, в случае положительного результата проверки матрицы на согласованность, определяют числовые оценки качественным градациям.

Пусть главный собственный вектор определен и имеет вид:

$$W = (w_1, \dots, w_i, \dots, w_s) \quad (2)$$

Тогда соответствие качественных градаций числовым оценкам можно представить следующей таблицей (табл. 4).

Таблица 4

Соответствие качественных градаций числовым значениям

Градация	L_1	...	L_i	...	L_s
W	w_1	...	w_i	...	w_s

При этом ИП автоматически получаю числовые оценки по качественному критерию (табл. 5).

Таблица 5

Числовые оценки ИП по субъективной шкале

ИП	ИП-1	...	ИП- <i>i</i>	...	ИП- <i>K</i>
Качественное значение	L_{s1}	...	L_{si}	...	L_{sk}
Числовое значение	w_{s1}	...	w_{si}	...	w_{sk}

Пример. Предположим, что для некоторого шага расчетного периода вполне достаточно иметь субъективную шкалу, имеющую качественные градации, выражающие уровень поступлений от инвестиций на этом шаге: «низкий» — L ; «средний» — M ; «высокий» — H ; промежуточные значения — LM («ниже среднего») и HM («выше среднего»).



Предположим, что оцениваются пять ИП. Пусть рассматриваемые ИП получили, с точки зрения этой шкалы, следующие качественные оценки:

ИП-1 \rightarrow М; ИП-2 \rightarrow LM; ИП-3 \rightarrow НМ; ИП-4 \rightarrow М; ИП-5 \rightarrow НМ.

Отообразим субъективную шкалу с качественными градациями на некоторую числовую шкалу. Предположим, что матрица в результате попарного сопоставления качественных градаций получила следующий вид (табл. 6).

Таблица 6

Матрица парных сравнений качественных градаций

A ₅	L	LM	M	НМ	H
L	1	1/3	1/4	1/8	1/9
LM	3	1	1/2	1/3	1/6
M	4	2	1	1/2	1/3
НМ	8	3	2	1	1/2
H	9	6	3	2	1

Установлено, что максимальное собственное значение анализируемой матрицы парных сравнений равно 5,22, а коэффициенты ИС = 0,055 и ОС = 0,049. Следовательно, матрица согласована и суждения экспертов логичны.

Найдем собственный вектор, отвечающий максимальному собственному значению этой матрицы. Имеем:

$$\mathbf{W} = (w_1, w_2, w_3, w_4, w_5) = (0,036, 0,100, 0,156, 0,289, 0,419) \quad (3)$$

Компоненты вектора \mathbf{W} определяют следующие числовые значения для качественных градаций (табл. 7).

Таблица 7

Соответствие качественных градаций числовым значениям

Градация	L	LM	M	НМ	H
w _i	0,036	0,100	0,156	0,289	0,419

При этом одновременно ИП получают следующие числовые оценки: ИП-1 \rightarrow М \rightarrow 0,156; ИП-2 \rightarrow LM \rightarrow 0,100; ИП-3 \rightarrow НМ \rightarrow 0,288; ИП-4 \rightarrow М \rightarrow 0,156; ИП-5 \rightarrow НМ \rightarrow 0,288.

Таким образом, в рамках моделирования инвестиционного процесса в условиях неопределенности, характеризующегося тем, что полезный эффект от инвестиций на каждом шаге расчетного периода адекватно может быть описан только совокупностью количественных и качественных показателей «прибыли», предложен подход, позволяющий интегрировать эту совокупность в единую количественную оценку ИП. Эта оценка является условной — оценка с позиции конкретного шага расчетного периода — и необходима для последующей иерархической композиции, результатом которой является интегральная количественная оценка ИП с позиции ведущего критерия, учитывающая все многообразие последствий от реализации ИП.



MODELLING OF THE INVESTMENT PROCESS IN THE CONDITIONS OF UNCERTAINTY

V. A. KALUGIN

I. N. TITOVA

Belgorod State University

e-mail: Kalugin@bsu.edu.ru

e-mail: Titova@bsu.edu.ru

It is shown that the useful effect from investments in the conditions of uncertainty on each step of the settlement period can be adequately described only with a set both quantitative (monetary streams), and qualitative indicators, characterising "profit" on project realisation. The structure of innovative process and hierarchical model of process of acceptance of investment decisions is presented. The way of modelling in conditions uncertainty useful effect from investments and a way of construction of a uniform quantitative estimation investment project on each step of the settlement period is offered.

Key words: investments, investment process, the uncertainty and risk conditions, useful effect from investments, a subjective scale of qualitative gradation, methodology of the analysis of hierarchical structures, a scale of relations, a matrix of pair comparisons, the main own vector of a matrix.

ОТРАСЛЕВЫЕ РЫНКИ И РЫНОЧНАЯ ИНФРАСТРУКТУРА

УДК 658.1+338.2

МОДИФИЦИРОВАННЫЕ МЕТОДЫ ЗАТРАТНОГО ПОДХОДА К ОЦЕНКЕ ТОВАРНЫХ ЗНАКОВ РОССИЙСКИХ ПРОМЫШЛЕННЫХ ПРЕДПРИЯТИЙ

Е. И. КАШИРИНА

*Ростовский
государственный
экономический
университет «РИНХ»*

e-mail: ktyxbr@inbox.ru

В условиях повышения роли управления предприятием на основе стоимости всех его активов, в том числе и нематериальных, становится существенной необходимостью изучения подходов и методов оценки стоимости товарных знаков промышленных предприятий. В статье автором выявлены основные этапы жизненного цикла товара, маркированного товарным знаком промышленного предприятия и определена стадия максимальной достоверности полученных в рамках затратного подхода результатов. Проведен анализ и предложены модификации основных методов затратного подхода к оценке товарных знаков промышленных предприятий на основе внесения корректировок в базовые формулы. В заключении автором сделаны выводы и отмечены достоинства и недостатки анализируемого подхода.

Ключевые слова: товарный знак, оценка стоимости, затратный подход, промышленные предприятия, жизненный цикл товара, затраты на замещение, затраты на воспроизводство, обесценение.

В современных условиях мирового финансового кризиса и волатильности фондового рынка капитализация крупных российских промышленных предприятий подвержена изменению. При этом стоимость товарных знаков таких предприятий имеет тенденцию к ежегодному увеличению, о чем свидетельствует рис. 1, иллюстрирующий изменения капитализаций таких крупных российских компаний, как ОАО «Балтика», ОАО «Лукойл» и ОАО «Вимм-Билль-Данн» и стоимости товарных знаков, правообладателями которых они являются – «Балтика», «Лукойл» и «Домик в Деревне»¹.

¹ Составлено автором на основе данных сайта фондовой биржи РТС <http://www.rts.ru>, сайта компании Interbrand <http://www.interbrand.ch>.

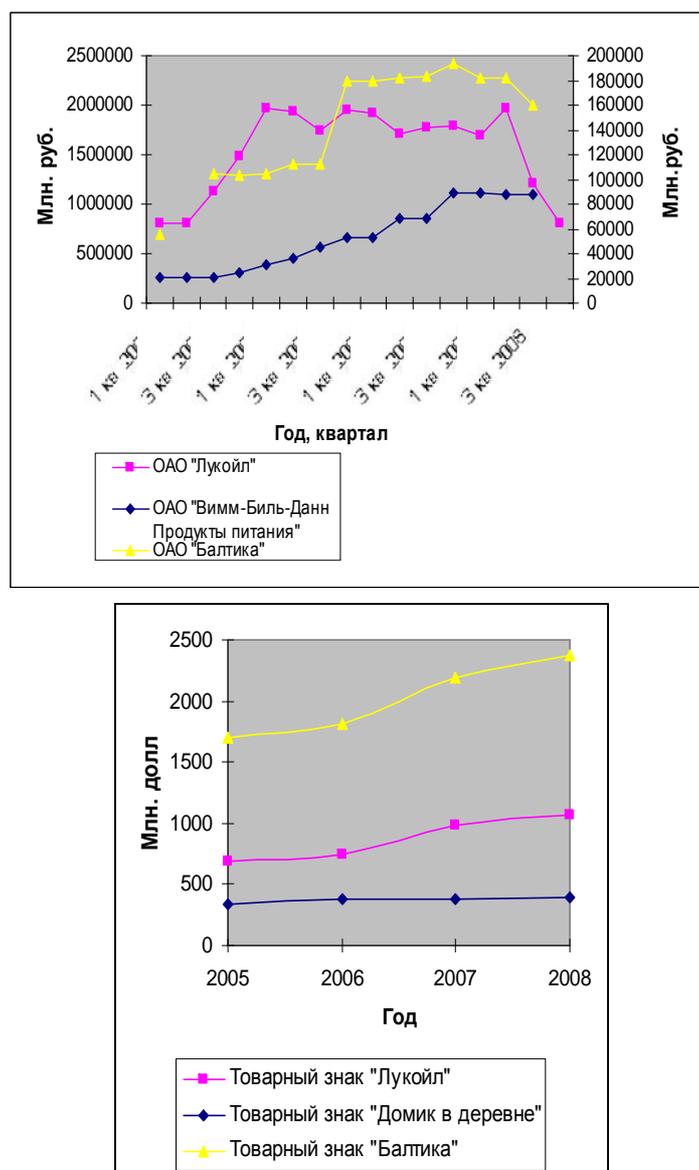


Рис. 1. Динамика изменения капитализаций и стоимости товарных знаков промышленных предприятий ОАО «Балтика», ОАО «Лукойл» и ОАО «Вимм-Биль-Данн» с 2005 по 2008 гг.

В этих условиях повышается роль управления предприятием на основе стоимости всех его активов, в том числе и нематериальных и, как следствие, становится существенной необходимостью изучения подходов и методов оценки стоимости товарных знаков промышленных предприятий.

В соответствии со сложившейся практикой и нормативными требованиями при оценке стоимости различных объектов имущества, в том числе и нематериальных активов, используются три основных подхода: доходный, затратный и сравнительный, которые могут приводить к различным результатам. Однако на практике специалистами занижается значимость применения затратного подхода к оценке товарных знаков.

Сущность затратного подхода к оценке исследуемого актива заключается в определении современной рыночной стоимости посредством суммирования всех расходов, необходимых для его создания и продвижения на рынок. В нашей статье анализируются товарные знаки промышленных предприятий. Под последними мы пони-

маем предприятия, занимающиеся производством, отгрузкой и реализацией продукции, и являющиеся точечным объектом: завод, фабрика, шахта, комбинат и т.д. Товарные знаки промышленных предприятий делятся на промышленные и потребительские товарные знаки. Промышленные товарные знаки в большей степени, чем потребительские, связаны с заказчиком. Создавая товарный знак, компания должна максимально акцентировать свое внимание на том, как удовлетворяется потребность заказчика за счет текущего продуктового портфеля компании и существующей системы работы с клиентом.

Проблема определения стоимости товарного знака с помощью затратного подхода состоит в том, что на практике расчет ее является сложной процедурой. «Сильный» товарный знак по определению уникален и затраты на его создание не отражают действительной стоимости товарного знака. Товарный знак может ничего не стоить, даже в случае, если на его создание были потрачены тысячи миллионов долларов.

Существуют, однако, ситуации, когда при определении стоимости товарного знака затратный подход является единственно правомерным и наиболее достоверным. Например, стоимость вновь создаваемых товарных знаков (знаков, еще не выпущенных на рынок) и не имеющих аналогов, невозможно рассчитать никаким другим подходом, кроме затратного.

При оценке стоимости товарного знака обязательным условием является анализ жизненного цикла оцениваемого актива, т.к. от него зависит точность и достоверность полученных результатов. При определении стоимости товарного знака необходимо проводить глубокий анализ объема продаж ассортимента товаров в течение всего их жизненного цикла.

Жизненный цикл товара, маркированного товарным знаком промышленного предприятия, отличается от типичного жизненного цикла любого товара без зарегистрированного товарного знака (появление на рынке, рост, зрелость, угасание). Он включает в себя следующие стадии (см. рис.2):

- I. Маркетинговые исследования
- II. Правовая защита (стадия разработки и регистрации товарного знака)
- III. Промышленное освоение
- IV. Выведение товара на рынок
- V. Серийное производство
- VI. Массовое производство
- VII. Насыщение рынка товаром
- VIII. Спад производства.
- IX.

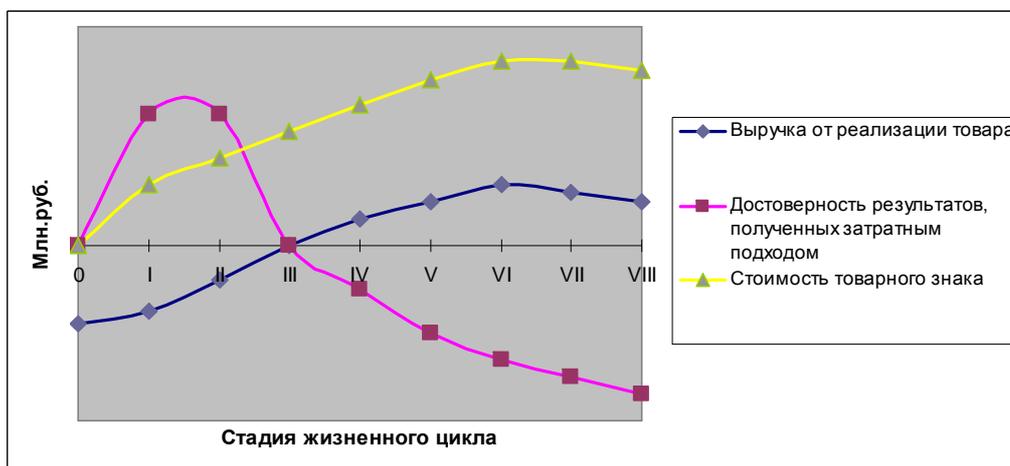


Рис. 2. Жизненный цикл товара, маркированного товарным знаком промышленного предприятия

Достоверность затратного подхода является максимальной на стадии разработки и регистрации товарного знака. Величина, полученная с помощью затратного подхода, достаточно точно отражает стоимость разработки до ее коммерциализации, но дальше этого момента, с которого, собственно и начинается отдача в инновационном бизнесе, метод не действует².

После начала выпуска продукции, маркированной товарным знаком, и выхода данного товара на рынок, достоверность полученных с помощью затратного подхода расчетов снижается, поскольку на этой стадии уже видна ожидаемая привлекательность исследуемого актива, и потенциальные покупатели готовы инвестировать средства для получения дальнейшей прибыли. В данном случае сумма затрат на разработку и регистрацию товарного знака не вполне адекватно отражает стоимость последнего, т.к. инвестора интересует именно возможность получения прибыли, а не затраты. На стадиях серийного, массового производства, насыщенности рынка и стадии спада производства маркированного товара достоверность результатов, полученных с помощью затратного подхода, минимальна, поскольку уже известно, каким спросом пользуется выпускаемая продукция и затраты на разработку, регистрацию и продвижение на рынок товарного знака не являются значимым фактором формирования его стоимости.

Тем не менее, затратный подход необходим как обязательная составная часть других подходов к оценке товарных знаков. В расчетах значение, полученное одним из методов затратного подхода, в большинстве случаев, служит ориентиром нижней границы стоимости.

В рамках затратного подхода при определении стоимости интеллектуальной собственности, в том числе товарных знаков, большинство отечественных специалистов в этой области (Н.Ю. Пузыня, Ю.Б. Леонтьев, Б.Б. Леонтьев, Г.Г. Азгальдов, А.Н. Козырев, В.Г. Зинов)³ выделяют следующие методы:

- 1) метод начальных (исходных) затрат;
- 2) метод затрат на воспроизводство;
- 3) метод затрат на замещение.

Затраты на воспроизводство – это сумма затрат на создание точной копии оцениваемого актива в существующих на дату оценки рыночных ценах с учетом его устаревания.

Затраты на замещение – это сумма затрат на создание аналогичного по своим функциональным и полезным свойствам оцениваемому товарному знаку с учетом его устаревания. Полезность представляет собой способность замещающего объекта обеспечивать эквивалентную степень удовлетворения в сопоставлении с оцениваемым товарным знаком.

Основой данных методов является суммирование всех затрат, связанных с созданием и продвижением товарного знака за весь период его существования, хотя между затратами на воспроизводство и затратами на замещение имеются достаточно тонкие и важные отличия, которые необходимо учитывать при определении стоимости товарного знака. Так, при определении затрат на воспроизводство не учитываются ни рыночный спрос на продукт, маркированный товарным знаком, ни его принятие рынком, т.е. не рассматривается вопрос о том, хочет ли потребитель иметь точную копию оцениваемого товарного знака. При определении затрат на замещение не всегда учитывается тот факт, что замещающий товарный знак может иметь меньшую или большую полезность, чем оцениваемый актив.

Представляется возможным в рамках затратного подхода модифицировать и классифицировать методы следующим образом:

² Зинов В.Г. Управление интеллектуальной собственностью. М.: Дело, 2003. С. 73.

³ См.: Зинов В.Г. Управление интеллектуальной собственностью. М.: Дело, 2003. С. 73; Леонтьев Б.Б., Мамаджанов Х.А. Основы оценки интеллектуальной собственности в России. М.: ИННЦ Роспатента, 2006. С. 91; Пузыня Н.Ю. Оценка интеллектуальной собственности и нематериальных активов. СПб: Питер, 2005. С. 263.



- 1) метод суммирования затрат на замещение товарного знака на дату оценки;
- 2) метод суммирования затрат на воспроизводство товарного знака на дату оценки;
- 3) метод индексации фактически произведенных затрат.

Предложенные методы предполагают определение рыночной стоимости товарного знака посредством выполнения следующих процедур:

- 1) определение суммы всех затрат;
- 2) определение величины обесценения товарного знака;
- 3) расчет рыночной стоимости товарного знака путем вычитания из суммы затрат величины обесценения.

Однако имеются значительные отличия на этапе определения затрат.

Метод суммирования затрат на замещение товарного знака на дату оценки применяется когда имеется возможность посчитать сумму затрат на создание аналогичного товарного знака промышленного предприятия в рыночных ценах на дату оценки.

При расчете суммы затрат на создание аналогичного товарного знака товара той же категории, назначения и качества определяется величина затрат на доведение «небрендового» товара промышленного предприятия до того же уровня, что и оцениваемый товарный знак или сумма затрат, связанных с заменой одного знака на другой для данного предприятия. Необходимо учитывать, что колебания спроса и предложения на выпускаемую под товарным знаком продукцию могут влиять на затраты как в сторону увеличения, так и в сторону уменьшения. Нерентабельное промышленное предприятие – правообладатель товарного знака – в значительной степени может снизить рыночную стоимость последнего. Стоимость уникального товарного знака может снизиться в большем объеме, поскольку он мало используется за пределами конкретного бизнеса. Товарные знаки, имеющие более широкое распространение, могут иметь пониженную стоимость в размере затрат, необходимых на вывод их из бизнеса, перевозку и установку на новом предприятии для использования в более рентабельной области производства.

Затратами, связанными с замещением товарного знака на другой, для данного предприятия являются следующие:

- расходы на изменение физических носителей: витрин, униформы, канцелярских принадлежностей, логотипа и других форм идентификации товарного знака, как внутренних, так и внешних;
- маркетинговые расходы по изменению товарного знака;
- временное падение доходов и прибылей, связанных с товарным знаком.

Таким образом, рассчитать затраты на замещение товарного знака в рамках анализируемого метода, возможно двумя способами:

- 1) если известна стоимость товарного знака с меньшей полезностью, формула для расчета затрат примет следующий вид:

$$Z_3 = C_{атз} + Z \quad (1)$$

Z_3 – затраты на замещение товарного знака;

$C_{атз}$ – стоимость аналогичного товарного знака с меньшей полезностью;

Z – затраты на доведение аналогичного товарного знака до уровня оцениваемого (расходы на улучшение качества производимого товара, промо-акции, рекламу, маркетинговые исследования);

- 2) при замещении одного товарного знака другим на данном предприятии формула для расчета затрат примет следующий вид:

$$Z_3 = Z_{ифи} + Z_{мп} + УВ \quad (2)$$

Z_3 – затраты на замещение товарного знака;

$Z_{ифи}$ – затраты на изменение форм идентификации товарного знака;

$Z_{мп}$ – маркетинговые расходы по изменению товарного знака;

$УВ$ – упущенная выгода правообладателя товарного знака, связанная с недополучением дохода за время замещения одного товарного знака другим.

Затраты на воспроизводство или замещение товарного знака включают в себя не только затраты на разработку, правовую регистрацию, маркетинговые исследования, расходы на рекламу, связанные с продвижением товарного знака, но и прибыль создателя и правообладателя знака – величина наиболее вероятного вознаграждения за инвестирование капитала в создание товарного знака.

Затем сумма полученных затрат с учетом прибыли создателя и правообладателя товарного знака должна быть уменьшена на величину обесценения. Поскольку данные процедуры предполагают и два остальных метода в рамках анализируемого подхода, то итоговые формулы будут рассмотрены ниже.

Метод суммирования затрат на воспроизводство товарного знака на дату оценки применяется, когда имеется возможность подсчитать сумму затрат на создание идентичного товарного знака в рыночных ценах на дату оценки.

Формула для расчета затрат на воспроизводство товарного знака имеет следующий вид:

$$Z_v = Z_m + Z_d + Z_{по} + Z_p \quad (3)$$

Z_v – затраты на воспроизводство товарного знака;

Z_m – затраты на маркетинговые исследования;

Z_d – затраты на дизайн и разработку товарного знака;

$Z_{по}$ – затраты на правовую охрану;

Z_p – затраты на рекламу и продвижение товарного знака на рынок.

Метод индексации фактически произведенных затрат используется, когда имеется документально подтвержденная информация о тех расходах, которые были понесены в прошлом на создание (приобретение) и выведение на рынок товарного знака, включая расходы по приведению его в состояние, в котором он пригоден к использованию в запланированных целях.

По мнению Б.Б. Леонтьева, необходимо выполнить следующие процедуры⁴:

1) критически проанализировать полученную информацию и выделить те суммы, которые действительно пошли на разработку данного актива, т.е. провести корректировку затрат по существу;

2) провести корректировку стоимости ввиду разницы во времени вложения средств, поскольку это вложение осуществлялось в разные периоды, в течение которых стоимость денег, как правило, не оставалась постоянной.

Расходы, которые были понесены в прошлом на создание (приобретение) и выведение на рынок товарного знака, включая расходы по приведению его в состояние, в котором он пригоден к использованию в запланированных целях, включают в себя:

– затраты на маркетинговые исследования;

– затраты на разработку и регистрацию товарного знака;

– расходы, оплачиваемые сторонними организациями за информационные и консультационные услуги, связанные с созданием или приобретением товарного знака;

– в случае приобретения исключительных прав на товарный знак обязательно учитываются регистрационные сборы, таможенные и патентные пошлины и иные аналогичные платежи, произведенные в связи с приобретением товарного знака;

– невозмещаемые налоги, уплачиваемые при приобретении товарного знака;

– затраты на рекламу продукции, маркируемой товарным знаком, и др.

Однако, как показывает практика крупных промышленных предприятий – правообладателей общеизвестных товарных знаков, стоимость таких товарных знаков может в несколько раз отличаться от тех затрат, которые предприятия производят для их рекламирования (см. рис. 3)⁵.

⁴ Леонтьев Б.Б., Мамаджанов Х.А. Основы оценки интеллектуальной собственности в России. М.: ИННЦ Роспатента, 2006. С. 91.

⁵ Составлено автором на основе данных официального сайта Пивоваренной компании «Балтика» <http://www.baltika.ru>, сайта электронной версии журнала «Эксперт» <http://www.expert.ru/printissues/>

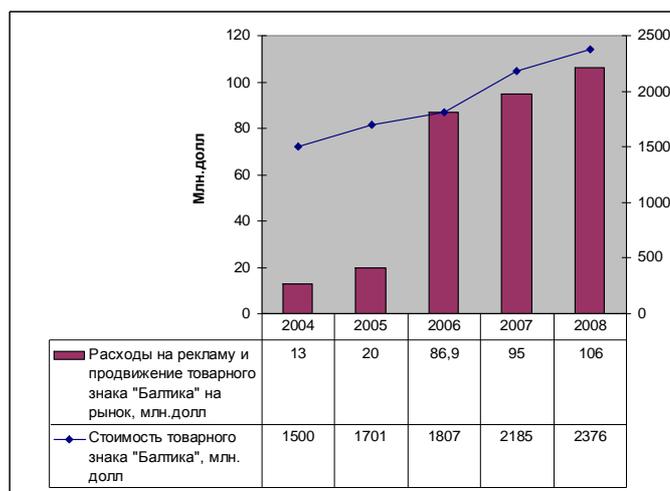


Рис. 3. Изменение стоимости и расходов на рекламу и продвижение товарного знака «Балтика» на рынок с 2004 по 2008 гг., в млн. долл.

Данные по фактически произведенным затратам извлекаются из документов бухгалтерского учета и отчетности предприятия за все года, на протяжении которых выполнялась работа по созданию и продвижению на рынок товарного знака. Следует обратить внимание на величину затрат и даты осуществления этих затрат.

После выявления всех произведенных в прошлом затрат, производится их корректировка по каждому году на величину индекса цен на дату оценки. Для товарных знаков промышленных предприятий представляется возможным при пересчете прошлых затрат в текущую стоимость использовать индекс цен производителей промышленной продукции (ИЦП), который рассчитывается на основе зарегистрированных цен на товары. Данный индекс используется для изучения и характеристики ценовых процессов в экономике, сравнительного анализа изменения цен на продукцию отдельных отраслей промышленности, при переоценке основных фондов, проведении приватизации, пересмотре ставок арендной платы, индексации задолженности промышленных организаций. ИЦП применяется при выполнении различных экономических расчетов и прогнозировании на макроуровне, для оценки в неизменных ценах объемов промышленного производства и валового внутреннего продукта и т.д.⁶

К величине проиндексированных прошлых затрат, как и в вышеописанных методах, прибавляется разумная прибыль создателя и прибыль правообладателя товарного знака и вычитается в денежном выражении величина функционального и экономического обесценения товарного знака в соответствии с уровнем актуальности и полезным сроком действия.

Прибыль создателя товарного знака – доход разработчика товарного знака на капитал, включая затраты на координацию и затраты труда в течение периода разработки, достаточный для стимулирования процесса разработки товарного знака⁷.

Прибыль правообладателя товарного знака может быть рассчитана, исходя из ставок отдачи на капитал при его наиболее вероятном аналогичном по уровню риска инвестировании и периода времени, необходимого для создания оцениваемой интеллектуальной собственности.

expert/2004/07/, сайта ежедневной деловой газеты «РБКdaily» <http://www.rbcdaily.ru/2007/08/21/media/288320>, сайта компании Interbrand <http://www.interbrand.ch>.

⁶ Об утверждении методологических положений по наблюдению за ценами производителей промышленной продукции и расчету индексов цен производителей: Постановление Госкомстата РФ от 21.05.99 № 33 (По состоянию на 18 октября 2006 года).

⁷ Рейли Р., Швайс Р. Оценка нематериальных активов / Пер. с англ. / Под ред. В.М. Рутгайзера. М.: Квинто-Консалтинг, 2005. С. 150.

Товарные знаки подвержены двум типам обесценения – функциональному и экономическому (внешнему)⁸.

Функциональное обесценивание – это уменьшение стоимости товарного знака из-за невозможности осуществлять маркетинговые функции (или извлекать полезность).

Экономическое (внешнее) обесценивание – это уменьшение стоимости оцениваемого товарного знака вследствие правовых, социальных или экономических воздействий, являющихся внешними (и неконтролируемыми) по отношению к текущему использованию или состоянию товарного знака.

Многие авторы (Г.Г. Азгальдов, Н.Н. Карпова, А.Н. Козырев, В.Л. Макаров, Б.Б. Леонтьев, Н.Ю. Пузыня) предлагают расчет общего обесценения товарного знака следующим образом:

$$K_o = 1 - T_{\phi}/T_n \quad (4)$$

K_o – коэффициент обесценения товарного знака;

T_{ϕ} – фактический срок действия свидетельства на товарный знак на дату оценки;

T_n – нормативный срок действия свидетельства.

Срок действия свидетельства на товарный знак очень часто не совпадает со сроком полезного использования, в течение которого правообладатель товарного знака предполагает использовать его для получения прибыли. Таким образом, для получения более достоверных результатов, следовало бы в знаменателе использовать именно этот показатель:

$$K_o = 1 - T_{\phi}/T_{\text{пнтз}} \quad (5)$$

K_o – коэффициент обесценения товарного знака;

T_{ϕ} – фактический срок действия свидетельства на товарный знак на дату оценки;

$T_{\text{пнтз}}$ – срок предполагаемого полезного использования товарного знака.

Следовательно, формулы определения стоимости товарного знака методом суммирования затрат на замещение товарного знака на дату оценки, имеют следующий вид:

$$C_{\text{тз}} = Z_3 * K_o \quad (6)$$

$$C_{\text{тз}} = (Z_3 + П_c + П_{\text{п}}) * K_o \quad (7)$$

Формула определения стоимости товарного знака методом суммирования затрат на воспроизводство товарного знака на дату оценки, имеет следующий вид:

$$C_{\text{тз}} = (Z_v + П_c + П_{\text{п}}) * K_o \quad (8)$$

$C_{\text{тз}}$ – стоимость товарного знака

Z_3 – затраты на замещение;

Z_v – затраты на воспроизводство;

$П_c$ – прибыль создателя товарного знака;

$П_{\text{п}}$ – прибыль правообладателя товарного знака;

K_o – коэффициент обесценения товарного знака.

Общая формула определения стоимости товарного знака методом индексации фактически произведенных затрат имеет следующий вид:

$$C_{\text{тз}} = \left(\sum (Z_i * K_{\text{пi}}) + П_c + П_n \right) * K_o \quad (9)$$

$C_{\text{тз}}$ – стоимость товарного знака;

Z_i – затраты, произведенные в i -том году, связанные с созданием и выводением на рынок товарного знака;

$K_{\text{пi}}$ – индекс цен производителей промышленной продукции (ИЦП);

$П_c$ – прибыль создателя товарного знака;

$П_n$ – прибыль правообладателя товарного знака;

K_o – коэффициент обесценения товарного знака.

⁸ Пузыня Н.Ю. Оценка интеллектуальной собственности и нематериальных активов СПб: Питер, 2005. С. 263.



Основными достоинствами проанализированных методов затратного подхода является их простота, универсальность (подходят для большинства рынков), доступность, предполагающая низкие затраты на проведение оценочных процедур.

Недостатком является то, что величина затрат на создание и поддержание товарного знака не свидетельствует об эффективности маркетинговых мероприятий.

Таким образом, в статье предложены модификации основных методов затратного подхода к оценке товарных знаков промышленных предприятий:

- определение затрат на замещение товарного знака при имеющейся информации о стоимости аналогичного товарного знака с меньшей полезностью;
- учет упущенной выгоды правообладателя товарного знака, связанной с недополучением дохода за время замещения одного товарного знака другим;
- использование индекса цен производителей промышленной продукции при пересчете прошлых затрат в текущую стоимость;
- учет в итоговой формуле расчета соотношения стоимости и величины проиндексированных прошлых затрат не только прибыли правообладателя, но и прибыли создателя товарного знака;
- учет не нормативного срока действия свидетельства, а срока предполагаемого полезного использования товарного знака при расчете коэффициента обесценения товарного знака;

Их можно использовать не только для оценки вновь создаваемых товарных знаков (знаков, еще не выпущенных на рынок), но и для определения стоимости давно функционирующих товарных знаков. При этом необходимо учитывать, что наиболее точные и достоверные результаты, рассчитанные методами затратного подхода, будут получены на II стадии жизненного цикла товара, маркированного оцениваемым товарным знаком – на стадии разработки и регистрации товарного знака.

MODIFIED METHODS OF COST APPROACH IN THE RUSSIAN INDUSTRIAL ENTERPRISES' TRADEMARKS EVALUATION

E. I. KASHIRINA

*Rostov State Economic
University «RINH»*

e-mail: ktyxbr@inbox.ru

With the increasing role of enterprise management based on the value of all its assets, including intangible, the need to explore approaches and methods of industrial enterprises' trademarks valuation has become essential. In the article the author identifies the main stages of the life cycle of product, marked by the industrial enterprises' trademark and defined the stage of maximum reliability obtained through the cost approach results. The analyzed and proposed modification of the basic methods of cost approach to assess the value of industrial enterprises' trademarks through adjustments in the basic formula. In the conclusion we can observe the findings and advantages and disadvantages of the analyzed approach.

Key words: trade mark, value assessment, cost approach, industrial enterprises, product life-cycle, depreciated replacement costs, reproduction costs, loss of value.

ФУНКЦИОНАЛЬНЫЕ ОСОБЕННОСТИ И ПЕРСПЕКТИВЫ РАЗВИТИЯ ОПТОВОЙ ТОРГОВЛИ СИСТЕМЫ ПОТРЕБИТЕЛЬСКОЙ КООПЕРАЦИИ

Р. П. МЕШЕЧКИНА¹⁾
В. В. ЛАЗАРЕВ²⁾

*¹⁾ Белгородский
университет
потребительской
кооперации*

²⁾ ФГУП ВИОГЕМ

*e-mail:
lvitaliy01@yandex.ru*

В статье изложены функции оптовой торговли, раскрыто их содержание, выявлены характерные для оптовой торговли потребительской кооперации функциональные особенности. Определены основные направления развития оптовой торговли потребительской кооперации на современном этапе развития экономики России. Выявлена необходимость предоставления экономической информации о состоянии рынка оптовыми организациями своим покупателям как одного из основных направлений развития оптовой торговли потребительской кооперации. Охарактеризованы элементы информации по исследованию рынка. Обоснованы перспективы развития исследуемой сферы торговли, связанные с предоставлением оптовыми организациями широкого спектра услуг покупателям. Представлены направления аналитического учета предоставляемых услуг оптовыми организациями.

Ключевые слова: оптовая торговля, потребительская кооперация, функции, функциональные особенности, услуги, информация, направление развития оптовой торговли, хранение товара, покупатель, поставщики, информационный центр, реализация, ассортимент, оптовые посредники, товарное обращение.

Оптовая торговля осуществляет реализацию практически всех видов продукции производственно-технического назначения и индивидуального потребления, находится в центре хозяйственных связей, осуществляемых оптово-посредническим звеном, как с производством, так и с потребителями. Она является важнейшим опосредованным звеном между сферами производства и потребления в процессе воспроизводства. Оптовая торговля является активным проводником товаров отечественного производства и большой массы зарубежной продукции на внутренний рынок, что стимулирует отечественных производителей повышать качество товаров в целях завоевания прочных позиций как на внутреннем, так и на внешних рынках.

Оптовые торговцы необходимы производителям и покупателям товаров потому, что они обеспечивают эффективность торгового процесса.

При административно-плановом регулировании экономики предприятия оптовой торговли потребительской кооперации выполняли распределительную функцию, а также осуществляли хранение и подсортировку товаров.

В условиях рыночного хозяйствования изменяются функции оптовой торговли: с одной стороны, они усложняются, а с другой стороны – персонифицируются. В связи с этим оптовой торговле потребительской кооперации необходимо функционально перестраиваться в соответствии со сложившейся ситуацией на рынке¹.

В потребительской кооперации использование оптовых торговцев как профессионалов является эффективным способом реализации продукции их производителями и приобретения покупателями в силу следующих выгод.

Во-первых, имея дело преимущественно с профессиональными (розничными) клиентами в торговле, а не с конечными потребителями, оптовик уделяет меньше внимания стимулированию и расположению своего торгового предприятия (обычным заботам в этой сфере).

Во-вторых, оптовая торговля эффективнее розничной как способ вложения капитала (оптовые сделки обычно и крупнее розничных, и их торговая зона шире).

¹ Хваткин Н. Г. Экономика оптовой торговли продовольствием. М. : ЭКСМО, 2007. 400 с.

В-третьих, учитывая это различие, государство в части правовых норм и налогов подходит к оптовым и розничным торговцам по-разному (естественным образом выделяя налогоплательщиков разного масштаба).

В-четвертых, мелкому производителю с ограниченными финансовыми ресурсами не под силу создать и содержать организацию прямого маркетинга (т.е. реализовать товары через каналы нулевого уровня или непосредственно потребителям), поэтому он предпочитает направлять средства на развитие производства, а не на организацию оптовой торговли.

В-пятых, эффективность деятельности оптовиков, как правило, оказывается выше, благодаря не только размаху операций, но и большому числу деловых контактов в сфере розницы, наличию у них специальных знаний и умений.

И, наконец, в-шестых, розничные торговцы, имеющие дело с широким товарным ассортиментом, нередко предпочитают закупать весь набор товаров у одного оптовика, а не по частям у разных производителей².

Следовательно, у производителей и у розничных торговцев как контрагентов оптовой торговли есть все основания прибегать к услугам оптовиков, когда с их помощью можно более эффективно выполнять одну или несколько следующих хозяйственных функций:

1. Стимулирование сбыта за счет подключения персонала оптовиков к охвату множества мелких клиентов при сравнительно небольших затратах, поскольку оптовикам покупатели верят больше, чем какому-то далекому производителю.

2. Закупки и формирование товарного ассортимента, которые передаются оптовику, избавляют клиента (розничного торговца или производителя) от значительных хлопот.

3. Разбивка крупных партий товаров на мелкие, осуществляемая оптовиками, обеспечивает клиентам экономию средств при закупке товаров вагонами и разбивке больших партий на мелкие (по вагонным или грузовым местам).

4. Складирование и хранение товарных запасов у оптовиков (как одна из основных функций последних) способствует снижению соответствующих издержек поставщика и потребителей.

5. Транспортировка как функция оптовиков обеспечивает более оперативную доставку товаров в силу того, что они находятся ближе к клиентам, чем производители.

6. Финансирование, осуществляемое оптовиками (в виде кредитования клиентов), распространяется и на поставщиков, когда заказы выдаются заблаговременно и вовремя оплачиваются счета.

7. Частичное принятие оптовиками на себя риска выгодно сторонам, поскольку оптовики принимают товар в свое собственное пользование и несут расходы в связи с его хищением, повреждением, порчей и устареванием.

8. Предоставление информации о рынке оптовиками своим поставщикам и клиентам о деятельности конкурентов, новых товарах, динамике цен и т.д. выгодно обеим сторонам.

9. Услуги по управлению и консультированию помогают розничным торговцам совершенствовать деятельность (за счет обучения их продавцов, участия в разработке схем технологической планировки магазина и устройства экспозиций, а также в организации систем бухучета).

В системе товарного обращения оптовая торговля выполняет целый ряд важных функций, дополняющих ее центральную, распределительную функцию между производителями и потребителями. Их содержание изложено в таблице³.

² Королева О.В. Развитие оптовой торговли потребительской кооперации в условиях рынка: диссертация / О. В. Королева. – Белгород: БУПК, 2001. 186 с.

³ Егоров В. Ф. Организация торговли. СПб.: Питер, 2006. 352 с.

Таблица

Содержание функций оптовой торговли

Функция 1	Содержание функции 2
формирование ассортимента	Оптовая торговля осуществляет закупки зачастую распыленных товаров вследствие специализации производства и дифференцированности спроса, т. е. она исследует предложение товаров и выбирает продукты для снабжаемого ею сегмента рынка по своему товарному ассортименту.
преодоление разрыва во времени	Оптовая торговля осуществляет функцию преодоления временного разрыва между моментами производства и потребления, например: нерегулярный подвоз южных фруктов, поставки строительных материалов с весны до осени и т. д.
преодоление пространственного разрыва	Функция преодоления расстояния между местом производства и местом потребления в любом случае является транспортной функцией, причем купленный товар может доставляться клиентам оптовой торговли или они могут забирать его сами на предприятии оптовой торговли.
формирование запасов (функция обеспечения надежности)	Служит для выравнивания колебаний спроса на различные товары в различные периоды времени.
гарантирование качества (функция манипулирования)	Означает, что оптовая торговля подготавливает товары к дальнейшей продаже. Это осуществляется путем сортировки, упаковки, смешивания ассортимента, очищения, т. е. путем «манипулирования». Сюда относятся, например, разлив импортного вина, созревание фруктов, хранение сырья до достижения им зрелости для целей переработки (древесина, табак) и т. д.
выравнивание цен	Заключается в использовании преимуществ по затратам благодаря скидкам при закупке больших партий товаров, транспортировке, упаковке, предложению товаров по сниженным ценам.
финансирование	Преодоление разрыва во времени между закупкой товара и его оплатой представляет собой кредитную функцию. По поставке предоставляется краткосрочный кредит, если торговое предприятие вносит задаток. Покупатель получает кредит («товарный кредит»), если торговое предприятие осуществляет поставки на срок (в кредит). Однако, на современном этапе в оптовой торговле системы потребительской кооперации, данная функция не получила широкого распространения.
исследование и освоение рынка	Предприятие оптовой торговли выполняет важную задачу, осваивая для имеющихся товаров новые рынки или расширяя рынок путем рекламы.
хранение на складе	Эта функция неразрывно связана с функцией преодоления разрыва во времени, поскольку ритмичность возникновения потребностей покупателей оптовой торговли зачастую не соответствует ритмичности предложения производителей. Например, большая часть яблок для производства сока собирается в октябре и поставляется на предприятия, производящие сок. Оптовая же торговля должна быть готова к поставкам яблок прежде всего в летние месяцы. Это возможно лишь при наличии достаточных складских мощностей.

Оптовая торговля потребительской кооперации России, поставляющая товары для осуществления торгового обслуживания сельского населения, выполняет функции не только экономические, но и социальные.

Как элемент инфраструктуры потребительского рынка оптовая торговля потребительской кооперации формирует необходимый ассортимент потребительских товаров и снабжает ими розничную торговлю, которая, являясь завершающим звеном движения товаров от производителя к потребителю, обеспечивает окончательное их признание, способствует началу нового витка воспроизводственного процесса.

Оптовая торговля как основной посредник в системе товародвижения связана, с одной стороны, с производителями (поставщиками) товаров, с другой стороны, с розницей. Учитывая то, что производителям выгодно выпускать большие партии

продукции узкой номенклатуры, а рознице – предлагать покупателям как можно более широкий ассортимент товаров, опт берет на себя посреднические функции.

Оптовая торговля как одна из организаций потребительской кооперации, вносит определенный вклад в развитие инфраструктуры села, содействует повышению реальных доходов сельского населения, способствуют сохранению и развитию культурного и духовного наследия.

Благосостояние общества в целом, и, в частности, населения, проживающего в сельской местности, зависит от функционирования системы удовлетворения многообразных потребностей, что оказывает непосредственное влияние на качество жизни населения, которое, в свою очередь, зависит от обеспеченности жильем, услугами образования, здравоохранения, уровня бытового, транспортного, культурного и торгового обслуживания. Многообразие форм собственности предопределило ситуацию, когда в торговом обслуживании сельского населения принимают участие оптовые предприятия различных типов, масштабов деятельности, стратегических направлений.

На рынке товаров и услуг оптовая торговля выполняет важную маркетинговую функцию, поскольку, наряду с обеспечением процесса реализации товаров, оптовые предприятия оказывают существенное влияние на формирование платежеспособного спроса и товарное предложение на рынке, чем обеспечивают определенный уровень рыночного равновесия и благоприятную конъюнктуру.

Предприятия оптовой торговли потребительской кооперации способствуют экономии издержек обращения, т.к. розничные предприятия освобождаются от закупок товаров у множества производителей, вследствие чего сокращаются материальные затраты, связанные с хранением, формированием ассортимента товаров и их доставкой.

К числу особенностей развития оптовой торговли на современном этапе следует отнести то, что предприятия оптовой торговли потребительской кооперации выступают центром сосредоточения и передачи информации по вопросам исследованию рынка, т.е. призваны выполнять информационную функцию. Именно информационные центры способны обеспечить сбор, накопление и обработку коммерческой информации, обобщив и проанализировав ее, передать поставщику или покупателю.

Эта информация включает следующие элементы:

- определение емкости рынка;
- определение возможности выхода на внешний рынок;
- характеристика рыночной ситуации;
- выявление текущего и перспективного спроса обслуживаемого населения;
- изучение потребительских характеристик, оценка конкурентоспособности товаров разных товаропроизводителей;
- выявление мотивации покупок и закономерностей поведения покупателей на рынке;
- изучение рынка товаров и возможностей их реализации;
- разработка предложений по обновлению и усовершенствованию выпускаемой продукции с учетом требований покупателей⁴.

Одним из основных направлений развития оптовой торговли потребительской кооперации является предоставление оптовыми предприятиями услуг покупателям и поставщикам товаров. Состав и виды оптовых торговых услуг в потребительской кооперации могут быть самыми разнообразными, исходя из условий товародвижения и информационно-коммерческого обслуживания участников рыночных отношений.

Предоставляемые услуги могут оплачиваться как поставщиками – производителями товаров, так и предприятиями розничной торговли. В этих условиях финансовое состояние оптовых предприятий потребительской кооперации будет зависеть

⁴ Голиков Е. А. Оптовая торговля. Менеджмент. Маркетинг. Логистика. Финансы, Безопасность. М.: Экзамен, 2004. 272 с.

не только от величины оборота оптовой торговли, но и от количества и качества услуг, оказываемых ими покупателям товаров.

Широкий спектр оптовых услуг позволит предприятиям оптовой торговли потребительской кооперации в условиях рынка успешно функционировать в конкурентной рыночной среде, избежать убыточности и добиться эффективной работы. В связи с переходом к рыночным отношениям, оптовые предприятия потребительской кооперации должны заинтересовать своих покупателей, привлечь их оказанием различных оптовых услуг. В этой связи важное значение имеет развитие такой услуги опта, как коммерческое кредитование, финансирование сделок. Опыт западных оптовых посредников показывает, что выбор канала товародвижения зависит от того, насколько его организатор может сам кредитовать сделки заинтересованных сторон. В практической деятельности оптовые фирмы зачастую финансируют производителя тем, что предоставляют ему заказ на определенный товар с гарантией его реализации и при этом в виде аванса оплачивают часть заказанной партии товаров. Финансирование розничных предприятий оптовые базы осуществляют путем реализации им товара с отсрочкой платежа. Условия предоставления кредитов разнообразны и зависят от объема торговой сделки, кредитоспособности покупателя, качества реализуемой продукции, экономической конъюнктуры.

Перспективным направлением развития кооперативной оптовой торговли является предоставление предприятиями исследуемой формы торговли своим клиентам институциональных услуг (договорные отношения с банками, органами сертификации и др.). Данная услуга может заключаться в том, что на территории оптовой базы будет находиться филиал коммерческого банка, что повлечет за собой повышение уровня оперативности расчетных операций при купле-продаже товаров⁵.

Не менее важным перспективным направлением развития оптовой торговли потребительской кооперации является предоставление рекламных услуг. Оптовые организации, осуществляя мероприятия по проведению рекламных акций, позволяют покупателям узнать о появляющихся товарах на рынке, об их потребительских свойствах. Для поставщиков реклама позволяет создать спрос на товар. Исходя из этого, оказание помощи в рекламировании товаров позволяет оптовым организациям расширять свою базу клиентов, привлекать новых поставщиков к сотрудничеству, что, в свою очередь, положительно отражается на конечном результате деятельности предприятия.

Помимо этого, важным направлением является также развитие услуг по управлению и консалтингу.

Для развития оптовой торговли потребительской кооперации предоставляется целесообразным оказание сервисных услуг по ремонту и гарантийному обслуживанию технически сложных товаров, консультирование работников розничной торговли по вопросам эксплуатации сложной техники, обеспечение их специальной литературой по организации торговли товарами отдельных групп, проведение мастер-классов для работников торговли.

Услуги оптовых организаций должны предоставляться клиентам на платной основе и способствовать активизации оптовой деятельности. Конкретные размеры оплаты услуг следует предусматривать в заключаемых между оптовым предприятием и клиентами договорах. Они должны отражать расходы на предоставление услуг с учетом их трудоемкости и способствовать достижению нормального уровня рентабельности этих операций.

Оптовые организации, оказывающие своим клиентам широкий спектр услуг, обеспечивают повышение эффективности своей деятельности, удовлетворяя потребности производителей и покупателей товаров. Это достигается при условии, когда величина дохода от предоставляемых услуг будет не ниже расходов, связанных с их

⁵ Повышение эффективности коммерческой деятельности оптовых предприятий потребительской кооперации: Диссертация. Белгород: БУПК, 2005. 216 с.

предоставлением. В связи с этим работникам коммерческих служб необходимо постоянно вести аналитический учет доходов от предлагаемых услуг по следующим направлениям:

- по характеру оказания услуг;
- в зависимости от спроса на услуги;
- по окупаемости;
- по увеличению оборота товаров, обусловленного выполняемыми услугами;
- по стоимости оказания услуг;
- по издержкам от предоставления услуг;
- по числу потребителей, пользующихся конкретными видами платных услуг⁵.

Привлекательность предлагаемых услуг может меняться в зависимости от сложившейся ситуации на рынке, вследствие чего необходимо расширять или уменьшать перечень услуг, проводить анкетные или устные опросы о возможном изменении их видов с целью повышения эффективности деятельности оптовых организаций, обеспечения деловых отношений с постоянными клиентами и привлечения новых.

Таким образом, оптовая торговля потребительской кооперации как важнейшая составляющая сферы товарного обращения, реализуя общие и специфические функции и являясь одной из отраслей общественного производства, в период становления в России рыночных отношений прошла существенный путь развития. Тем не менее, ряд проблем, связанных с развитием оптовой торговли потребительской кооперации, требует скорейшего решения, поскольку именно оптовая торговля, ее состояние и тенденции развития оказывают влияние на развитие экономики страны и ее социальную сферу, в том числе и на повышение качества жизни. Повышение уровня и качества жизни населения на территории, обслуживаемой организациями потребительской кооперации, предполагает наиболее полное удовлетворение его разносторонних потребностей. Уровень удовлетворения жизненно важной потребности в товарах, наряду с такими факторами, как обеспечение их предложения и наличие платежеспособного спроса, во многом зависит от того, насколько развит его инфраструктурный комплекс, и, в частности, сфера товарного обращения.

FUNCTIONAL PECULIARITIES AND PROSPECTS OF WHOLESALE TRADE DEVELOPMENT OF CONSUMER COOPERATION SYSTEM

R. P. MESHECHKINA ¹⁾

V. V. LAZAREV ²⁾

¹⁾ *Belgorod University
of Consumer Cooperation*

²⁾ *FGUP VIOGEM*

*e-mail:
lvitaliy01@yandex.ru*

The paper dwells on the function of wholesale trade, reveals their content, discloses functional peculiarities characteristic of consumer cooperation wholesale trade; defines main directions for the development of consumer cooperation wholesale trade at modern stage of the development of Russian economy; reveals the necessity of providing economic information on the situation in the market by wholesale units to their buyers as one of the main directions of consumer cooperation wholesale trade development; characterizes the elements of information on the market study; substantiates the prospects for the development of the trade sphere under study connected with the provision of a wide spectrum of services to buyers by wholesale units; provides the directions of analytical accountancy of the services provided by wholesale units.

Key words: wholesale trade, consumer cooperation, functions, functional peculiarities, services, information, wholesale trade development directions, goods storage, buyers, suppliers, information center, sales, assortment, wholesale intermediaries, goods turnover.

РЫНОК ТРУДА И ЭКОНОМИКА ОБРАЗОВАНИЯ

УДК 378:001.891

К ВОПРОСУ О НЕОБХОДИМОСТИ СТАНОВЛЕНИЯ СИСТЕМЫ ТРАНСФЕРА ЗНАНИЙ И ТЕХНОЛОГИЙ НА ОСНОВЕ ИННОВАЦИОННОГО ПОТЕНЦИАЛА ВЕДУЩИХ ВУЗОВ РОССИИ

М. В. ВЛАДЫКА

*Белгородский
государственный
университет*

*e-mail:
Yladyka@bsu.edu.ru*

Развитие инновационного потенциала вузов России связано с необходимостью коммерциализации результатов научно-технической деятельности и становлением системы трансфера знаний и технологий, то есть перехода от объектной к субъектной позиции той части инновационной системы, которая составляет инновационный потенциал. Решение проблем трансфера знаний и технологий, а также коммерциализации результатов научно-технической деятельности вузов, степень их разработанности и реализации не только не удовлетворяют, но и не соответствуют современному состоянию инновационного потенциала высшей школы России и национальной экономики.

Ключевые слова: инновационный потенциал, высшая школа, результаты научно-технической деятельности вузов, коммерциализация, трансфер, рынки знаний и технологий.

Интеграция российской экономики в глобальное экономическое пространство в условиях мирового финансового кризиса обозначила серьезные проблемы, касающиеся российского высшего образования. К высшей школе в современных условиях обращен ряд вызовов:

- рост количества вузов за последние десятилетия (массификация);
- спрос на высшее образование (массовизация);
- развитие рыночных отношений в области высшего образования (коммерциализация);
- сокращение государственного финансирования вузов;
- высокая степень неравенства среди учреждений высшего образования (дифференциация);
- трансформация целей и ценностей высшего образования (прагматизация);
- изменяющаяся организационно-экономическая роль государства по отношению к образованию;
- неопределенность места высшей школы в общественной структуре и социально-экономических отношениях.



Многие авторы¹ к одним из наиболее фундаментальных вызовов последних лет относят глобализацию и интернационализацию высшей школы как следствие превращения мировой экономики в XXI веке в «целостный глобальный организм... со всемирными производственно-сбытовыми структурами, глобальной финансовой системой, планетарной информационной сетью»². И.А. Майбуров на основе анализа мировых тенденций развития высшей школы определяет, что процессы интернационализации ведут к интеграции, а глобализации – к конкуренции в высшей школе и образовании³.

В результате ключевым преимуществом в данной конкурентной борьбе становится способность высшей школы к генерации, аккумуляции и трансляции не только знаний, но также результатов научно-технической, интеллектуальной и инновационной деятельности. Известно, что знания генерируются в процессе фундаментальных, а не прикладных исследований, что позволяет утверждать: для того, чтобы высшая школа активно развивала базовые конкурентные преимущества, необходима финансовая поддержка государства и субъектов рынка. Достижение конкурентных преимуществ высшей школы в современных условиях связано с применением инновационных стратегий и реализацией научно-технического и инновационного потенциала.

Развитие указанных процессов ведет к растущей конкуренции и появлению новых организационно-институциональных структур в высшей школе России. Принят пакет законов и указов, касающихся реорганизации сети высших учебных заведений и выделения группы наиболее крупных и конкурентоспособных, к которым относятся следующие:

1. Два высших учебных заведения, имеющих особый статус – Московский и Санкт-Петербургский университеты.
2. Сеть федеральных университетов, в том числе находящихся в стадии формирования (табл. 1).
3. Вузы, получившие статус национальных исследовательских университетов (14 вузов).

Таблица 1

Федеральные университеты высшей школы России

№№	Федеральный университет	Базовый вуз	Стадия
1	Южный	Ростовский государственный университет	Функционирует
2	Сибирский	Красноярский государственный университет	Функционирует
3	Северный (Арктический)	Архангельский технический университет	Создается
4	Приволжский	Казанский государственный университет им.В.И.Ленина	Создается
5	Уральский	Уральский технический университет им.Б.Н.Ельцина	Создается
6	Дальневосточный	Дальневосточный государственный университет	Создается
7	Северо-восточный	Якутский государственный университет им.М.К.Аммосова	Создается

В высшей школе происходят институциональные процессы концентрации (укрупнение и объединение вузов), федерализации (получение статуса особого, феде-

¹Никольский В. Глобальное образование: пределы либерализации // Высшее образование в России. 2004, №8. С.17; Скотт П. Глобализация и университет // ВВШ «Alma Mater», 2005, №4. С. 3–8; Денмен Б.Д. Появление трансрегиональных схем (структур) обмена в области образования в регионах Европы, Северной Америки и Тихоокеанской Азии // Высшее образование в Европе. 2001, №2; Ван дер Венде М. Послесловие к одной дискуссии // ВВШ «Alma Mater», 2000, №5. С.30-32.

²Potter W. Public colleges try to show their value to States, but not everyone is convinced// Chronicle of higher education. 2003. Vol. 49. № 35. May9. P.A26 -A28.

³Майбуров И.А. Парадигма согласованного развития высшей школы и промышленности в Регионах. Екатеринбург: ГОУ ВПО УГТУ – УПИ, 2003. 516 с.

рального и национального исследовательского университета), дифференциации (рейтинговое деление по типам и видам). Это связано с глобальными тенденциями развития мировой и национальной экономики, с ростом их инновационных потенциалов, конкурентоспособности вузов и превращением их в мировые центры науки и инноваций, в том числе с централизацией финансовых ресурсов и сосредоточением на наиболее перспективных направлениях развития.

В связи с необходимостью развития инновационного потенциала в национальных масштабах, активизацией инновационных процессов в высшей школе России и развитием стратегических конкурентных преимуществ университетов в долгосрочной перспективе на передний план выдвигается проблема трансфера знаний и технологий как форма реализации научно-технической продукции и коммерциализации результатов интеллектуальной деятельности вуза.

Становится очевидным, что одним из главных факторов реализации инновационного потенциала высшей школы России является процесс коммерциализации и трансфера результатов научно-технической деятельности (далее РНТД), обеспечивающий возможность выхода на рынки технологий, тесных связей с промышленностью и бизнесом, ростом доходов и, как следствие, реализацию более высокого научного, репутационно-имиджевого (брендового), социально-экономического статуса вуза. Представляется, что дальнейшее расширение «критической массы» инновационного потенциала высшей школы связано с необходимостью коммерциализации и трансфера РНТД, то есть перехода от объектной к субъектной позиции той части инновационной системы, которая составляет инновационный потенциал.

В научной литературе процесс реализации научно-технической продукции определяется термином «передача/трансфер технологии», осуществляемый в различных формах, разными способами и по разным направлениям⁴.

В широком смысле определение трансфера технологий подразумевает процесс, посредством которого фундаментальные исследования находят применение в производстве продукции и предоставлении услуг. Как экономическая категория трансфер технологии трактуется как передача научно-технических знаний, умений и опыта для оказания научно-технических услуг, применения технологических процессов, выпуска продукции⁵. Критерием является активное применение переданной технологии для производственных целей.

Передача технологии может осуществляться в различных формах, разными способами и по разным каналам. Она может передаваться на коммерческой и некоммерческой основе, быть внутрифирменной, внутригосударственной и международной. По направлению передача технологий может быть вертикальной, соответствующей стадиям «исследование – производство», и горизонтальной – передача информации из одной научной области в другую.

Трансфер знаний и технологий является одним из видов реализации результатов научно-технической деятельности и по экономическому содержанию осуществляется как в некоммерческой, так и в коммерческой (коммерциализация разработок) формах.

Основной поток передачи в некоммерческой форме приходится на непатентованную информацию: фундаментальные исследования, научные открытия и технологические изобретения, деловые игры и психологические тренинги и др. Некоммерческий трансфер чаще всего используется в области научных исследований фундаментального характера. Он обычно сопровождается небольшими расходами и мо-

⁴ Менеджмент технологических инноваций: Учеб. пособие / Под ред. проф. С. В. Валдайцева и проф. Н. Н. Молчанова. СПб.: Изд-во СПбГУ, 2003; Бовин А.А., Чередникова Л.Е. Интеллектуальная собственность: экономический аспект. М.; Новосибирск, ИНФРА-М: НГАЭиУ, 2007. С. 137; Бромберг Г.В. Интеллектуальная собственность (с учетом материалов 4-ой части Гражданского Кодекса Российской Федерации): Монография. М.: ИНИЦ Роспатента, 2008. С. 47.

⁵ Управление инновациями в организации: Учеб. пособие / А.А. Бовин, Л.Е. Чередникова, В.А. Якимович. М.: Омега-Л, 2006. С.138.



жет поддерживаться как по государственной линии, так и на основе фирменных и личных контактов.

Формы передачи технологии на некоммерческой основе:

- информационные массивы специальной научной литературы, банки данных, патенты, справочники и др.;
- конференции, выставки, симпозиумы, семинары и т. д.;
- обучение, стажировки, практика студентов, ученых и специалистов, осуществляемая на паритетной основе университетами, научными организациями, предприятиями и др.;
- перекрестное лицензирование на паритетной основе;
- миграция ученых и специалистов из одних организаций в другие, в том числе в другие страны;
- учреждение некоммерческих организаций в виде фондов, венчурных и внедренческих фирм, межвузовских объединений в виде некоммерческого партнерства, исследовательских подразделений и т. д.

Коммерческий трансфер или коммерциализация технологий означает процесс перехода результатов научных исследований в сферу практического применения, производства и маркетинга новых продуктов с целью получения коммерческой выгоды, осуществляется преимущественно в виде продажи патентов, лицензий на изобретения, ноу-хау и оформляется в виде договора (договора об уступке патента, лицензионного соглашения, договора о совместном производстве, о научно-техническом сотрудничестве и др.).

Основными формами коммерческого трансфера являются следующие:

- передача технологии в овеществленном виде (объекты промышленной собственности);
- передача технологии при прямых и портфельных инвестициях;
- продажа патентов;
- продажа лицензий на все виды запатентованной промышленной собственности, кроме товарных знаков, знаков обслуживания и т.д.;
- продажа лицензий на незапатентованные виды промышленной собственности – ноу-хау, секреты производства, технологический опыт, инструкции, чертежи, схемы, спецификации и т.д.

Следует выделить специфические особенности коммерческого трансфера технологий, к которым относятся высокая степень монополизации, поддерживаемая принятой системой охраны прав промышленной собственности, и, вследствие этого, – слабость позиции покупателя (в отличие от потребительского рынка – это рынок продавца); высокая норма прибыли (большая разница между ценой и стоимостью; асимметричное разделение на страны-импортеры (развивающиеся страны) и государства, которые одновременно являются экспортерами и импортерами (развитые страны); большой объем технологического трансфера внутри транснациональных корпораций.

Следует также различать трансфер разработок и технологий, который имеет место в случае, если предприятие – разработчик новой технологии по тем или иным причинам не может успешно коммерциализировать научно-технические разработки, уступает их другому субъекту деятельности, или же осуществляет целенаправленную деятельность по нетрадиционному (отличному от первоначального) использованию этих результатов в других отраслях или областях применения. Преимущественно идут на уступку и передачу технологии малые и средние предприятия; предприятия, выпускающие многие виды продукции; фирмы с высоким научно-техническим потенциалом; предприятия, владеющие инновациями как побочными продуктами НИОКР или второстепенными технологическими процессами, у которых трансфер значительно повышает потенциал коммерциализации разработки. Трансфер разработок нацелен на внедрение научно-технических разработок в практику и не связан

непосредственно с получением коммерческого эффекта, поэтому началом трансфера разработки можно считать формирование технически реализуемой идеи или научно-технического задела, а завершением – доведение новой разработки до производства.

Необходимо отметить, что новая технология, как и любой другой товар, должна иметь хозяина (владельца интеллектуальной собственности) и цену (иметь оценку стоимости интеллектуальной собственности). Особое значение имеют такие общие для всех видов товаров характеристики, как условия эксплуатации, безопасность и экологичность, конкурентные преимущества и результаты сравнительных испытаний, инжиниринговые услуги при внедрении и эксплуатации, бизнес-сопровождение, возможности адаптации к требованиям конкретного заказчика.

На рис. 1 показаны основные направления межрегионального трансфера технологий на последовательных стадиях НИОКР. Первое и второе направление оказывают наибольшее влияние на достижение краткосрочного финансового выигрыша и увеличение доли данного производителя на рынке. Третье направление способствует увеличению исследовательских ресурсов и оказывает долгосрочное воздействие на промышленную конкурентоспособность и уровень национальной безопасности: передача технологии новым компаниям (пунктирная линия); технологический трансфер из исследовательских организаций в действующие фирмы (непрерывная линия); передача высоких технологий для дальнейших исследований ведомственным лабораториям, университетам и объединениям (двойная пунктирная линия).

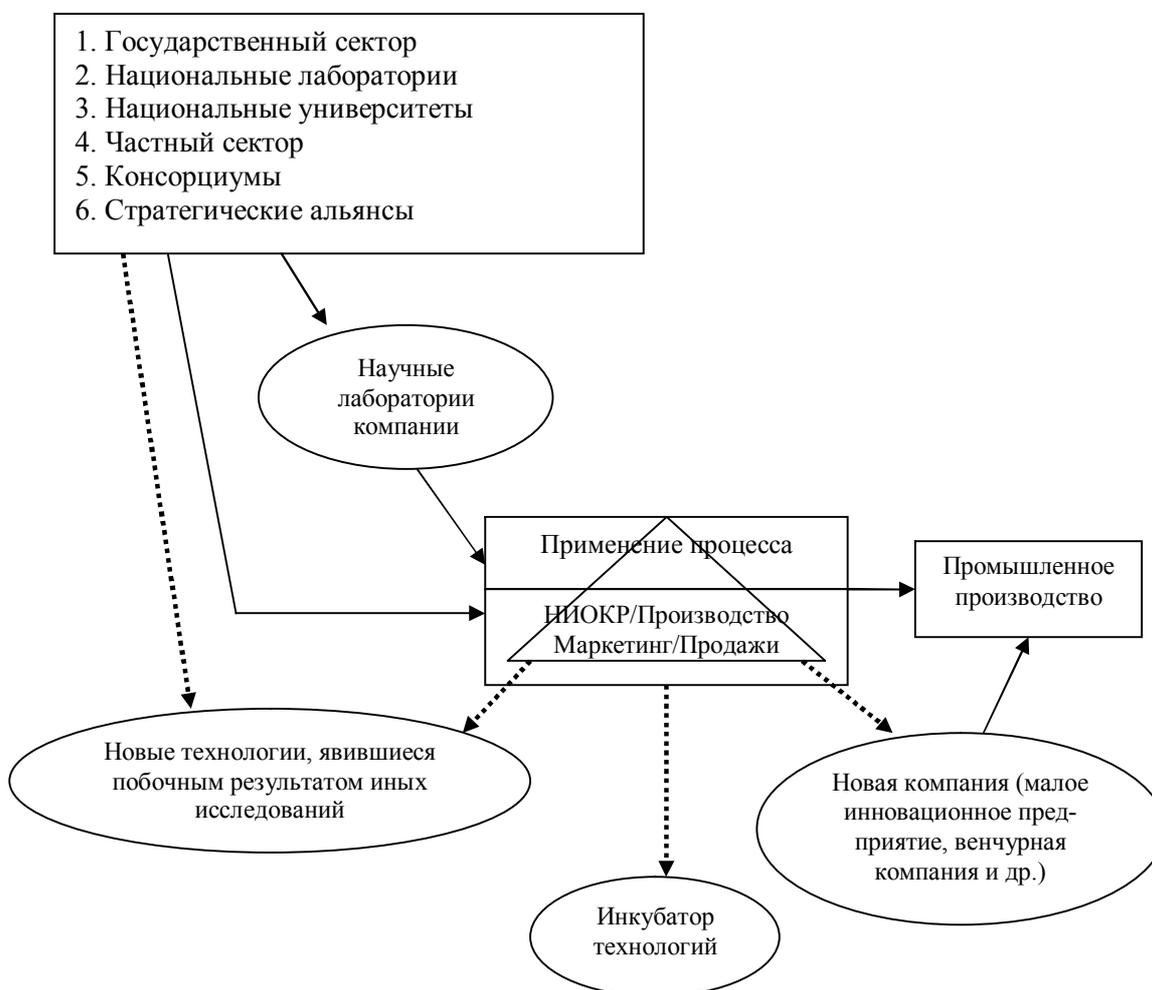


Рис. 2. Основные каналы трансферта технологий до стадии промышленного использования



Если обратиться к опыту развитых стран, то можно отметить, что в течение последних десятилетий в США была сформирована Национальная сеть передачи технологий, в задачи которой входит обеспечение доступа промышленных корпораций, малых фирм частного сектора к федеральным научно-техническим и технологическим ресурсам (табл. 1). Это связано, по мнению американских специалистов⁶, с двумя стратегическими тенденциями: обострением конкуренции в области высоких технологий со стороны иностранных компаний, стремлением повысить эффективность использования в коммерческих целях научных и технологических результатов, полученных в ходе освоения федеральных бюджетных ассигнований на НИОКР (в 2006 г. они превысили 3, 2 млрд. долл.)⁷.

Таблица 1

Законодательные нормы системы трансфера технологий в США

Элемент системы трансфера (законодательная норма)	Контекст
1	2
1. Распределение прав на федеральную собственность	Юридические и экономические механизмы обеспечения производства новыми перспективными идеями и технологиями, которые возникают в процессе выполнения финансируемых из госбюджета научных исследований и разработок между государством, крупными промышленными компаниями и малыми инновационными фирмами частного сектора, высшими учебными заведениями (университетами) и бесприбыльными организациями
2. Принятие в 1980 г. Закона Bayh-Dole, Act	Закон изменил статус, расширил возможности и направления деятельности Национального центра передачи технологий США, особенно для университетов, государственных структур и малых инновационных фирм в патентовании изобретений
3. Принятие в 1981 г. Закона Stevenson-Wydler, Act	Закон предоставил университетам, бесприбыльным организациям и фирмам малого бизнеса право передавать лицензии на коммерческое использование изобретений, сделанных в ходе исследований при финансовой поддержке правительства, любым промышленным компаниям
4. Принятие в 1982 г. Закона об инновационных исследованиях Small Business Innovation Research, Act	Закон создал специальную программу, обеспечивающую выделение всеми федеральными ведомствами с годовым бюджетом на НИОКР свыше 100 млн. долл. не менее 1,25% бюджета на проведение исследований и разработок силами малого бизнеса. Устанавливалась только обязательная нижняя граница ассигнований, верхняя не регламентировалась. За восемь лет (1983 – 1990 гг.) в программу включились 11 федеральных министерств и ведомств, которые рассмотрели почти 100 тыс. заявок от небольших наукоемких фирм и приняли к финансированию около 15 тыс. проектов
5. Принятие в 1984 г. Закона о кооперативных исследованиях Cooperative Research, Act	Закон вывел за рамки действия антитрестовского законодательства создание на доконкурентных стадиях НИОКР научно-исследовательских консорциумов с участием промышленных компаний и университетов
6. Принятие в 1986 г. Закона о технологических консорциумах Technology Corporate, Act	Образован Консорциум федеральных научных лабораторий для оказания помощи компаниям, особенно малым фирмам, в установлении контактов с компетентными федеральными научными подразделениями

⁶ T. Wohlers, RP/RT State of the Industry: 2007. Executive Summary; <http://www.mk-technology.com>

⁷ Science and Engineering Indicators – 2000. - volume I. Wash.: Gov. Print. off., 2000. P. 34.

* Составлено автором по Science and Engineering Indicators (transfer of market knowledge and technology) – 2005. Volume I. Wash.: Gov. Print. off., 2005. P. 126 - 151.

Окончание таблицы 1

1	2
7. Принятие двух нормативных актов: о передаче технологий Federal Technology Transfer Act, в 1986 г. и национальной конкурентоспособности National Competitiveness Act, в 1989 г.	Эти нормативные акты обеспечили промышленным компаниям правовые гарантии на использование интеллектуальной собственности, возникающей в результате соглашений о кооперативных исследованиях с федеральными лабораториями, и дали последним право на роялти от практического применения их изобретений, созданных в рамках подобных соглашений
8. В Патентный закон США в 1991 г. были введены дополнения к гл. 18, касающиеся прав на использование изобретений, созданных при содействии государства	Поощрение максимального участия университетов в патентовании своих результатов научно-технической деятельности, поощрение максимального участия малых предприятий в научно-технической деятельности, финансируемой государством, и обеспечение производителями такого объема прав на изобретение, созданное при поддержке государства, который достаточен для удовлетворения его нужд и для принятия мер против использования или ненадлежащего использования изобретений

До 1980 г. результаты НИОКР, финансируемых из госбюджета, являлись федеральной собственностью. Это не создавало у ученых и инженеров, работающих в государственных лабораториях или получающих финансовую поддержку от государства, особой заинтересованности в коммерческом применении полученных знаний. Обострение конкуренции на мировом рынке и ухудшение торгового баланса страны заставили конгресс пойти на изменение действующего законодательства и принять ряд новых федеральных законов.

Следует отметить, что с конца 60-х годов и до 1986 г. в США наблюдался спад числа выдаваемых патентов⁸, но затем обозначился их рост, что связано с указанным изменением возможностей для негосударственных структур в патентовании изобретений. В США некоторые корпорации вкладывают в исследования суммы, сопоставимые с общегосударственными: General Motors выделяет бюджет в объеме 2,5 млрд. долл. в год, IBM – 3 млрд.⁹

Созданная инфраструктура передачи технологий из государственного сектора в промышленность способствует увеличению количества поданных заявок на изобретения с участием федеральных лабораторий и национальных университетов, росту количества выданных на них патентов и повышению расходов частного сектора на поддержку научных исследований в университетах. Необходимо отметить, что резкий рост операций с интеллектуальной собственностью в США стал возможен только после того, как работникам университетов разрешили патентовать в частном порядке даже те изобретения, которые профинансированы за счет государственных грантов¹⁰. Фундаментальная наука имеет колоссальное количество прикладных последствий и бизнесу позволили этими последствиями пользоваться.

В задачи Национального центра передачи технологий США входит обеспечение доступа промышленных фирм к федеральным научно-техническим и технологическим ресурсам. С этой целью организована система ознакомления с механизмами передачи технологий и обучение их практическому применению. Все это способствовало заметной активизации деятельности по передаче технологий на всех уровнях. Позитивные результаты выразились в увеличении поданных заявок на изобретения с участием федеральных лабораторий, росте количества выданных на них патентов и повышении расходов частного сектора на поддержку научных исследований в университетах.

Определенный интерес может представлять опыт стран, в которых эксперты отмечают значительный прогресс в области развития трансферта технологий и ком-

⁸ Шерер Ф.М., Росс Д. Структура отраслевых рынков технологий. М.: ИНФРА-М, 2007. С. 27.

⁹ Creating Entrepreneurial Universities: Organizational Pathways of Transformation. Oxford, New-York: International Association of Universities and Pergamon, 2008. P. 167-169.

¹⁰ Science and Engineering Indicators – 2007. Volume I. Wash.: Gov. Print. off., 2007. P. 276.



мерциализации результатов научных исследований и разработок (Англия, Германия, Япония, Китай), где на базе крупнейших университетов созданы организационные структуры трансфера знаний, разработок и технологий с участием государственного и частного капиталов (табл. 2).

Таблица 2

Особенности системы трансфера технологий в развитых странах

Страна	Система трансфера	Контекст
Англия	Интегрированные службы трансфера технологий на базе крупных университетов и исследовательских центров Соединенного Королевства	Стимулирование развития интегрированных служб трансфера технологий происходит путем направления правительственных средств напрямую в университеты для осуществления так называемой деятельности «третьего потока» через «Фонд инноваций высшего образования в Англии» и «Грант по передаче знаний». В прошлом эти фонды развития распределялись посредством проведения тендеров, в настоящее время – посредством «формульного финансирования». В основе этого способа, главным образом, лежит потенциальная коммерческая привлекательность научных исследований. ЦТТ, представленные подразделениями по охране и коммерциализации технологий, функционирующими при крупных университетах и исследовательских центрах, действуют более 20 лет
Германия	Научные инженерные общества, совместные исследовательские ассоциации и стратегические альянсы в промышленности	Функции технологических посредников между лабораториями и компаниями выполняют различные научные общества и совместные исследовательские ассоциации в промышленности. Ведущая организационная роль принадлежит "Фраунгоферовскому обществу", в которое после объединения Германии вошли 45 исследовательских институтов, в том числе 9 из бывшей ГДР. Для облегчения малым фирмам доступа к его услугам правительство предоставляет им субсидии в размере до 40 % полной стоимости заказываемых НИОКР
Япония	Центры трансфера технологий при университетах, частные структуры в крупных компаниях и государственные офисы по лицензированию технологий	Правительство финансирует 27 ЦТТ (согласно терминологии Японии – ninteі TLO), которые обеспечивают изобретателей университетов услугами по управлению и коммерциализации результатов исследований и разработок: две трети средств – на эксплуатационные расходы (без возмещения расходов, связанных с услугами патентного поверенного и выплат патентных пошлин) в пределах суммы, эквивалентной 300 тыс. долл. США в год, сроком на пять лет, и одна треть – на НИОКР
Китай	Центры трансфера технологий при университетах, госструктуры по охране и защите интеллектуальной собственности в провинциях и столицах	Каждый крупный научно-исследовательский университет имеет структуру по трансферу технологий, изначально финансируемую Правительством КНР из доли от общих средств, выделенных университету Правительством. Эта модель финансирования ЦТТ меняется. В настоящее время большинство ЦТТ работают как ассоциированные частные компании, владельцами которых являются исключительно университеты

* Составлено автором по The Jurnal of Innovation in National Politics. Analysis of Elements Transfers and Engineering Indicators. Vol. 83. September, London, 2006. P.28 – 33.

Российский рынок услуг в области трансфера технологий только начинает формироваться, поскольку не существует какого-либо официального законодательного акта в отношении организации и финансирования. Действующих на российском инновационном рынке посредников условно можно разделить на две категории: консалтинговые фирмы и фонды. Первые специализируются на оказании консультационных услуг по вопросам охраны интеллектуальной собственности и маркетинга, вторые – на поиске заказчика-инвестора. При этом фонды не обязательно располагают или распоряжаются финансовыми ресурсами. Основной особенностью посредниче-

ских компаний и организаций в области интеллектуальной собственности и трансфера технологий в России является предоставление только консалтинговых услуг¹¹. Они не берут на себя расходы и ответственность по управлению интеллектуальной собственностью и не отвечают за конечный результат трансфера технологии.

Анализ деятельности посредников на российском рынке услуг в области трансфера и коммерциализации технологий показывает¹², что этими посредниками являются, в основном, консалтинговые компании и некоммерческие фонды, ведущие поиск разработчиков для решения технологических задач инвестора и поиск инвесторов, заинтересованных в покупке или доработке технологических решений разработчика, а также обеспечивающие юридическую, организационную и реже финансовую помощь разработчикам проектов.

Низкий уровень трансфера технологий напрямую связан с низким уровнем, а фактически с отсутствием, системы коммерциализации результатов научно-технической деятельности вузов и системы вовлечения в хозяйственный оборот результатов интеллектуальной деятельности. Основные причины низкой эффективности трансфера технологий в России заключаются в следующем:

- университеты и научные организации не включены в организационные механизмы формирующегося российского рынка трансфера технологий;
- ученые и разработчики не представляют, как полученные ими научные результаты могут быть трансформированы в рыночный продукт;
- экономические агенты рынка (менеджеры компаний, посредники и др.) практически не знакомы с характером современной науки, структурой и важнейшими направлениями ее деятельности, прорывными достижениями; не всегда могут судить о достоверности научных результатов и их технологической эффективности;
- государство, призванное устанавливать правила трансфера технологий, не является активным участником данного рынка и не стремится к разработке современных законодательных положений для производства научного знания и его коммерциализации;
- патентная защита производимой продукции не гарантирует добросовестной конкуренции, а предприятия и посредники не располагают средствами для финансирования своих новых разработок.

Новая роль и общественная миссия университета в современных условиях находит выражение в его возрастающем значении не только как центра передовой науки и высшего профессионального образования, но и как одного из важнейших субъектов национальной инновационной системы и инновационной деятельности, ориентированной на коммерциализацию результатов научно-технической деятельности путем вывода на рынок наукоемкой конкурентоспособной продукции. Поэтому в современных условиях для повышения эффективности функционирования инновационного потенциала Российской Федерации, его нормативно-правовой охраны, коммерческого использования результатов научно-технической деятельности, вовлечения в хозяйственный оборот объектов интеллектуальной собственности, организации патентно-информационного обеспечения изобретательской и патентно-лицензионной работы необходимо и стратегически важно создание Центров трансфера знаний, разработок и технологий на базе крупнейших вузов Российской Федерации.

¹¹ Интернет-ресурс: сайт Агентства финансовой информации «Консультант», <http://www.dinform.ru/newanalit>

¹² Роль Центров трансфера технологий в повышении инновационной активности регионов / Материалы конференции, 24 – 25 ноября 2005 г. Воронеж-Москва: ЗАО «Воронежский инновационно-технологический центр. 136 с.



Целью создания таких Центров должна быть реализация государственной инновационной политики в части формирования инновационного потенциала, развития рынков знаний и технологий, механизмов трансфера технологий, охраны интеллектуальной собственности и вовлечения ее в хозяйственный оборот, создания научно-технической, защищенной патентами, продукции субъектами инновационной деятельности.

Стратегическими задачами Центров должны стать: создание вертикально интегрированных стратегических альянсов между высшими учебными заведениями (университетами), государством, крупными промышленными компаниями, малыми инновационными фирмами частного сектора и бесприбыльными организациями; привлечение заказов от промышленности и бизнеса, в том числе и путем демонстрации успешных результатов (технологий, прототипов, моделей, аналогов и т.п.); коммерциализация РНТД и РИД; использование дохода от коммерциализации для финансирования научных работ и инновационной инфраструктуры; привлечение инвестиций в инфраструктуру.

Сложность выхода университетов на рынок наукоёмкой конкурентоспособной продукции заключается в неопределенности правового режима новшеств в действующем законодательстве. Принятый 28 июля 2009 г. ФЗ «О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам создания бюджетными учреждениями науки и образования хозяйственных обществ в целях практического применения (внедрения) результатов интеллектуальной деятельности»¹³ переводит в реальную плоскость механизмы развития системы трансфера знаний и технологий, инновационного потенциала высшей школы России, повышения конкурентоспособности национальной экономики и выхода университетов на международные рынки наукоёмкой продукции. Данным законом вносятся поправки¹⁴, согласно которым за учреждениями науки и образования, в том числе учреждениями государственных академий наук, закрепляется право создавать хозяйственные общества, внося в их уставные капиталы права на использование результатов интеллектуальной деятельности без согласия собственника имущества.

В заключение отметим, что в настоящее время наше государство не располагает достаточными ресурсами для введения в хозяйственный оборот созданной на его средства интеллектуальной продукции, а следовательно, необходимо искать иные, нежели прямое закрепление прав за государством, пути решения этой проблемы. Университеты должны активно участвовать в создании, становлении и формировании системы трансфера знаний, разработок и технологий в Российской Федерации.

¹³ Федеральный Закон от 28 августа 2009 г. № 137 – ФЗ «О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам создания бюджетными учреждениями науки и образования хозяйственных обществ в целях практического применения (внедрения) результатов интеллектуальной деятельности»

¹⁴ Закон предусматривает внесение изменений в следующие законодательные акты Российской Федерации: в Федеральный закон от 10 июля 1992 г. №3266-1 «Об образовании»; в Федеральный закон от 23 августа 1996 г. №127-ФЗ «О науке и государственной научно-технической политике»; в Федеральный закон от 22 августа 1996 г. №125-ФЗ «О высшем и послевузовском профессиональном образовании», соответствующим научным и образовательным организациям; в Федеральный закон от 24 июля 2007 г. №209-ФЗ «О развитии малого и среднего предпринимательства в Российской Федерации»



ON THE NECESSITY OF FORMATION AND DEVELOPMENT OF SYSTEM OF KNOWLEDGE AND TECHNOLOGIES TRANSFER ON THE BASIS OF INNOVATIVE POTENTIAL OF HIGHER EDUCATION IN RUSSIA

M. V. VLADYKA

Belgorod State University

e-mail: Yladyka@bsu.edu.ru

The development of the innovative capacity of universities in Russia due to the necessity of commercialization results of scientific technical activities and the foreground there are formation and development problems of innovation potential of universities, transfer market knowledge and technology, i.e. the transition from the subjective position of the object in that part of the innovation system, which is the innovation potential and the need to implement it. Addressing the commercialization of scientific and technological activities of universities, their degree is elaborated and implementation not only satisfies, but does not correspond to the present state of the innovative capacity of higher education in Russia and the national economy.

Key words: innovative potential, Graduate School, the results of scientific and technical activities of universities, commercialization, transfer, market knowledge and technology.

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

УДК 519.7

ИССЛЕДОВАНИЕ ОДНОГО ПАРАМЕТРА БУЛЕВЫХ ФУНКЦИЙ, БЛИЗКОГО К НЕЛИНЕЙНОСТИ*

В. Б. Алексеев
Р. Р. Омаров

*Московский
государственный
университет
им. М.В. Ломоносова*

*e-mail:
vbalekseev@rambler.ru*

*e-mail:
rustamomarov@ya.ru*

Минимальное расстояние (по Хэммингу) от булевой функции f до аффинных булевых функций называют нелинейностью функции f . Это один из параметров, характеризующих качество криптографических систем, использующих функцию f . В работе рассматривается один из классов булевых функций от 2^n переменных, имеющих максимальную нелинейность $2^{2^{n-1}} - 2^{n-1}$, а именно, класс Мэйорана–Мак-Фарланда. Исследуется, как меняется расстояние от функций этого класса до класса приближающих функций, если в класс приближающих функций кроме аффинных включить все функции, у которых в полиноме Жегалкина имеется одно нелинейное слагаемое. Показано, что новое расстояние может быть различным для разных функций из класса Мэйорана–Мак-Фарланда и изменяется в пределах от $2^{2^{n-1}} - 3 \cdot 2^{n-1} + 2$ до $2^{2^{n-1}} - 2 \cdot 2^{n-1}$, причем обе границы достижимы.

Ключевые слова: булева функция, криптографические свойства булевых функций, нелинейность, класс Мэйорана–Мак-Фарланда.

Введение. Постановка задачи

Булевы функции широко применяются в криптографии. При этом стойкость систем шифрования часто основывается на «сложности» используемых булевых функций. Поскольку аффинные функции считаются очень простыми, то в качестве одной из характеристик «сложности» булевых функций рассматривается «удаленность» данной функции от всех аффинных функций. Этому параметру, называемому нелинейностью булевой функции, посвящено множество работ. Обзор имеющихся результатов о нелинейности (с указанием имеющихся публикаций) можно найти в книге [1], где нелинейности посвящена отдельная глава. Дадим необходимые определения.

Пусть n – произвольное натуральное число. Через V_n будем обозначать векторное пространство наборов длины n с компонентами из $\{0,1\}$ с операцией \oplus покомпонентного сложения векторов по модулю 2.

* Работа выполнена при поддержке РФФИ, гранты 07-01-00154 и 09-01-00701.

Определение. Пусть f – булева функция от n переменных, то есть $f: V_n \rightarrow \{0,1\}$. Весом $wt(f)$ булевой функции f называется количество наборов, на которых функция f равна 1.

Определение. Пусть f, g – булевы функции от n переменных. Расстоянием от булевой функции f до булевой функции g называется величина $dist(f, g) = wt(f \oplus g)$. Таким образом, $dist(f, g)$ – это число наборов, на которых f и g принимают разные значения.

Определение. Пусть f – булева функция от n переменных и M – произвольное множество булевых функций от n переменных. Расстоянием от f до множества M называется величина $dist(f, M) = \min_{g \in M} dist(f, g)$.

Определение. Пусть $x \in V_n, y \in V_n$. Через $\langle x, y \rangle$ будем обозначать скалярное произведение x и y : $\langle x, y \rangle = x_1 y_1 \oplus \dots \oplus x_n y_n$ (здесь \oplus – это сложение по модулю 2).

Определение. Булева функция $g(x)$ от n переменных называется аффинной, если существуют $a = (a_1, \dots, a_n) \in V_n$ и $c \in \{0,1\}$ такие, что $g(x) = \langle a, x \rangle \oplus c = a_1 x_1 \oplus \dots \oplus a_n x_n \oplus c$. Множество всех аффинных булевых функций от n переменных будем обозначать A_n . Отметим, что для любой аффинной булевой функции $g(x)$ от n переменных, отличной от константы, $wt(g(x)) = 2^{n-1}$.

Определение. Расстояние $dist(f, A_n)$ от булевой функции $f(x)$ от n переменных до множества A_n аффинных булевых функций называется нелинейностью функции $f(x)$ и обозначается через N_f .

Лемма 1 [1]. Для любой булевой функции $f(x)$ от n переменных справедливо неравенство $N_f \leq 2^{n-1} - 2^{n/2-1}$. Для четных n эта оценка достижима, то есть существуют функции $f(x)$ от $2n$ переменных, для которых $N_f = 2^{n-1} - 2^{n/2-1}$.

Определение. Булевы функции $f(x)$ от $2n$ переменных, для которых $N_f = 2^{2n-1} - 2^{n-1}$, называют максимально-нелинейными функциями (этот класс называют также классом бент-функций).

Таким образом, максимально-нелинейные функции – это булевы функции от $2n$ переменных, наиболее плохо приближаемые аффинными функциями. Мы рассмотрим вопрос о том, что происходит с максимально-нелинейными функциями, если класс аффинных функций несколько расширяется, а именно, в качестве приближающих рассматриваются все функции, у которых в полиноме Жегалкина имеется не более одного нелинейного слагаемого.

Определение. Через AE_n будем обозначать класс всех почти аффинных функций $g(x)$, а именно, функций вида $g(x) = \langle a, x \rangle \oplus c \oplus x_{i_1} \cdot \dots \cdot x_{i_k}$, где $a \in V_n, c \in \{0,1\}$ и $\{i_1, \dots, i_k\}$ – произвольное подмножество (возможно, пустое) множества $\{1, \dots, n\}$.

В данной работе нас интересует вопрос: одинакова ли величина $dist(f, AE_{2n})$ для всех максимально-нелинейных функций f от $2n$ переменных и насколько для них $dist(f, AE_{2n})$ отличается от $dist(f, A_{2n})$? На первый вопрос мы дадим отрицательный ответ. На второй вопрос мы дадим полный ответ для одного известного дос-



таточно широкого класса максимально-нелинейных функций (множество всех максимально-нелинейных функций f от $2n$ переменных пока не описано).

Определение. Пусть $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$. Класс Мэйорана–Мак-Фарланда определяется как класс всех булевых функций $f(x, y)$ от $2n$ переменных вида $f(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y)$, где π – произвольная подстановка на множестве V_n , а $\Phi(y)$ – произвольная булева функция от n переменных.

Известно, что все функции из класса Мэйорана–Мак-Фарланда являются максимально-нелинейными функциями [1]. В данной работе мы устанавливаем точные пределы, в которых изменяется значение $dist(f, AE_{2n})$ для всех функций из класса Мэйорана–Мак-Фарланда.

Оценки нового параметра для функций класса Мэйорана–Мак-Фарланда

Покажем сразу же, что для любой функции $f(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y)$ из класса Мэйорана–Мак-Фарланда $dist(f, AE_{2n}) < dist(f, A_{2n})$.

Лемма 2. Для любой функции $f(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y)$ из класса Мэйорана–Мак-Фарланда справедливо неравенство:

$$dist(f(x, y), AE_{2n}) \leq 2^{2n-1} - 2 \cdot 2^{n-1}.$$

Доказательство. Пусть $\pi(0, \dots, 0) = a = (a_1, \dots, a_n)$, $a' = a \oplus (1, 0, \dots, 0)$ и $\gamma = (\gamma_1, \dots, \gamma_n) = \pi^{-1}(a')$. Так как $\gamma \neq (0, \dots, 0)$, то существует j такое, что $\gamma_j = 1$. Рассмотрим функцию $g(x, y) = \langle a, x \rangle \oplus \langle b, y \rangle \oplus c \oplus x_1 y_j$ из класса AE_{2n} , где b пока произвольно и $c = \Phi(0, \dots, 0)$. Тогда $dist(f(x, y), g(x, y)) = wt(h(x, y))$, где

$$\begin{aligned} h(x, y) &= f(x, y) \oplus g(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y) \oplus \langle a, x \rangle \oplus \langle b, y \rangle \oplus c \oplus x_1 y_j = \\ &= \langle \pi(y) \oplus a, x \rangle \oplus \Phi(y) \oplus \langle b, y \rangle \oplus \Phi(0, \dots, 0) \oplus x_1 y_j. \end{aligned}$$

При $\beta = (0, \dots, 0)$ получаем $h(x, \beta) \equiv 0$. При $\beta = \gamma$ получаем

$$\begin{aligned} h(x, \beta) &= h(x, \gamma) = \langle \pi(\gamma) \oplus a, x \rangle \oplus \Phi(\gamma) \oplus \langle b, \gamma \rangle \oplus \Phi(0, \dots, 0) \oplus x_1 = \langle (1, 0, \dots, 0), x \rangle \oplus x_1 \oplus \\ &\oplus \Phi(\gamma) \oplus \Phi(0, \dots, 0) \oplus \langle b, \gamma \rangle = \Phi(\gamma) \oplus \langle b, \gamma \rangle \oplus \Phi(0, \dots, 0). \end{aligned}$$

Положим $b_j = \Phi(\gamma) \oplus \Phi(0, \dots, 0)$ и $b_i = 0$ при $i \neq j$. Тогда $h(x, \gamma) \equiv 0$. При $\beta \neq (0, \dots, 0)$ и $\beta \neq \gamma$ имеем $\pi(\beta) \oplus a \neq (0, \dots, 0)$ и $\pi(\beta) \oplus a \neq (1, 0, \dots, 0)$. Поэтому в этих случаях $h(x, \beta)$ является аффинной функцией, отличной от константы, откуда $wt(h(x, \beta)) = 2^{n-1}$. Таким образом,

$$dist(f(x, y), g(x, y)) = wt(h(x, y)) = \sum_{\beta \in V_n} wt(h(x, \beta)) = 2^{n-1}(2^{n-1} - 2) = 2^{2n-1} - 2 \cdot 2^{n-1},$$

откуда следует утверждение леммы 2.

Лемма 3. Если нелинейное слагаемое в $g(x, y) \in AE_{2n}$ не содержит переменных x_1, \dots, x_n , то $dist(f(x, y), g(x, y)) \geq 2^{2n-1} - 2^{n-1}$ для любой функции $f(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y)$ из класса Мэйорана–Мак-Фарланда.

Доказательство. Так как все функции из класса Мэйорана–Мак-Фарланда являются максимально-нелинейными, то $dist(f(x, y), g(x, y)) \geq 2^{2n-1} - 2^{n-1}$, если $g(x, y)$ не содержит нелинейных слагаемых. Пусть теперь $g(x, y)$ содержит нелинейное слагаемое $y_{j_1} \cdot \dots \cdot y_{j_r}$. Положим

$$f_1(x, y) = f(x, y) \oplus y_{j_1} \cdot \dots \cdot y_{j_r} = \langle \pi(y), x \rangle \oplus (\Phi(y) \oplus y_{j_1} \cdot \dots \cdot y_{j_r}) \text{ и}$$

$$g_1(x, y) = g(x, y) \oplus y_{j_1} \cdots y_{j_t}.$$

Тогда $f_1(x, y)$ лежит в классе Мэйорана–Мак–Фарланда и $g_1(x, y)$ – аффинная функция. Поэтому $\text{dist}(f(x, y), g(x, y)) = \text{dist}(f_1(x, y), g_1(x, y)) \geq 2^{2n-1} - 2^{n-1}$. Лемма доказана.

Следующая теорема дает явную формулу для вычисления $\text{dist}(f(x, y), AE_{2n})$ для любой функции $f(x, y)$ из класса Мэйорана–Мак–Фарланда и служит основой для получения дальнейших оценок.

Теорема 1. Для любой функции $f(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y)$ из класса Мэйорана–Мак–Фарланда от $2n$ переменных выполняется равенство:

$$\text{dist}(f(x, y), AE_{2n}) = 2^{2n-1} - 2^{n-1} - \max_{I, J, \beta', b} \left(2^{n-k} \cdot \sum_{\beta \in P \cap Q} \text{sg}(\langle \beta \oplus \beta', b \rangle \oplus \Phi(\beta) \oplus \Phi(\beta') \oplus \langle \pi(\beta) \oplus \pi(\beta'), (1, 1, \dots, 1) \rangle) \right),$$

где $I = \{i_1, \dots, i_k\} \subseteq \{1, \dots, n\}$, $k = |I| \geq 1$, $J = \{j_1, \dots, j_t\} \subseteq \{1, \dots, n\}$, $\beta' \in V_n$, $b \in V_n$, $P = \{\beta \in V_n \mid (\pi(\beta))_s = (\pi(\beta'))_s \text{ при всех } s \notin I\}$, $Q = \{\beta \in V_n \mid \beta_{j_1} = \dots = \beta_{j_t} = 1\}$, $\text{sg}(0) = -1$, $\text{sg}(1) = +1$ (если $P \cap Q = \emptyset$, то соответствующая сумма считается равной 0).

Доказательство. Заметим вначале, что, если $f(x_1, \dots, x_n) \equiv c$, где c – константа ($c \in \{0, 1\}$), то $\text{wt}(f(x_1, \dots, x_n)) = 2^{n-1} + 2^{n-1} \text{sg}(c)$, где $\text{sg}(0) = -1$, $\text{sg}(1) = +1$. Это представление будет часто использоваться в дальнейшем.

Пусть $f(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y)$ и пусть $g(x, y) = \langle a, x \rangle \oplus \langle b, y \rangle \oplus c \oplus x_{i_1} \cdots x_{i_k} \cdot y_{j_1} \cdots y_{j_t}$ – произвольная функция из класса AE_{2n} , где $a \in V_n$, $b \in V_n$, $c \in \{0, 1\}$, $I = \{i_1, \dots, i_k\} \subseteq \{1, \dots, n\}$, $J = \{j_1, \dots, j_t\} \subseteq \{1, \dots, n\}$. С учетом лемм 2 и 3, будем считать, что $k = |I| \geq 1$. Выражение $x_{i_1} \cdots x_{i_k} \cdot y_{j_1} \cdots y_{j_t}$ будем сокращенно записывать как $x_I y_J$, при этом положим $x_I y_\emptyset = x_I$. Положим

$$h(x, y) = f(x, y) \oplus g(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y) \oplus \langle a, x \rangle \oplus \langle b, y \rangle \oplus c \oplus x_I y_J = \langle \pi(y) \oplus a, x \rangle \oplus \Phi(y) \oplus \langle b, y \rangle \oplus c \oplus x_I y_J.$$

Тогда $\text{dist}(f(x, y), g(x, y)) = \text{wt}(h(x, y))$ и наша задача – исследовать $\text{wt}(h(x, y))$.

Лемма 4. Пусть $\beta' = \pi^{-1}(a)$, $k = |I| \geq 1$, $P = \{\beta \in V_n \mid (\pi(\beta))_s = (\pi(\beta'))_s \text{ при всех } s \notin I\}$, $Q = \{\beta \in V_n \mid \beta_{j_1} = \dots = \beta_{j_t} = 1\}$ (если $J = \{j_1, \dots, j_t\}$ пусто, то $Q = V_n$). Тогда

$$\text{wt}(h(x, y)) = 2^{2n-1} + 2^{n-1} \text{sg}(\langle \beta', b \rangle \oplus \Phi(\beta') \oplus c) - 2^{n-k} \cdot \sum_{\beta \in P \cap Q} \text{sg}(\langle \beta, b \rangle \oplus \Phi(\beta) \oplus c \oplus \langle \pi(\beta) \oplus a, (1, 1, \dots, 1) \rangle).$$

Доказательство. Пусть

$$h_1(x, y) = \langle \pi(y) \oplus a, x \rangle \oplus \Phi(y) \oplus \langle b, y \rangle \oplus c, \quad h_2(x, y) = x_I y_J.$$

Тогда $h(x, y) = h_1(x, y) \oplus h_2(x, y)$. При $\beta \neq \beta'$ имеем $\pi(\beta) \neq a$ и функция $h_1(x, \beta)$ является аффинной функцией, отличной от константы, откуда $\text{wt}(h_1(x, \beta)) = 2^{n-1}$. Если $\beta = \beta'$, то $\pi(\beta') \oplus a = 0$, $h_1(x, \beta') = \Phi(\beta') \oplus \langle b, \beta' \rangle \oplus c$ (константа) и $\text{wt}(h_1(x, \beta')) = 2^{n-1} + 2^{n-1} \text{sg}(\Phi(\beta') \oplus \langle b, \beta' \rangle \oplus c)$. Отсюда



$$wt(h_1(x, y)) = \sum_{\beta \in V_n} wt(h_1(x, \beta)) = 2^{n-1} \cdot 2^n + 2^{n-1} sg(\Phi(\beta') \oplus \langle b, \beta' \rangle \oplus c) = 2^{2n-1} + 2^{n-1} sg(\Phi(\beta') \oplus \langle b, \beta' \rangle \oplus c).$$

$$\text{Легко видеть, что } wt(h_2(x, y)) = 2^{n-k} \cdot 2^{n-t} = 2^{2n-k-t}.$$

Исследуем теперь вес функции $h_1(x, y) \cdot h_2(x, y) = h_1(x, y) \cdot x_I y_J$. Имеем

$$wt(h_1(x, y) \cdot h_2(x, y)) = \sum_{\beta \in Q} wt(h_1(x, \beta) \cdot x_I) = \sum_{\beta \in Q} wt(\langle \pi(\beta) \oplus a, x \rangle \oplus \Phi(\beta) \oplus \langle b, \beta \rangle \oplus c) \cdot x_I).$$

Пусть x' – набор всех переменных x , кроме x_{i_1}, \dots, x_{i_k} , и пусть $h_3(x', \beta)$ – функция от $n-k$ переменных, получающаяся из $h_1(x, \beta)$ при подстановке $x_{i_1} = \dots = x_{i_k} = 1$. Тогда $wt(h_1(x, \beta) \cdot x_I) = wt(h_3(x', \beta))$. При этом, если хотя бы для одной координаты $s \notin I$ выполняется $(\pi(\beta))_s \neq a_s$, то функция $h_3(x', \beta)$ – аффинная функция, отличная от константы, и поэтому ее вес равен 2^{n-k-1} . Если же $(\pi(\beta))_s = a_s$ при всех $s \notin I$, то $h_3(x', \beta)$ – константа, равная $\langle \pi(\beta) \oplus a, (1, 1, \dots, 1) \rangle \oplus \Phi(\beta) \oplus \langle b, \beta \rangle \oplus c$, и ее вес равен $2^{n-k-1} + 2^{n-k-1} sg(\langle \pi(\beta) \oplus a, (1, 1, \dots, 1) \rangle \oplus \Phi(\beta) \oplus \langle b, \beta \rangle \oplus c)$.

Таким образом,

$$wt(h_1(x, y) \cdot h_2(x, y)) = \sum_{\beta \in Q} wt(h_3(x', \beta)) = 2^{n-t} \cdot 2^{n-k-1} + \sum_{\beta \in P \cap Q} 2^{n-k-1} sg(\langle \pi(\beta) \oplus a, (1, 1, \dots, 1) \rangle \oplus \Phi(\beta) \oplus \langle b, \beta \rangle \oplus c).$$

Подставляя полученные формулы для $wt(h_1(x, y))$, $wt(h_2(x, y))$ и $wt(h_1(x, y) \cdot h_2(x, y))$ в равенство

$$wt(h(x, y)) = wt(h_1(x, y) \oplus h_2(x, y)) = wt(h_1(x, y)) + wt(h_2(x, y)) - 2 \cdot wt(h_1(x, y) \cdot h_2(x, y)),$$

получим утверждение леммы 4.

Лемма 5. Пусть $\beta' = \pi^{-1}(a)$, $\langle \beta', b \rangle \oplus \Phi(\beta') \oplus c = 1$. Тогда $wt(h(x, y)) \geq 2^{2n-1} - 2^{n-1}$.

Доказательство. При $I = \emptyset$ утверждение следует из леммы 3. При $I \neq \emptyset$ воспользуемся леммой 4. Так как в лемме 4 имеем $|P| = 2^k$, то слагаемых под знаком суммирования в лемме 4 не больше 2^k . Тогда при условиях леммы 5:

$$wt(h(x, y)) \geq 2^{2n-1} + 2^{n-1} - 2^{n-k} 2^k = 2^{2n-1} - 2^{n-1}.$$

Лемма доказана.

Из лемм 2, 3 и 5 вытекает, что для вычисления $dist(f(x, y), AE_{2n})$ для любой функции $f(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y)$ из класса Мэйорана–Мак–Фарланда достаточно рассматривать $dist(f(x, y), g(x, y))$ только для таких функций $g(x, y) = \langle a, x \rangle \oplus \langle b, y \rangle \oplus c \oplus x_I y_J$, в которых $I \neq \emptyset$ и $c = \langle \beta', b \rangle \oplus \Phi(\beta')$, где $\beta' = \pi^{-1}(a)$.

По лемме 4 в таких случаях получим:

$$wt(h(x, y)) = 2^{2n-1} - 2^{n-1} - 2^{n-k} \cdot \sum_{\beta \in P \cap Q} sg(\langle \beta \oplus \beta', b \rangle \oplus \Phi(\beta) \oplus \Phi(\beta') \oplus \langle \pi(\beta) \oplus \pi(\beta'), (1, 1, \dots, 1) \rangle).$$

Рассматривая минимум $wt(h(x, y))$ по всем таким функциям $h(x, y)$, получаем утверждение теоремы 1. Теорема 1 доказана.

Теорема 2. Для всех функций $f(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y)$ из класса Мэйорана–Мак–Фарланда от $2n$ переменных при всех $n \geq 2$ выполняются неравенства:

$$2^{2n-1} - 3 \cdot 2^{n-1} + 2 \leq dist(f, AE_{2n}) \leq 2^{2n-1} - 2 \cdot 2^{n-1},$$

причем обе границы достижимы. (При $n = 1$ $dist(f, AE_{2n}) = 0$ для всех f).

Доказательство. Верхняя оценка доказана в лемме 2. Покажем, что эта оценка достижима. Рассмотрим функцию $f(x, y) = \langle y, x \rangle$. Она входит в класс Мэйорана–Мак-Фарланда (при $\pi(y) \equiv y$ и $\Phi(y) \equiv 0$), поэтому мы можем использовать для нее теорему 1. Выражение под суммой в теореме 1 для данной функции примет вид:

$$sg(\langle \beta \oplus \beta', b \rangle \oplus \langle \beta \oplus \beta', (1, 1, \dots, 1) \rangle) = sg(\langle \beta \oplus \beta', b \oplus (1, 1, \dots, 1) \rangle).$$

При этом $I \neq \emptyset$ и $P = \{\beta \in V_n \mid (\beta)_s = (\beta')_s \text{ при всех } s \notin I\}$. Пусть $I = \{i_1, \dots, i_k\}$, $k = |I| \geq 1$. Пусть b и β' фиксированы. Тогда при β , пробегающем множество P , выражение $\langle \beta \oplus \beta', b \oplus (1, 1, \dots, 1) \rangle$ можно рассматривать как аффинную функцию от переменных $\beta_{i_1}, \dots, \beta_{i_k}$ без свободного члена. Такая функция принимает значение 1 не более 2^{k-1} раз. Поэтому для функции $f(x, y) = \langle y, x \rangle$ при любых фиксированных I, J, β', b сумма в теореме 1 не превосходит 2^{k-1} и, следовательно, максимум не превосходит 2^{n-1} . Тогда $dist(\langle x, y \rangle, AE_{2n}) \geq 2^{2n-1} - 2^{n-1} - 2^{n-1} = 2^{2n-1} - 2 \cdot 2^{n-1}$. С учетом леммы 2, получаем $dist(\langle x, y \rangle, AE_{2n}) = 2^{2n-1} - 2 \cdot 2^{n-1}$.

Докажем теперь нижнюю оценку. Посмотрим, какие значения может принимать сумма по множеству $P \cap Q$ в теореме 1. Так как $|P| = 2^k$, то число слагаемых в сумме не превосходит 2^k и, следовательно, сама сумма не превосходит 2^k , а выражение под максимумом не превосходит 2^n . Отсюда максимум не превосходит 2^n и теорема 1 дает неравенство $dist(f, AE_{2n}) \geq 2^{2n-1} - 3 \cdot 2^{n-1}$. Однако эта оценка недостижима. Заметим, что $\beta' \in P$ и при $\beta = \beta'$ слагаемое в сумме в теореме 1 равно $sg(0) = -1$. Поэтому, если $\beta' \in Q$, то $\beta' \in P \cap Q$ и сумма по множеству $P \cap Q$ в теореме 1 не превосходит $2^k - 2$, а соответствующее выражение под максимумом не превосходит $2^{n-k} \cdot (2^k - 2) = 2^n - 2 \cdot 2^{n-k} \leq 2^n - 2$. Пусть теперь $\beta' \notin Q$. Тогда рассматриваемая сумма содержит не более $2^k - 1$ слагаемых и, следовательно, не превосходит $2^k - 1$. Тогда соответствующее выражение под максимумом не превосходит $2^{n-k} \cdot (2^k - 1) = 2^n - 2^{n-k} \leq 2^n - 2$ при $k < n$. Остается рассмотреть случай $k = n$ и $\beta' \notin Q$. Тогда $Q \neq V_n$ и $t \geq 1$. Поэтому $|P \cap Q| \leq |Q| \leq 2^{n-1}$. Тогда рассматриваемая сумма и соответствующее выражение под максимумом не превосходят $2^{n-1} \leq 2^n - 2$ при $n \geq 2$. Таким образом, выражение под максимумом всегда не превосходит $2^n - 2$. Следовательно, и максимум не превосходит $2^n - 2$. Отсюда следует нижняя оценка в теореме 2. Докажем теперь, что она достижима.

Лемма 6. Для любой подстановки π на V_n существует функция $\Phi(y)$ от n переменных такая, что для функции $f(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y)$ из класса Мэйорана–Мак-Фарланда от $2n$ переменных выполняется равенство $dist(f, AE_{2n}) = 2^{2n-1} - 3 \cdot 2^{n-1} + 2$.

Доказательство. Для данной подстановки π определим функцию $\Phi(y)$ от n переменных следующим образом. Пусть $\Phi(0, \dots, 0)$ произвольно и $\Phi(y) = \Phi(0, \dots, 0) \oplus \langle \pi(y) \oplus \pi(0, \dots, 0), (1, 1, \dots, 1) \rangle \oplus 1$ при $y \neq (0, \dots, 0)$. Возьмем $I = \{1, \dots, n\}$, $J = \emptyset$, $\beta' = (0, \dots, 0)$, $b = (0, \dots, 0)$. Тогда $k = n$, $P = Q = V_n$. При этом в сумме в теореме 1 будет 2^n слагаемых вида $sg(\Phi(\beta) \oplus \Phi(0, \dots, 0) \oplus \langle \pi(\beta) \oplus \pi(0, \dots, 0), (1, 1, \dots, 1) \rangle)$. При $\beta = (0, \dots, 0)$ это слагаемое равно $sg(0) = -1$. При



остальных β , подставляя в это слагаемое $\Phi(\beta) = \Phi(0, \dots, 0) \oplus \langle \pi(\beta) \oplus \pi(0, \dots, 0), (1, 1, \dots, 1) \rangle \oplus 1$, получаем, что это слагаемое равно $\text{sg}(1) = 1$. Поэтому рассматриваемая сумма будет равна $2^n - 2$ и максимум будет не меньше $2^n - 2$. Поэтому для рассматриваемой функции $f(x, y) = \langle \pi(y), x \rangle \oplus \Phi(y)$ выполняется неравенство $\text{dist}(f, AE_{2n}) \leq 2^{2n-1} - 3 \cdot 2^{n-1} + 2$. С учетом доказанной нижней оценки в теореме 2 получаем, что для этой функции $\text{dist}(f, AE_{2n}) = 2^{2n-1} - 3 \cdot 2^{n-1} + 2$. Лемма 6, а вместе с ней и теорема 2 доказаны.

Заключение

Таким образом, максимально-нелинейные функции из класса Мэйорана–Мак-Фарланда имеют разную стойкость при приближении их почти аффинными функциями класса AE_{2n} . В работе установлены точные пределы, в которых изменяется значение $\text{dist}(f, AE_{2n})$ для всех функций из класса Мэйорана–Мак-Фарланда. Вопрос о точных границах изменения величины $\text{dist}(f, AE_{2n})$ для произвольных максимально-нелинейных функций f от $2n$ переменных требует дальнейшего изучения.

Литература

1. Логачёв, О.А. Булевы функции в теории кодирования и криптологии [Текст] / О.А. Логачёв, А.А. Сальников, В.В. Яценко. – М.: Изд-во МЦНМО, 2004. - 470 с.

INVESTIGATION OF A PARAMETER OF BOOLEAN FUNCTIONS CLOSED TO NONLINEARITY

V. B. Alekseev
R. R. Omarov

*Lomonosov Moscow
State University, Moscow*

*e-mail:
vbalekseev@rambler.ru*

*e-mail:
rustamomarov@ya.ru*

Minimal distance (by Hamming) between Boolean function f and all affine Boolean functions is said to be *nonlinearity* of function f . This is one of the parameters characterizing quality of cryptographic systems using this function. We consider one set of Boolean functions having maximal nonlinearity, namely the Maiorana–McFarland set. We investigate how the distance between these functions and the set of approximating functions changes when we include in the set of approximating functions not only affine functions but also all functions having one nonlinear term in their Zhegalkin polynomial form. We show that new distance can be different for different functions from the Maiorana–McFarland set and it can vary between $2^{2n-1} - 3 \cdot 2^{n-1} + 2$ and $2^{2n-1} - 2 \cdot 2^{n-1}$ and both bounds can be achieved.

Key words: Boolean function, cryptographic properties of Boolean functions, nonlinearity, Maiorana–McFarland set.

НОВЫЕ АРХИТЕКТУРЫ И АЛГОРИТМЫ ОБУЧЕНИЯ ДИСКРЕТНЫХ НЕЙРОННЫХ СЕТЕЙ АДАПТИВНОЙ РЕЗОНАНСНОЙ ТЕОРИИ

В. Д. ДМИТРИЕНКО
И. П. ХАВИНА
А. Ю. ЗАКОВОРОТНЫЙ

*Национальный
технический
университет
«Харьковский
политехнический
институт»*

e-mail: Arcade@i.ua

Проанализированы достоинства и недостатки архитектур и алгоритмов обучения дискретных нейронных сетей адаптивной резонансной теории (АРТ). Предложены новые архитектуры нейронных сетей АРТ и алгоритмы обучения сетей АРТ без адаптации весов связей распределенных распознающих нейронов.

Ключевые слова: дискретные нейронные сети адаптивной резонансной теории, алгоритмы обучения.

Постановка проблемы исследований

Системы управления, распознавания, классификации, прогнозирования, идентификации на основе искусственных нейронных сетей (НС) многими авторами рассматриваются как альтернатива классическим системам. В настоящее время известен целый ряд возможных архитектур для построения нейросетевых систем управления, а также систем управления, где нейрорегуляторы работают параллельно обычным регуляторам [1 – 4]. Большой интерес к нейросетевым системам управления связан с тем, что они имеют ряд полезных свойств, которых трудно добиться с помощью других методов. В частности, речь идет о робастности систем управления, об определении управлений в условиях, когда неточно известна модель объекта, более гибком реагировании на изменяющиеся внешние условия, возможности обработки данных различной природы. Можно также отметить, что для нейросетевых моделей не существует ограничений, связанных с нелинейностью системы, ее порядком или видом применяемого функционала.

Известен и целый ряд НС, пригодных для разработки высокоэффективных систем распознавания, классификации, прогнозирования, идентификации [3 – 6]. Большой интерес к нейросетевым системам распознавания, классификации и т.д. связан с тем, что они имеют ряд полезных свойств, которых часто трудно добиться с помощью других методов. В частности, речь идет об их универсальности, о возможности построения НС под задачу, о возможности обучения на примерах, о самообучении и обобщении, распознавании новой информации, об использовании ассоциаций, о распознавании и классификации зашумленных изображений, возможности обработки данных различной природы и т.д. [5 – 7]. В тоже время существуют и определенные проблемы при разработке систем на основе НС.

Для идентификации и управления в реальном времени динамическими объектами, содержащими в своих структурах "черные" и "серые" ящики применяются нейронные сети, основанные на аппроксимации нелинейностей, например, рядами Вольтера, Винера, полиномами Колмогорова-Габора, нелинейными разностными уравнениями. Однако в тех случаях, когда управление "черным" или "серым" ящиком должно осуществляться на распознавании ситуаций, которые определяются некоторым множеством динамических процессов, несмотря на большое число работ в этой области, до сих пор удовлетворительных систем управления разработать не удалось. Это связано с особенностью обучения НС, которая, как правило, требует наличия всей исходной информации, поскольку обучение новому образу методом обратного распространения ошибки или генетическими алгоритмами, или другими мето-



дами требует, как правило, полного переобучения сети. Этот же недостаток характерен и для других систем на основе НС: распознавания, классификации, прогнозирования и т.д. Все это требует совершенствования известных и разработки новых НС, позволяющих расширить область эффективного применения искусственных нейронных сетей.

Анализ последних исследований и публикаций

Невозможность с помощью уже известных НС [1 – 6] решить проблему восприятия новой информации при сохранении уже имеющейся (т.е. решить проблему стабильности – пластичности) привело к разработке принципиально новых конфигураций НС – сетей адаптивной резонансной теории (АРТ). Эти сети в определенной степени решают противоречивые задачи чувствительности (пластичности) к новым данным и сохранения полученной ранее информации (стабильности). Поэтому проблемы распознавания конкретных режимов динамических объектов и адаптация системы в процессе эксплуатации могут решаться на основе НС АРТ [5 – 9]. Однако известные дискретные нейронные сети АРТ имеют и существенные недостатки, если требуется запоминать сотни или тысячи реализаций одного и того же динамического процесса, поскольку матрица весов связей между S - и Y -нейронами размерностью $n \times m$ (n – число S -нейронов входного слоя; m – число распознающих Y -нейронов) хранит только одно изображение, полученное в результате пересечения запоминаемых динамических процессов. При малых значениях параметра сходства [7 – 11] в режиме распознавания память нейронной сети АРТ не имеет необходимого объема информации (наборов фрагментов изображений) для принятия обоснованных решений, а при больших значениях параметра аналогичные изображения, отличающиеся небольшим числом второстепенных деталей, запоминаются как прототипы разных классов изображений с помощью разных Y -нейронов. Это делает практическое использование этих НС сомнительным из-за чрезмерно большого числа Y -элементов распознающего слоя.

Другой недостаток сетей АРТ – связан с наличием только одного способа обучения НС путем адаптации весов связей распознающих нейронов, что может приводить к искажению или потере, хранящейся в памяти сети информации, нерациональной классификации входной информации. Еще одним недостатком сети АРТ-1 является отсутствие режимов обучения сети с учителем. Минусом НС АРТ является и невозможность получения нескольких решений, а также наличие только одного поля обрабатывающих нейронов.

Целью статьи является разработка новых архитектур и алгоритмов обучения дискретных нейронных сетей АРТ, позволяющих устранить существующие недостатки этих сетей и расширить область их эффективного применения.

Архитектура и алгоритмы функционирования сети АРТ-1

Архитектура нейронной сети АРТ-1 приведена на рис. 1 в прямоугольнике, нарисованном пунктирной линией. Эта сеть относит предъявленное изображение к одному из запомненных классов только в случае, если оно достаточно похоже на прототип этого класса, то есть на изображение, хранящееся в весах связей нейронной сети. Если такое свойство, установленное с помощью специального параметра сходства между двумя изображениями [7 – 11], имеет место в режиме обучения, то найденный прототип модифицируется, чтобы стать более похожим на предъявленное сети входное изображение. При отсутствии достаточного сходства между предъявленным изображением и прототипами всех имеющихся классов, оно запоминается сетью как прототип нового класса. Это возможно благодаря тому, что информация об изображениях-прототипах хранится локально, только в весах связей одного распознающего нейрона, и сеть имеет значительное число неиспользуемых распознающих нейронов, избыток которых уменьшается по мере поступления новой входной информации.

Наличие избыточных распознающих нейронов является принципиальной особенностью НС АРТ, поскольку при их отсутствии новые входные изображения просто теряются. Таким образом, локальное хранение информации и наличие избыточных распознающих нейронов позволяет запоминать новую информацию (новые образы), не искажая или теряя уже имеющуюся [7 – 10].

Основу архитектуры нейронной сети составляют три слоя нейронов: слой чувствительных S -нейронов, воспринимающих входную информацию в виде черно-белых изображений или бинарных входных векторов; слой бинарных интерфейсных Z -нейронов; слой распознающих Y -нейронов. Z - и S -элементы объединяются в поле F_1 входных обрабатывающих нейронов. Кроме этого, имеются еще три управляющих нейрона: R , G_1 и G_2 .

Выход каждого бинарного нейрона S_i ($i = \overline{1, n}$) входного слоя связан одной однонаправленной связью с единичным весовым коэффициентом с бинарным нейроном Z_i ($i = \overline{1, n}$) интерфейсного слоя. Выходные сигналы всех чувствительных элементов поступают также по связям с единичными весовыми коэффициентами на входы управляющих нейронов R , G_1 и G_2 . Каждый интерфейсный нейрон Z_i ($i = \overline{1, n}$) связан со всеми элементами Y_j ($j = \overline{1, m}$) распознающего слоя и с управляющим нейроном R . Связи с весовыми коэффициентами W_{ij}^1 ($i = \overline{1, n}$; $j = \overline{1, m}$), $0 \leq W_{ij}^1 \leq 1$ передают сигналы с выхода каждого Z -элемента на входы каждого Y -нейрона. Бинарные связи с весовыми коэффициентами W_{ji}^2 ($j = \overline{1, m}$; $i = \overline{1, n}$) связывают выход каждого распознающего нейрона с входом каждого интерфейсного нейрона. На входы каждого Z -элемента по связям с единичными весовыми коэффициентами поступает также сигнал с выхода бинарного управляющего нейрона G_1 . Сигналы с выходов всех Z -элементов по связям с единичными весовыми коэффициентами поступают на входы управляющего нейрона R . Такими же связями выходы всех распознающих нейронов соединены со входами нейрона G_1 , а их входы – с выходами управляющих нейронов G_2 и R . В любой момент времени нейроны S_i и Z_i ($i = \overline{1, n}$), G_1 и G_2 находятся в одном из двух состояний: 0 или 1.

Нейроны Y -слоя с помощью дополнительной нейронной сети, на рис. 1 не показанной, образуют слой конкурирующих нейронов. Каждый элемент Y_j ($j = \overline{1, m}$) распознающего слоя может находиться в одном из четырех состояний:

– активен, сеть АРТ-1 функционирует в режиме определения выходных сигналов нейронов распознающего слоя, выходной сигнал $U_{\text{вых}Y_j}$ удовлетворяет условию

$0 \leq U_{\text{вых}Y_j} \leq 2$ и определяется выходными сигналами Z -элементов:

$$U_{\text{вых}Y_j} = U_{\text{вх}Y_j} = \sum_{i=1}^n W_{ij}^1 U_{\text{вых}Z_i};$$

– активен, $U_{\text{вых}Y_j} = 1$, нейрон Y_j является победителем при текущем предъявлении входного изображения;

– неактивен, $U_{\text{вых}Y_j} = 0$, нейрон установлен в начальное состояние или проиграл соревнование с другими нейронами при текущем предъявлении входного изображения;

– заторможен, $U_{\text{вых}Y_j} = -1$, информация, хранящаяся в весах связей нейрона, не удовлетворяет текущему входному изображению по величине параметра сходства, поэтому нейрон исключен из соревнования.



Все связи нейронной сети АРТ-1 являются возбуждающими, за исключением связей от Y -элементов к нейрону G_1 и от интерфейсных Z -элементов к нейрону R , которые передают тормозящие сигналы.

Z - и Y -нейроны получают сигналы из трех источников: Z -нейроны – от управляющего нейрона G_1 и от нейронов слоев S и Y ; Y -нейроны – от управляющих нейронов G_2 , R и от Z -нейронов. Z - и Y -элементы переходят в активное состояние по правилу "два из трех", то есть только при наличии сигналов из двух различных источников. Например, если Y -нейрон получает только n единичных сигналов с выходов всех Z -элементов, но нет возбуждающего сигнала от управляющего нейрона G_2 , то перейти в активное состояние он не может.

Для обучения дискретных нейронных сетей АРТ-1 обычно используется метод быстрого обучения, при котором равновесные веса связей нейронов определяются за одно предъявление входного изображения [8 – 11].

Недостатки нейронной сети АРТ-1.

Поясним один из недостатков сети АРТ-1 с помощью следующего примера.

Пример 1. Пусть требуется с помощью сети разделить на два класса следующее множество векторов:

$$\begin{aligned} S^1 &= (1,1,1,1,0,0,0,0), & S^2 &= (0,1,1,1,0,0,0,0), & S^3 &= (1,0,1,1,0,0,0,0), \\ S^4 &= (1,1,0,1,0,0,0,0), & S^5 &= (1,1,1,0,0,0,0,0), & S^6 &= (0,0,0,0,1,1,1,1), \\ S^7 &= (0,0,0,0,0,1,1,1), & S^8 &= (0,0,0,0,1,0,1,1), & S^9 &= (0,0,0,0,1,1,0,1), \\ S^{10} &= (0,0,0,0,1,1,1,0). \end{aligned}$$

Классификация приведенных векторов может быть выполнена многими способами. Один из наиболее напрашивающихся – отнести к первому классу вектора S^1, S^2, \dots, S^5 , а ко второму классу – вектора S^6, S^7, \dots, S^{10} , то есть сформировать классы векторов, отличающиеся от векторов S^1 и S^6 не более, чем одним компонентом.

Возьмем сеть со следующими параметрами: $m = 6$ – число Y -нейронов в распознающем слое сети; $n = 8$ – число нейронов во входном слое сети; $p = 0,5$ – параметр сходства; выбирается такое числовое значения исходя из того, что векторы S^1, S^2, \dots, S^5 и S^6, S^7, \dots, S^{10} в каждом классе отличаются друг от друга не более, чем двумя компонентами; $W_{ij}^1 = 1/(1+n) = 1/(1+8) = 0,111$ – начальные значения весов связей W_{ij}^1 ($i = \overline{1, 8}; j = \overline{1, 6}$); $W_{ji}^2 = 1$ – начальные значения весов связей W_{ji}^2 ($i = \overline{1, 8}; j = \overline{1, 6}$); $L = 2$ – константа, для коррекции весов связей W_{ij}^1 ($i = \overline{1, 8}; j = \overline{1, 6}$) в процессе обучения; $q = 10$ – число классифицируемых входных бинарных векторов.

Нетрудно проверить, что в процессе обучения нейронной сети исходное множество векторов будет разбито на классы: $\{S^1, S^2, S^3\}, \{S^4, S^5\}, \{S^6, S^7, S^8\}, \{S^9, S^{10}\}$, если вектора будут подаваться на вход сети в порядке возрастания их номеров. Дополнительный анализ показывает, что разбиение векторов S^1, S^2, \dots, S^{10} на классы при заданной величине параметра сходства зависит и от порядка их предъявления в обучающей последовательности. Например, нетрудно получить следующие классификации K^r векторов S^1, S^2, \dots, S^5 : $K^1 = \{\{S^1, S^2, S^3\}, \{S^4, S^5\}\}$ (получена выше); $K^2 = \{\{S^1, S^2, S^4\}, \{S^3, S^5\}\}$; $K^3 = \{\{S^1, S^2, S^5\}, \{S^3, S^4\}\}$; $K^4 = \{\{S^1, S^3, S^5\}, \{S^2, S^4\}\}$ и т.д.

Неудача с решением задачи классификации указанного множества векторов связана с особенностями архитектуры и алгоритма функционирования нейронной сети АРТ-1. Во-первых, сеть в весах связей распознающего нейрона запоминает пере-

сечение бинарных входных векторов. Во-вторых, близость бинарных векторов определяется с помощью параметра сходства, учитывающего только единичные элементы сравниваемых векторов. В анализируемом примере пересечением бинарных векторов S^1, S^2, \dots, S^5 (а также векторов S^6, S^7, \dots, S^{10}) является вектор со всеми нулевыми компонента: $S^0 = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$.

Таким образом, при пересечении рассматриваемых множеств векторов полностью утрачена информация об единичных компонентах бинарных векторов. Запоминание информации в нейронной сети в виде пересечения входных векторов или изображений всегда приводит к потере информации. При малых значениях параметра сходства может теряться основная часть информации по классифицируемым или распознаваемым изображениям.

При запоминании множеств реализаций одного и того же динамического процесса, когда каждая реализация вносит что-то новое по сравнению с другими реализациями и когда запомнить каждую реализацию практически невозможно из-за их слишком большого числа, в памяти сети необходимо хранить не пересечения обучающих изображений, а комплексные изображения, полученные из исходных с помощью логической операции "ИЛИ", других математических операций, эвристик или комбинации указанных способов формирования обучающих изображений [11]. Частным случаем является запоминание информации в виде объединения, а не пересечения бинарных изображений или векторов. В рассматриваемом примере имеем $S^1 \cup S^2 \cup S^3 \cup S^4 \cup S^5 = S^1$. Запоминание единственного вектора S^1 приводит к получению матриц весов связей W^1, W^2 , с помощью которых могли бы правильно классифицироваться первые пять входных векторов, если бы после предъявления входных векторов не было, при наступлении резонанса, адаптации весов связей, меняющей память сети.

Обучение нейронной сети без адаптации весов связей распределенных распознающих Y -нейронов при наступлении резонанса, когда первыми во входной последовательности предъявляются вектора S^1, S^6 , приводит к разбиению векторов S^1, S^2, \dots, S^{10} на два естественных класса $\{S^1, S^2, S^3, S^4, S^5\}, \{S^6, S^7, S^8, S^9, S^{10}\}$.

Обобщая рассмотренный пример, можно предложить общий метод обучения НС АРТ-1 без адаптации весов связей при наступлении резонанса последовательностью входных изображений: $S^k, S^{k+1}, S^{k+2}, \dots, S^{q-1}, S^q$, упорядоченных по убыванию величины нормы входного изображения, то есть $\|S^k\| \geq \|S^{k+1}\| \geq \|S^{k+2}\| \geq \dots \geq \|S^{q-1}\| \geq \|S^q\|$. Такой метод обучения разбивает исходное множество обучающих изображений на минимальное число классов.

При упорядочивании предъявляемых изображений по возрастанию величины нормы входного изображения, то есть $\|S^k\| \leq \|S^{k+1}\| \leq \|S^{k+2}\| \leq \dots \leq \|S^{q-1}\| \leq \|S^q\|$, число классов изображений, как правило, возрастает. Если в рассматриваемом примере входные вектора предъявлять сети в порядке $S^2, S^3, S^4, S^5, S^7, S^8, S^9, S^{10}, S^1, S^6$, то в результате обучения сети при параметре сходства, равном 0,5, будет получено шесть классов: $\{S^2, S^3\}, \{S^4, S^5\}, \{S^7, S^8\}, \{S^9, S^{10}\}, \{S^1\}, \{S^6\}$.

Успешное функционирование предложенных алгоритмов обучения нейронной сети указывает на возможность алгоритмов обучения последовательностями изображений, упорядоченных учителем по величине их нормы, причем в обучающей последовательности могут использоваться подпоследовательности как с возрастанием, так и с убыванием нормы входных изображений. Возможно также использование во входной последовательности комплексных изображений, полученных учителем в результате применения к изображениям входной последовательности некоторых математических операций, эвристик или комбинации указанных способов формирова-



ния обучающих изображений. При решении реальных задач эффективно также запрещение адаптации весов связей при наступлении резонанса только части распознающих нейронов, которые могут быть выделены в отдельный слой [11].

Дискретные нейронные сети АРТ, определяющие множества равноценных решений

Еще один недостаток систем распознавания на основе нейронных сетей АРТ – получение единственного решения даже в тех случаях, когда имеется два или более равноценных решения. Поясним это с помощью следующего примера.

Пусть $n = 8$, параметр сходства ρ изображений равен 0,8 и в весах связей нейронов Y_1 и Y_2 после режима быстрого обучения хранятся изображения $S^1 = (1,1,1,1,0,0,0,1)$, $S^2 = (0,1,1,1,1,0,1,0)$, а веса связей сети имеют следующие значения: $W_{ij}^1 = 1/(1+n) = 0,111$ ($i = \overline{1,8}$; $j = \overline{3,m}$); $W_{i1}^1 = 0,333$ ($i = 1,2,3,4,8$); $W_{i1}^1 = 0$ ($i = 5,6,7$); $W_{i2}^1 = 0,333$ ($i = 2,3,4,5,7$); $W_{i2}^1 = 0$ ($i = 1,6,8$); $W_{1i}^2 = 1$ ($i = 1,2,3,4,8$); $W_{1i}^2 = 0$ ($i = 5,6,7$); $W_{2i}^2 = 1$ ($i = 2,3,4,5,7$); $W_{2i}^2 = 0$ ($i = 1,6,8$).

При подаче на вход сети изображения $S^3 = (1,1,1,1,0,0,1,0)$ на входах распознающих нейронов будут следующие сигналы: $U_{\text{вх}Y_1} = U_{\text{вх}Y_2} = 0,333 \cdot 5 = 1,665$, $U_{\text{вх}Y_3} = U_{\text{вх}Y_4} = \dots = U_{\text{вх}Y_m} = 0,111 \cdot 5 = 0,555$.

В силу алгоритма функционирования сети победителем станет нейрон Y_1 , при этом параметр сходства ρ будет иметь значение $\rho = \|U_{\text{вых}Z}\| / \|S^3\| = 4/5 = 0,8$, где $\|U_{\text{вых}Z}\|$ – норма вектора выходных сигналов нейронов слоя Z ; $\|S^3\|$ – норма вектора S^3 . Но точно такой же параметр сходства был бы получен и в случае, если бы нейрон-победителем стал распознающий элемент Y_2 .

Таким образом, при предъявлении входного изображения сеть относит это изображение к первому близкому изображению, хранящемуся в весах связей распознающих Y -нейронов, и тем самым определяет принадлежность входного изображения к одному из классов. Остальные возможные варианты решения задачи распознавания теряются. Для расширения возможностей дискретной нейронной сети АРТ и получения всех возможных вариантов решения задачи распознавания добавим к базовой архитектуре АРТ-1 еще один управляющий нейрон R^1 , инвертирующий выходные сигналы нейрона R (рис. 1), и слой регистрирующих нейронов Y_j^1 ($j = \overline{1,m}$). Каждый нейрон Y_j^1 связан однонаправленной бинарной связью с соответствующим нейроном Y_j . Перед началом режима распознавания нейроны Y_j^1 ($j = \overline{1,m}$), как и другие нейроны сети, переводятся в пассивное состояние по цепям связей, на рисунке не показанным. Нейроны Y_j^1 ($j = \overline{1,m}$) переходят в активное состояние по правилу "два из трех" – при наличии единичных сигналов на выходе нейронов Y_j и управляющего нейрона R^1 . Единичный сигнал с выхода нейрона Y_j^1 по цепи обратной связи фиксирует единичный выходной сигнал элемента Y_j^1 и затормаживает нейрон-победитель Y_j . После этого в сети начинается поиск нового нейрона-победителя. Процесс поиска продолжается до тех пор, пока все распознающие распределенные нейроны не окажутся заторможенными. При этом на выходах нейронов слоя Y^1 может не оказаться ни одного единичного сигнала, так как входное изображение не похоже ни на одно из изображений, хранящихся в весах связей сети, либо на выходах

Y^1 -нейронов может быть один или несколько единичных сигналов, указывающих на принадлежность входного изображения к одному или нескольким классам изображений.

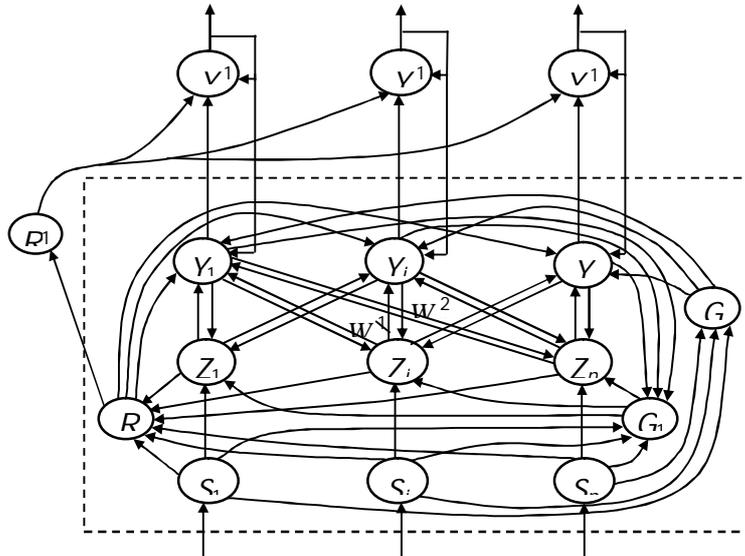


Рис. 1. Архитектура дискретной нейронной сети АРТ

Сети АРТ с несколькими полями обрабатывающих нейронов

При разработке систем распознавании сложных динамических процессов проявляется еще один недостаток дискретных НС АРТ-1 – наличие только одного поля обрабатывающих нейронов, поскольку это не позволяет применять эти сети для разработки систем распознавания, где необходимо использовать разнородную информацию о процессах, существенно отличающуюся значениями параметра сходства [11]. Для распознавания сложных динамических режимов необходима нейронная сеть с несколькими полями чувствительных элементов и соответствующих им полей Z-нейронов. Эта сеть должна как воспринимать и распознавать отдельные компоненты входной информации, так и корректировать веса связей распознающих нейронов при условии, что все распознанные компоненты с заданными значениями параметров сходства согласовываются в одном комплексном изображении. На рис. 2 приведена архитектура сети АРТ с d параллельно работающими полями чувствительных и интерфейсных нейронов. Сеть имеет также и d групп управляющих нейронов: $R^1, G_1^1, G_2^1, \dots, R^d, G_1^d, G_2^d$. Входы управляющих нейронов G_2, R связаны с выходами соответственно нейронов $G_2^q (q=1, d)$ и $R^q (q=1, d)$.

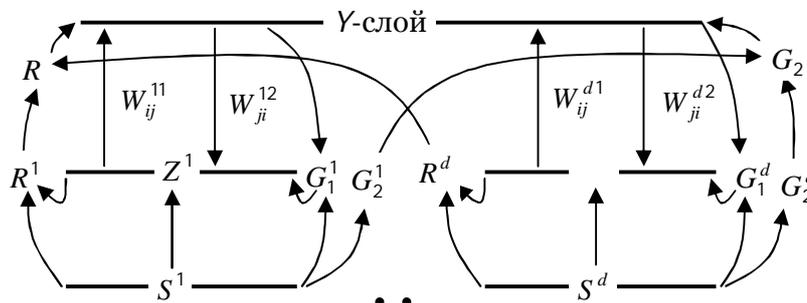


Рис. 2. Дискретная нейронная сеть АРТ с d параллельно работающими полями обрабатывающих нейронов



Нейрон G_2 переходит в активное состояние, обеспечивая работу распознающих нейронов по правилу "два из трех", только при активном состоянии всех управляющих нейронов $G_2^q (q=\overline{1, d})$. При определении нейрона-победителя в распознающем слое R -элемент не выдает запрещающего единичного сигнала нейрону-победителю только случае, когда на его вход не поступает ни одного единичного сигнала с выхода нейронов $R^q (q=\overline{1, d})$. Это возможно только при соответствии всех компонент входного изображения компонентам изображения, хранящегося в памяти нейронной сети.

Любое поле $F_1^q (q=\overline{1, d})$ нейронной сети с параллельно функционирующими полями обрабатывающих нейронов выполняет функции, аналогичные функциям поля F_1 сети АРТ-1 [11].

Каждый распознающий Y -нейрон этой сети, как и сети АРТ-1, связан со всеми интерфейсными нейронами двумя видами связей, передающих информацию как от интерфейсных нейронов к распознающим, так и наоборот. Алгоритмы функционирования нейронной сети с параллельно работающими полями чувствительных и интерфейсных нейронов нетрудно восстановить, если известен алгоритм работы сети АРТ-1.

Принципиальное отличие и преимущество дискретной нейронной сети АРТ с параллельно функционирующими полями обрабатывающих нейронов по сравнению с сетью АРТ-1 состоит в том, что сложное изображение может быть разделено на нужное число фрагментов, и каждый фрагмент может распознаваться независимо со своим параметром сходства.

Метод, разработанный для обучения дискретной нейронной сети АРТ с одним полем чувствительных элементов и все алгоритмы, из него вытекающие, могут использоваться и в рассматриваемой нейронной сети. Однако число конкретных алгоритмов здесь существенно возрастает, поскольку в каждом поле может использоваться свой алгоритм определения весов связей распознающих нейронов.

Выводы

Таким образом, разработан новый метод обучения дискретных нейронных сетей АРТ без адаптации в процессе обучения весов связей распределенных распознающих нейронов и на его основе предложен ряд новых алгоритмов функционирования дискретных НС АРТ в режиме обучения как с учителем, так и без него. Эти алгоритмы открывают возможности компактного хранения больших объемов информации и расширяют возможности использования дискретных сетей АРТ для решения задач классификации и распознавания.

Разработана новая архитектура и алгоритмы функционирования дискретной НС АРТ, позволяющие определять два или более равноценных решения (если они существуют) в задачах распознавания. Предложена новая дискретная НС АРТ с параллельно работающими полями обрабатывающих нейронов. Принципиальное отличие и преимущество этой сети по сравнению с сетью АРТ-1 состоит в том, что сложное изображение может быть разделено на нужное число фрагментов, и каждый фрагмент может распознаваться независимо со своим параметром сходства.

Перспективой дальнейших исследований является разработка непрерывных НС АРТ, позволяющих определять два или более равноценных решения в задачах распознавания или классификации и работающих с несколькими полями чувствительных нейронов.

Литература

1. Комашинский В.И. Нейронные сети и их применение в системах управления и связи [Текст] / В.И. Комашинский, Д.А. Смирнов. – М.: Горячая линия – Телеком, 2002. – 94 с.



2. Neural networks for control / Edited by *W. Thomas Miller III, Richard S. Sutton, and Paul J. Werbos*, – Cambridge, Massachusetts, London: MIT Press, 1996. – 524 P.
3. Барский А.Б. Нейронные сети: распознавание, управление, принятие решений [Текст] / А.Б. Барский. – М.: Финансы и статистика, 2004. – 176 с.
4. Галушкин А.И. Нейрокомпьютеры и их применение на рубеже тысячелетий в Китае. В 2-х томах. Том 2 [Текст] / А.И. Галушкин. – М.: Горячая линия – Телеком, 2004. – 464 с.
5. Хайкин С. Нейронные сети: полный курс [Текст] / С. Хайкин С. – М.: Издательский дом «Вильямс», 2006. – 1104 с.
6. Комарцова Л.Г. Нейрокомпьютеры: Учеб. пособие для вузов [Текст] / Л.Г. Комарцова Л.Г., А.В. Максимов. – М.: Изд-во им. Н.Э. Баумана, 2002. – 320 с.
7. Grossberg S. Competitive learning: From interactive activation to adaptive resonance // *Cognitive Science* – 1987. – Vol. 11. – P. 23 – 63.
8. Carpenter G.A., Grossberg S. A massively parallel architecture for selforganizing neural pattern recognition machine // *Computing, Vision, Graphics and Image Processing*. – 1987. – Vol. 37. – P. 54 – 115.
9. Дмитриенко В.Д., Корсунов Н.И. Основы теории нейронных сетей [Текст] / В.Д. Дмитриенко, Н.И. Корсунов. – Белгород: БИИММАП, 2001. – 159 с.
10. Fausett L. *Fundamentals of Neural Networks. Architectures, Algorithms and Applications*. – New Jersey: Prentice Hall International, Inc., 1994. – 461 p.
11. Носков В.И. Моделирование и оптимизация систем управления и контроля локомотивов [Текст] / Носков В.И., Дмитриенко В.Д., Заполовский Н.И., Леонов С.Ю. – Х.: ХФИ Транспорт Украины, 2003. – 248 с.

NEW ARCHITECTURES AND TRAINING ALGORITHMS OF DISCRETE NEURAL NETWORKS OF ADAPTIVE RESONANCE THEORY

V. D. DMITRIENKO
I. P. KHAVINA
A. Y. ZAKOVOROTNYI

*National Technical
University «Kharkov
Polytechnic Institute»*

e-mail: Arcade@i.ua

In article analyzed lacks and advantages of architecture and training algorithms of Adaptive Resonance Theory (ART) discrete neural networks. Offered new architectures of ART neural networks and training algorithms of these networks without adaptation weights connections of distributed recognizing neurons.

Key words: adaptive resonance theory discrete neural networks, training algorithms.

О СТОЙКОЙ ОБФУСКАЦИИ КОМПЬЮТЕРНЫХ ПРОГРАММ

Н. П. ВАРНОВСКИЙ¹⁾
В. А. ЗАХАРОВ¹⁾
Н. Н. КУЗЮРИН²⁾
А. В. ШОКУРОВ²⁾

¹⁾ *Московский
государственный
университет
им. М.В. Ломоносова*

e-mail: zakh@cs.msu.su

²⁾ *Институт
системного программирования
Российской академии наук*

e-mail: nnkuz@ispras.ru

В данной работе предлагается схема аппроксимации функций и их производных по эмпирическим данным. Она основана на использовании известной из математического анализа формулы, позволяющей выразить дифференцируемую функцию через производную

Ключевые слова: интерполяция, оценка производной, частотное представление, устойчивость вычислений.

Введение

Обфускацией программ называется такое преобразование программ, которое сохраняет все функциональные характеристики исходной программы, лишь незначительно ухудшает ее быстродействие, но при этом делает чрезвычайно трудоемким извлечение из открытого текста программы ключевой информации об устройстве содержащихся в ней алгоритмов и структур данных. Впервые концепция обфускации программ была предложена и подробно исследована в статье [1]. В этой и последующих работах обсуждалась возможность применения обфускации программ для защиты авторских прав на программное обеспечение [1], для предотвращения реинженерии программ [2], для создания и защиты «водяных знаков» [3], для информационной защиты мобильных агентов [4], для обеспечения приватности информационного поиска [5] и др. Обфускация программ также находит применение в криптографии; с ее помощью можно конструировать системы шифрования с открытым ключом [6-8], проводить вычисления над зашифрованными данными [6,9] и перешифрование сообщений [10], создавать верифицируемые системы тайного голосования [11] и др.

Однако методы обфускации программ имеют ценность лишь в том случае, если они сопровождаются строго обоснованными гарантиями их стойкости. Стойкость методов обфускации программ можно оценивать в зависимости от того, какие цели преследует противник, получивший доступ к тексту обфускированной программы, и какими ресурсами он располагает для достижения своей цели. Целью противника может быть получение информации об устройстве программы, вычисляемой ею функции, содержащихся в программе алгоритмах и структурах данных. Ресурсы, которыми располагает противник, включают предварительное знание об исходной программе, а также время, отведенное для решения задачи деобфускации, набор инструментальных средств анализа программ и др. Кроме того, возможности противника зависят от того, в какой форме представлена обфускированная программа (например, противник может иметь доступ ко всей программе или только к некоторой выделенной ее части). Эти три параметра (цель, предварительные знания и вычислительные ресурсы) полностью определяют модель противника.

В криптографии задачу взлома систем защиты, т. е. цель, стоящую перед противником, принято называть *угрозой безопасности* этой системы, а предположения о знаниях и возможностях противника, т. е. о его ресурсах – *атакой* на систему защиты информации. Традиционно классификация и оценка стойкости систем защиты

информации проводят относительно пар «угроза-атака», определяющих модель противника. Поэтому классификацию определений стойкости обфускации программ также наиболее естественно проводить относительно возможных моделей противника с учетом различных форм представления программ. В настоящей заметке для различных моделей противника приведены формальные определения стойкости обфускации программ, а также рассмотрены положительные и отрицательные результаты, оценивающие возможность построения стойких обфускаторов.

Стойкость обфускации программ

1. Модель «черного ящика». Впервые строгое математическое определение стойкости обфускации программ было предложено в работе [6]: обфускация считается стойкой, если всякий противник может извлечь из текста обфускированной программы за разумное (полиномиальное относительно размера входных данных) время не больше информации, чем можно было бы получить, проводя тестовые испытания этой программы как «черного ящика».

Определение 1. Вероятностный алгоритм O называется *обфускатором программ, стойким в модели «черного ящика»*, если он удовлетворяет следующим трем требованиям:

1. Для всякой машины Тьюринга (МТ) M любое выходное слово $O(M)$ машины O на входе M описывает МТ, вычисляющую ту же функцию, что и машина M (требование функциональности);
2. Длина описания и время выполнения любой МТ $O(M)$ не более чем полиномиальны от соответствующих параметров машины M (требование эффективности);
3. Для любой полиномиальной вероятностной МТ (РРТ) A (противник) существует РРТ S (симулятор) и пренебрежимо малая функция ν , удовлетворяющие для любой МТ M соотношению $|\Pr[A(O(M)) = 1] - \Pr[S^M(1^{|M|}) = 1]| \leq \nu(|M|)$ (требование стойкости).

В статье [6] было установлено, что стойких обфускаторов программ в модели «черного ящика» не существует. Вместе с тем, в работах [8,9,10,13] было показано, что для отдельных классов функций (т.н. «точечных» функций) стойкая обфускация вычисляющих их программ возможна при тех или иных криптографических предположениях. Наиболее значительное продвижение было достигнуто в работе [10]; ее авторы доказали осуществимость стойкой обфускации программ перешифрования сообщений. В целом, однако, результаты исследований в этом направлении не дают больших оснований для оптимизма: до сих пор не удалось обнаружить достаточно сложной функции, программы вычисления которой допускают доказуемо стойкую обфускацию в модели «черного ящика».

2. Модель «серого ящика». В модели «черного ящика» противник стремится извлечь хотя бы какую-нибудь информацию об исходной программе. В модели «серого ящика» угроза противника ослаблена: предполагается, что трассы вычисления исходной программы не являются секретом для противника. Для формулировки требования стойкости обфускации в новой модели введем оракул $Tr(M)$, представляющий собой математическую модель «серого ящика». На входе x оракул $Tr(M)$ возвращает пару $(y, tr_{M(x)})$, где y – результат работы машины M на входе x , а $tr_{M(x)}$ – трасса вычисления машины M на этом входе. Строка $tr_{M(x)}$ определяется как конкатенация всех инструкций машины M , последовательно выполняемых на входе x .

Определение 2. Вероятностный алгоритм O называется *обфускатором программ, стойким в модели «серого ящика»*, если он удовлетворяет требованиям функциональности, эффективности, а также следующему требованию стойкости: для любой РРТ A существует РРТ S и пренебрежимо малая функция ν , удовлетворяющие для любой МТ M соотношению $|\Pr[A(O(M)) = 1] - \Pr[S^{Tr(M)}(1^{|M|}) = 1]| \leq \nu(|M|)$.

Нам удалось установить, что универсальных обфускаторов, стойких, стойких в модели «серого ящика», также невозможно построить.

Теорема 1. *Эффективные обфускаторы, стойкие в модели «серого ящика», не существуют.*

Доказательство. Предположим, что существует эффективный алгоритм (PPT) $O(M)$, удовлетворяющий требованиям определения 2. Для каждой тройки $\alpha, \beta \in \{0,1\}^n$ и $b \in \{0,1\}$ определим машину Тьюринга $C_{\alpha,\beta,b}$ следующим образом. Пусть x_1, x_2, \dots – последовательность строк, подаваемых на вход машине $C_{\alpha,\beta,b}$. Если $x_1 = \alpha$, то машина выдает β , в ответ на $x_1 = \beta$ или $x_2 = \beta$ она выдает b , а во всех остальных случаях она выдает в ответ число 0.

Пусть r – случайная строка полиномиальной длины, используемая вероятностным алгоритмом O . Определим функцию $f(\alpha, \beta, b, r) = O(C_{\alpha,\beta,b,r})$. Очевидно, что функция f эффективно вычислима. Поскольку значение функции f однозначно определяет функцию, вычисляемую машиной $C_{\alpha,\beta,b}$, значение бита b также определено однозначно.

Машина S , имеющая оракульный доступ к функции, вычисляемой машиной $C_{\alpha,\beta,b}$, может получить в качестве ответа от оракула ненулевое значение лишь с пренебрежимо малой вероятностью. Отсюда $\Pr\{S^{C_{\alpha,\beta,b}} = b\} \leq 1/2 + \nu(n)$, где ν – пренебрежимо малая функция. Следовательно, согласно определению 2, для любой полиномиальной вероятностной машины Тьюринга A $\Pr\{A(F(\alpha, \beta, b, r) = b)\} \leq 1/2 + \nu(n)$. Таким образом, b – трудный предикат функции f , и функция f односторонняя.

Теперь достаточно построить, используя одностороннюю функцию, семейство функций, необфускируемое в модели с «серым ящиком». Полученное противоречие докажет теорему.

Основная идея дальнейшего доказательства достаточно проста. Барак и др. [6] построили такое бесконечное семейство машин Тьюринга, что некоторый предикат $\pi(M)$, заданный на этом семействе, невыведываем при оракульном доступе к функции, вычисляемой машиной M , но может быть легко вычислен, если дан текст любой программы, эквивалентной M . Достаточно модифицировать эту конструкцию таким образом, чтобы следующие два требования выполнялись одновременно. Во-первых, выведываемость предиката $\pi(M)$ при наличии текста любой программы, эквивалентной M , должна сохраняться. Во-вторых, трассы выполнения машины M не должны давать противнику никакой дополнительной полезной информации, по сравнению с парами (вход, выход), получаемыми от «черного ящика». Противник, имеющий доступ к трассам выполнения машины M , представляется более сильным, чем противник, который видит только пары вход-выход. Для защиты от столь сильного противника мы используем криптографические примитивы. А именно, вместо сравнения входа x с фиксированной строкой α , как в работе Барака и др. [6], сначала проверяем равенство $f(x) = f(\alpha)$, где f – односторонняя функция. Лишь в том случае, когда равенство выполняется, мы проверяем равенство $x = \alpha$.

Типичная трасса $tr_{M(x)}$ состоит из инструкций, вычисляющих значение $f(x)$ и проверяющих равенство $f(x) = \alpha$. В подавляющем большинстве случаев результат проверки отрицательный. Более того, если равенство $f(x) = \alpha$ выполняется с вероятностью, которая не является пренебрежимо малой, то это противоречит тому, что функция f односторонняя.

Следует подчеркнуть, что до сих пор излагалась лишь общая идея доказательства. В действительности используется несколько более сложная конструкция.

Далее переходим к формальному доказательству. Контрпример, построенный Бараком и др. [6] использует два семейства машин Тьюринга. Для произвольной пары строк $\alpha, \beta \in \{0,1\}^n$ машина Тьюринга $C'_{\alpha,\beta}(x)$ выдает β , если $x = \alpha$, и выдает 0^n в противном случае. Для тех же параметров α, β машина Тьюринга $D'_{\alpha,\beta}(C)$ выдает 1, если $C(\alpha) = \beta$, и выдает 0 в противном случае.

Пусть $\{f : \{0,1\}^n \rightarrow \{0,1\}^n\}$ – односторонняя функция. Наша конструкция зависит также от целочисленного параметра $t \geq 2$, который может быть константой или значением произвольного, но фиксированного полинома (от n).

Выбираем случайно и равномерно $2t$ строк $\alpha_1, \dots, \alpha_t, \beta_1, \dots, \beta_t \in \{0,1\}^n$. Обозначим эту совокупность из $2t$ строк через γ . Теперь определяем машину Тьюринга $C_\gamma(x)$ следующим образом. На входе x эта машина вычисляет $f(x)$ и сравнивает результат с заранее вычисленными значениями $f(\alpha_i)$, $i = 1, \dots, t$. Если $f(x) \leq f(\alpha_i)$ для всех $i = 1, \dots, t$, то C_γ выдает 0^n и останавливается. В случае, если $f(x) = f(\alpha_i)$ для некоторого i , C_γ проверяет равенство $x = \alpha_i$, и, если оно выполняется, выдает β_i , иначе выдает 0^n , и в обоих случаях останавливается.

Далее мы определяем машину Тьюринга D_γ . Она хранит два массива строк $\alpha_1, \dots, \alpha_t$ и β_1, \dots, β_t . Пусть i – индекс для обоих этих массивов, который изначально принимается равным 0. Получив на вход машину Тьюринга C , машина D_γ действует следующим образом.

1. Проверяет равенство $i=t-1$. Если оно выполняется, выдает 0 и останавливается.
2. Увеличивает индекс $i=i+1$ и затем вызывает машину C на входе α_i .
3. Если $C(\alpha_i) = \beta_i$, то проверяет текущее значение индекса. Если $i=t$, выдает 1 и останавливается, иначе выполняется переход на шаг 2.
4. Если $C(\alpha_i) \neq \beta_i$, то выдает 0 и останавливается.

Заметим, что моделирующая машина, имеющая доступ к машинам C_γ и D_γ как оракулам, может подать на вход машине D_γ произвольную машину C и извлечь значение α из трассы, затем подать машине C_γ в качестве запроса строку α_1 и получить β_1 (даже в случае модифицированной машины D_γ , скрывающей значения β_i) и т. д. По этой причине моделирующая машина может выдать лишь $t-1$ запросов машине D_γ (на самом деле, дальнейшие запросы допускаются, но завершаются нулевыми ответами, а потому могут игнорироваться). Необфускируемое свойство состоит в существовании t различных строк $\alpha_1, \dots, \alpha_t \in \{0,1\}^n$ таких, что на каждой из них данная машина Тьюринга выдает ненулевой ответ.

Пара машин Тьюринга (C_γ, D_γ) заменяет машины $(C'_{\alpha,\beta}, D'_{\alpha,\beta})$, используемые в доказательстве отрицательного результата в работе Барака и др. [6]. Анализ этого доказательства показывает, что оно переносится на рассматриваемый случай, если верно следующее утверждение.

Утверждение. Для любой полиномиальной вероятностной машины Тьюринга S выполняется равенство

$$|\Pr\{S^{C_\gamma, D_\gamma}(1^n) = 1\} - \Pr\{S^{Z_\gamma, D_\gamma}(1^n) = 1\}| = v(n),$$

где Z_γ – машина, которая отличается от C_γ только тем, что на входе α_i выдает 0^n , и вероятности определяются случайным равновероятным выбором строк $\alpha_1, \dots, \alpha_s, \beta_1, \dots, \beta_t \in \{0,1\}^n$ и случайными величинами машины S .

Доказательство. Достаточно показать, что любая полиномиальная вероятностная машина Тьюринга на любой из своих запросов к оракулу получает ненулевой ответ лишь с пренебрежимо малой вероятностью, вне зависимости от того, какая из двух машин, C_γ или Z_γ , используется в качестве первого оракула.

Для простоты изложения всюду далее предполагаем, что $t=2$. В этом случае машина S может выдать второму оракулу лишь один запрос. Пусть $a_1, \dots, a_s \in \{0,1\}^n$ – все запросы машины S к первому оракулу до выдачи единственного запроса ко второму (подчеркнем, что s является случайной переменной). Предположим, что вероятность появления ненулевой строки среди ответов на эти s запросов не является пренебрежимо малой. Тогда можно построить полиномиальную вероятностную машину Тьюринга (обозначим ее через T) для инвертирования функции f .

Пусть $z \in_R \{0,1\}^n$ и $y=f(z)$ --- входная строка машины T . Машина T подбрасывает монету и, на основе исхода этого эксперимента, решает, какое из значений, $f(\alpha_1)$ или $f(\alpha_2)$ будет принято за y . Без ограничения общности, пусть это будет α_1 . Далее, T выбирает случайные равновероятные строки $\alpha_2, \beta_1, \beta_2 \in \{0,1\}^n$ и вызывает S как подпрограмму. Машина T перехватывает все запросы к первому оракулу. Выполняя запрос $x \in \{0,1\}^n$, машина T вычисляет $f(x)$ и проверяет равенства $f(x)=y$ и $f(x) = f(\alpha_2)$. Если ни одно из них не выполняется, машина T выдает 0^n . В случае $f(x)=y$ машина T выдает β_1 , а в случае $f(x) = f(\alpha_2)$ она действует так же как машина C_γ .

Очевидно, что вероятность получения ненулевого ответа от этого моделируемого оракула T не меньше, чем в случае реального оракула.

Поэтому для некоторого $j \in \{1, \dots, s\}$ вероятность получить от машины T ненулевой ответ не является пренебрежимо малой. Для этого значения j , α_j принадлежит прообразу значения y с вероятностью $1/2$. Таким образом, T инвертирует функцию f с вероятностью, которая не является пренебрежимо малой, что противоречит тому, что функция f является односторонней.

Если в ответах на первые s запросов ненулевая строка появляется лишь с пренебрежимо малой вероятностью, то вероятность получить ненулевой ответ на единственный запрос машины S ко второму оракулу также пренебрежимо мала.

Для остальных запросов к оракулу (после выполнения запроса ко второму оракулу) тот же аргумент, что и выше, показывает, что если вероятность успеха не является пренебрежимо малой, то существует эффективный алгоритм инвертирования функции f . Это противоречие доказывает утверждение, и, тем самым, завершает доказательство теоремы 1.

Ранее, в статье [14] нами было доказано, что если существуют односторонние функции, то обфускаторы, стойкие в модели «серого ящика», не существуют. Полученное усиление этого результата явилось следствием доказанного нами факта о том, что существование обфускаторов в смысле определения 2 влечет существование односторонних функций.

3. Обфускация с дополнительным входом. При проведении обфускации программ разумно предполагать, что противнику известны некоторые сведения о защищаемых программах. Например, противник может располагать инструкцией пользователя программой, каким-либо описанием функции, вычисляемой программой, или примерами обфускации некоторых программ. Естественно, эти дополнительные знания могут существенно усилить возможности противника. В криптогра-

фии атаки такого рода (например, атака с выбором зашифрованных сообщений) считаются наиболее сильными. В статье [7] было предложено формальное определение стойкости обфускации программ (в неунiformной модели вычислений) в случае, когда противник обладает дополнительными сведениями о защищаемых программах.

Определение 3. Вероятностный алгоритм O называется *обфускатором с дополнительным входом*, если он удовлетворяет требованиям функциональности, эффективности, а также следующему требованию стойкости: для любой РРТ A существует такие РРТ S и пренебрежимо малая функция ν , что для любого полинома $q(\cdot)$ и битовой строки z , длина которой не превосходит $q(|M|)$, выполняется соотношение $|\Pr[A(O(M), z) = 1] - \Pr[S^M(1^{|M|}, z) = 1]| \leq \nu(|M|)$.

В этой же работе было показано, что некоторые семейства функций не могут быть эффективно и стойко обфускированы в рамках определения 3.

4. Защита алгоритмов. Дальнейшее ослабление угрозы противника приводит нас к задаче применения обфускации для защиты алгоритмов. Предполагается, что противнику известна функция, вычисляемая программой, и его цель – извлечение информации об особенностях программной реализации этой функции. Знание противником функции, вычисляемой программой M формально выражается предоставлением ему доступа к произвольной эквивалентной программе M_0 , вычисляющей ту же самую функцию. Таким образом, мы приходим к следующему определению.

Определение 4. Вероятностный алгоритм O называется *обфускатором алгоритмов*, если он удовлетворяет требованиям функциональности, эффективности, а также следующему требованию стойкости: для любой РРТ A существует РРТ S и пренебрежимо малая функция ν , для которых соотношение $|\Pr[A(O(M), M_0) = 1] - \Pr[S^M(1^{|M|}, M_0) = 1]| \leq \nu(|M|)$ выполняется для любой пары эквивалентных МТ M, M_0 .

Вопрос о существовании обфускаторов алгоритмов остается открытым. В следующей теореме мы представляем частичный отрицательный результат: такие обфускаторы не существуют в модели со случайным оракулом. В этой модели все алгоритмы, входящие в определение обфускатора, имеют доступ к оракулу, вычисляющему случайную функцию.

Теорема 2. В модели со случайным оракулом обфускаторы алгоритмов не существуют.

Доказательство. Определим дельта-функцию I_α , $\alpha \in \{0,1\}^n$, следующим образом: $I_\alpha(x) = 1$, если $x = \alpha$, и $I_\alpha(x) = 0$ во всех остальных случаях. Пусть $R: \{0,1\}^{2n} \rightarrow \{0,1\}^{2n}$ – функция, вычисляемая случайным оракулом.

Для всякой пары $\alpha, r \in \{0,1\}^n$ определим машину Тьюринга $M_{\alpha,r}$ следующим образом. Машина $M_{\alpha,r}$ хранит в своей памяти строки r и $\beta = R(\alpha, r)$. На входе $x \in \{0,1\}^n$ машина $M_{\alpha,r}$ обращается к случайному оракулу с запросом (x, r) . Если $R(x, r) = \beta$, то машина $M_{\alpha,r}$ выдает 1, во всех остальных случаях – 0, и останавливается.

Рассматривается семейство машин Тьюринга $\{M_{\alpha,r}\}$ для всех $\alpha \in \{0,1\}^n$ и всех длин входа n .

Согласно определению обфускатор сохраняет функциональность машины Тьюринга, и поэтому любая машина $O(M_{\alpha,r})$ вычисляет дельта-функцию I_α .

Рассмотрим следующего противника A . На входе $(O(M), \tilde{M})$ машина A находит n и далее вычисляет строку $r_0 = r_0(n)$, $r_0 \in \{0,1\}^n$. Здесь $r_0(\cdot)$ – произвольная эффективно вычисляемая функция, которая каждому натуральному числу n ставит в со-

ответствие некоторую битовую строку длины n . Предположение о том, что машина A может найти значение параметра n не ограничительно, поскольку такую возможность всегда можно обеспечить с помощью специальным образом построенной машины \tilde{M} .

Далее машина A выполняет машину $O(M)$, подавая ей на вход случайную строку $x \in_R \{0,1\}^n$. Если среди запросов машины $O(M)$ к оракулу встретится запрос (cx, r_0) , то машина A выдает 1, в противном случае – 0.

Для дальнейшего зафиксируем какой-либо простой алгоритм вычисления дельта-функции I_α . Например, на входе $x \in \{0,1\}^n$ проверяем равенство $x = \alpha$, и если оно выполняется, выдаем 1, иначе – 0. Будем предполагать, что машина \tilde{M}_α реализует именно этот алгоритм.

Сравним вероятность успеха противника в двух экспериментах. В первом из них машина A получает на вход обфускацию машины M_{α, r_0} . Очевидно, что

$$\Pr\{A(O(M_{\alpha, r_0}), \tilde{M}_\alpha) = 1\} \geq 1 - \nu(n),$$

где ν – пренебрежимо малая функция. Во втором эксперименте машина A получает обфускацию машины $M_{\alpha, r}$ для случайной строки $r \in_R \{0,1\}^n$. В этом случае

$$\Pr\{A(O(M_{\alpha, r}), \tilde{M}_\alpha) = 1\} \leq \nu(n).$$

Машина S получает только описание машины \tilde{M}_α и оракульный доступ к дельта-функции I_α . Это не дает ей никакой информации о значении r_0 . Следовательно, машина S не может различить два случая, соответствующих описанным выше экспериментам, и требования определения не могут быть выполнены. Теорема доказана.

Вместе с тем, обфускацию простейших алгоритмов можно провести сравнительно просто. В работах [15,16] показано, каким образом можно провести обфускацию детерминированных конечных автоматов и упорядоченных двоичных разрешающих диаграмм (OBDD).

5. Защита параметров. Во многих криптографических приложениях обфускация применяется лишь для одной цели – скрыть содержащуюся в программе константу (ключ), выбранную случайным образом из некоторого конечного множества. При этом считается, что сам алгоритм, реализуемый программой, известен противнику. Для формализации задачи обфускации программ в такой постановке рассмотрим параметризованное множество МТ $M = \{M(c) : c \in \{0,1\}^n\}$, отличающихся друг от друга только константами c .

Определение 5. Вероятностный алгоритм O называется *обфускатором константы*, если он удовлетворяет требованиям функциональности, эффективности и следующему требованию стойкости: для любой РРТ A существует РРТ S и пренебрежимо малая функция ν , для которых неравенство $|\Pr[A(O(M(c)), M(c_0)) = 1] - \Pr[S^{M(c)}(1^{|M(c)|}, M(c_0)) = 1]| \leq \nu(n)$ выполняется для любой пары констант c, c_0 , равномерно случайно выбранных из множества $\{0,1\}^n$.

Возможность построения стойких обфускаторов констант, зависит от параметризованного множества МТ M . Если $M(c)$ – это универсальная машина Тьюринга, в которой константа c играет роль моделируемой программы, то задача обфускации константы равносильна задаче обфускации программ в модели «черного ящика» с дополнительным входом. В статье [7] показано, что универсального стойкого обфускатора программ с дополнительным входом не существует. В то же время, алгоритм шифрования в криптосистеме с открытым ключом можно рассматривать как разно-

видность обфускации, скрывающей константу – секретный ключ. Обфускацию сингулярных функций, исследованную в работах [9, 12], также можно рассматривать как стойкую защиту констант.

6. Соккрытие предикатов. Наиболее слабая угроза – это распознавание некоторого функционального свойства (предиката) программы. Таким свойством может быть, например, наличие в программе скрытых (недекларированных) функциональных возможностей, побочных эффектов, вредоносных участков кода (вирусов).

Определение 6. Вероятностный алгоритм O называется *обфускатором предиката* π на множестве МТ M , если он удовлетворяет требованиям функциональности, эффективности, а также следующему требованию стойкости: для любой РРТ A существует РРТ S и пренебрежимо малая функция ν , удовлетворяющие для любой МТ M из класса M соотношению $|\Pr[A(O(M)) = \pi(M)] - \Pr[S^M(1^{|M|}) = \pi(M)]| \leq \nu(n)$.

Как видно из приведенного определения, обфускацию в модели «черного ящика» можно рассматривать как соккрытие всех возможных предикатов. Невозможность построения обфускатора, стойкого в модели «черного ящика», установленная в статье [6], не отрицает существования стойких обфускаторов, скрывающих отдельные предикаты для некоторых классов программ. Так, например, в статье [12] показано, что для класса программ, вычисляющих константы и точечные функции существует стойкая обфускация следующего свойства: «функция, вычисляемая программой, является константой».

Работа выполнена при поддержке гранта РФФИ 09-01-00632.

Литература

1. Collberg C., Thomborson C., Low D. A taxonomy of obfuscating transformations // Tech. Report, N 148, Univ. of Auckland, 1997.
2. Collberg C., Thomborson C., Low D. C. Manufacturing cheap, resilient and stealthy opaque constructs // Symp. on Principles of Programming Languages, 1998, p. 184-196.
3. Collberg C., Thomborson C. Watermarking, tamper-proofing, and obfuscation – tools for software protection // IEEE Transactions on Software Engineering, v. 28, N 6, 2002.
4. D'Anna L., Matt B., Reisse A., Van Vleck T., Schwab S., LeBlanc P. Self-protecting mobile agents obfuscation report // Report #03-015, Network Associates Laboratories, 2003.
5. Ostrovsky R., Skeith III W.E. Private searching on streaming data // Lecture Notes in Computer Science, v. 3621, 2005, p. 223-240.
6. Barak B., Goldreich O., Impagliazzo R., Rudich S., Sahai A., Vadhan S., Yang K. On the (Im)possibility of obfuscating programs // Lecture Notes in Computer Science, v. 2139, 2001, p. 1-18.
7. Goldwasser S., Kalai T.Y. On the impossibility of obfuscation with auxiliary input // Proc. of the 46th IEEE Symp. on Foundations of Computer Science, 2005, 553-562.
8. Hofheinz D., Malone-Lee J., Stam M. Obfuscation for cryptographic purposes // Lecture Notes in Computer Science, v. 4392, 2007, p. 214-232.
9. Lynn B., Prabhakaran M., Sahai A. Positive results and techniques for obfuscation // Lecture Notes in Computer Science, v. 3027, 2004, p. 20-39.
10. Hohenberger S., Rothblum G. N., Shelat A., Vaikuntanathan V. Securely obfuscating re-encryption // Lecture Notes in Computer Science, v. 4392, 2007, p. 233-252.
11. Adida B., Wikstrom D. Obfuscated ciphertext mixing // IACR. Eprint Archive, N 394, 2005.
12. Wee H., On obfuscating point functions // Proc. of 37th ACM Symp. on Theory of Computing, 2005, p. 523-532.
13. Varnovsky N.P., Zakharov V.A. On the possibility of provably secure obfuscating programs // Lecture Notes in Computer Science, v. 2890, 2003, p. 91-102.
14. Varnovsky N.P. A note on the concept of obfuscation // Труды Института системного программирования, т. 6, 2004.
15. Goldwasser, S., Rothblum G. On best-possible obfuscation // Lecture Notes in Computer Science, v. 4392, 2007, p. 253-272.
16. Kuzurin N.N., Shokurov A.V., Varnovsky N.P., Zakharov V.A. On the concept of software obfuscation in computer security // Lecture Notes in Computer Science, v. 4779, 2007, p. 281-298.



ON THE SECURE OBFUSCATION OF COMPUTER PROGRAMS

N. P. VARNOVSKY¹⁾

V. A. ZAKHAROV¹⁾

N. N. KUZURIN²⁾

A. V. SHOKUROV²⁾

¹⁾ *Lomonosov Moscow
State University*

e-mail: zakh@cs.msu.su

²⁾ *Institute for System
Programming of Russian
Academy of Science*

e-mail: nnkuz@ispras.ru

The aim of program obfuscation is to bring a program into such a form, which impedes the understanding of its algorithm and data structures or prevents extracting of some valuable information from the text of a program. In this paper we address the issue of defining security of program obfuscation. We analyze several formal definitions of obfuscation security, consider positive and negative results on obfuscation security and the applications where these model may be valid.

Keywords: computer security, program obfuscation, security, Turing machine, cryptosystem.

МОДЕЛИРОВАНИЕ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ В ВЫЧИСЛИТЕЛЬНОЙ ПОДСИСТЕМЕ С ПРИМЕНЕНИЕМ АГРЕГАТИВНЫХ МОДЕЛЕЙ

А. Н. ПРИВАЛОВ
В. Л. КУЛЕШОВ

*Тульский артиллерийский
инженерный институт*

e-mail:
alexandr_prv@rambler.ru

Рассматривается применение агрегативных моделей для моделирования информационных процессов в вычислительной подсистеме, спроектированной на основе систем с распределенной обработкой данных. Приводятся результаты эксперимента.

Ключевые слова: агрегативная модель, сеть массового обслуживания, информационные процессы.

Функционирование любой тренажёрной системы на основе вычислительных систем с распределенной обработкой данных (СРОД) может быть представлено в виде совокупности взаимодействий пользователей (операторов, руководителей обучения) с системой [1]. С формальной точки зрения любое такое взаимодействие можно отобразить в виде последовательности этапов передачи и обработки информации.

Для оценки взаимодействия однородных информационных процессов (ИП) в вычислительной подсистеме тренажёрных систем предлагается подход, основанный на агрегативном описании систем с использованием особенностей их структуры.

Существо подхода заключается в выделении из рассматриваемой системы некоторой подсистемы (агрегата) и последующем ее детальном исследовании; при этом учитывается влияние остальной части системы, которая представляется в виде обобщенной сети массового обслуживания (СМО) с интенсивностью обслуживания, зависящей от числа заявок в ней.

Полагается, что эта СМО ведет себя по отношению к оставшейся подсистеме аналогично той части системы, которую она заменяет.

В общем случае СРОД достаточно сложной структуры может быть разбита на несколько подсетей, которые исследуются отдельно. При этом объединение отдельных СМО в подсети осуществляется таким образом, чтобы:

- взаимодействия элементов внутри подсети (внутренние взаимодействия) могли быть исследованы без учета;
- взаимодействий между подсетями; взаимодействия между подсетями (внешние взаимодействия) можно было анализировать без учета внутренних взаимодействий.

Указанные условия будут выполняться, если частота внутренних взаимодействий много больше частоты внешних взаимодействий. В этом случае говорят о почти разложимой системе.

Агрегативные модели дают точное решение для сетей, допускающих решение в виде произведения. Так можно показать, что точное решение может быть получено для сетей МО, описываемых моделью Гордона – Ньюэлла [2]. Агрегативный подход часто связывают с замкнутыми сетями, однако он может быть использован и в отношении открытых сетей МО.

Рассмотрим функционирование вычислительного комплекса (ВК) в режиме реального масштаба времени, ограничиваясь анализом ИП на уровне ВК. Пусть на ВК поступает пуассоновский поток заявок интенсивностью λ , а реализуемые информационно-вычислительные работы (ИВР) связаны с выполнением достаточно длинной цепочки переходов: процессор–внешнее устройство– процессор.... Время обслуживания одного запроса на устройство $j, j = \overline{1, J}$ распределено по экспоненциально-



му закону со средним $1/\lambda_j$. Здесь индекс «1» относится к узлу процессора. Вероятность выхода на узел (устройство) $j, j = \overline{2, J}$, после обслуживания в узле процессора есть θ_{1j} . Наибольший допустимый уровень мультипрограммирования в вычислительном комплексе – N .

Для исследования ИП в подобной ситуации может быть использован как аппарат открытых, так и замкнутых сетей МО. Однако в первом случае не будет учитываться ограничение на уровень мультипрограммирования, что может привести к существенному завышению оценок пропускной способности вычислительного узла, особенно при большой загрузке системы. Во втором случае получаемые оценки будут занижены, поскольку при расчетах будет приниматься максимальная пропускная способность вычислительного комплекса, соответствующая уровню мультипрограммирования N , в то время как при числе активных ИП в вычислительном узле, меньшем N , эта величина может заметно отличаться от ее максимального значения.

В данной ситуации ошибка в расчетах увеличивается с уменьшением загрузки системы. Для преодоления указанных трудностей необходима модель, отражающая реальную зависимость пропускной способности рабочей станции от числа активных ИП. Такая модель получается при использовании агрегативного подхода. В самом деле, интенсивность переходов между состояниями, связанными с занятием отдельных ресурсов (процессоров, ВУ) при реализации активного ИП, значительно больше интенсивности последовательных активизаций ИП. Поэтому можно выделить и рассмотреть отдельно замкнутую подсеть, отображающую процесс реализации в мультипрограммном режиме n активных ИП, $n = \overline{1, N}$.

На рис. 1, а схематически представлена рассматриваемая открытая сеть МО, а на рис. 1,б – образованная из нее подсеть, которую условно можно считать замкнутой.

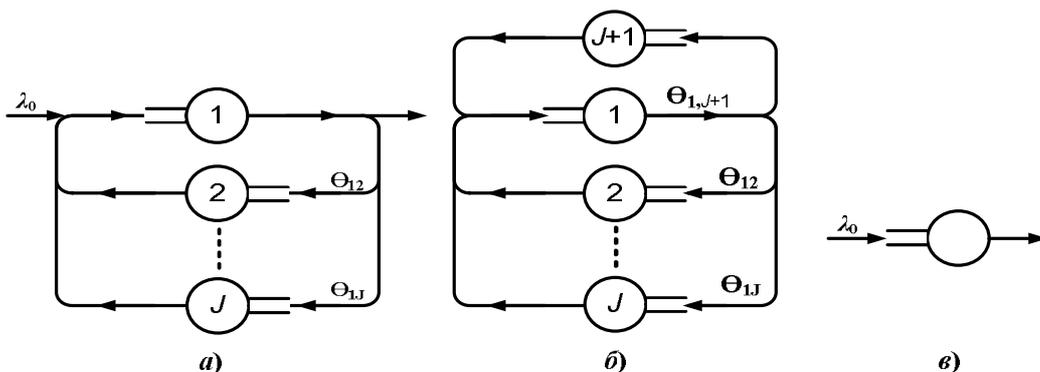


Рис. 1. Агрегирование открытой сети МО

В последнем случае для удобства вычислений введен фиктивный узел $J+1$, для которого полагаем $1/\lambda_{J+1}=0$.

Пропускная способность вычислительного комплекса определяется зависимостью (1)

$$\lambda(n) = e_{J+1} G(n-1) / G(n), n = \overline{1, N}, \tag{1}$$

т.е. является функцией от n . При этом e_{J+1} определяется из решением системы:

$$e_j = \sum_{k=1}^J e_k \theta_{kj}, \text{ а } G(K) \text{ рассчитывается по рекуррентной формуле}$$

$$G_j(k) = G_{j-1}(k) + p_j G_j(k-1), j = \overline{2, J}, k = \overline{1, K}, \tag{2}$$

где K – число заявок, циркулирующих в сети.

Рассмотренную подсеть заменим эквивалентным обслуживающим прибором, интенсивность обслуживания которым выражается следующим образом:

$$\gamma_i = \lambda(i), i = \overline{1, N}, \gamma_i = \lambda(N), \gamma > N. \quad (3)$$

Допуская теперь, что время обслуживания эквивалентным прибором можно считать распределенным по экспоненциальному закону, рассматриваемую систему представим в виде СМО типа М/М/1 с переменной интенсивностью обслуживания, задаваемой соотношением (3). Решение такой СМО может быть записано в виде:

$$P_i = \begin{cases} \lambda^i P_0 / \prod_{k=1}^i \lambda(k), 1 \leq i \leq N; \\ \lambda^i P_0 / \left\{ [\lambda(N)]^{i-N} \prod_{k=1}^N \lambda(k) \right\}, i > N; \end{cases} \quad (4)$$

$$P_0 = \left\{ 1 + \sum_{i=1}^N \left[\lambda^i / \prod_{k=1}^i \lambda(k) \right] + \lambda^{N+1} / \left[\prod_{k=1}^N \lambda(k) (\lambda(N) - \lambda) \right] \right\}^{-1}.$$

Рассчитав стационарные вероятности состояний (4), можно получить требуемые характеристики реализации ИП в системе. Например, среднее число заявок в системе выразится как

$$n_1 = P_0 \left\{ \left[\sum_{i=1}^N i \lambda^i / \prod_{k=1}^i \lambda(k) \right] + \frac{1}{\prod_{k=1}^N \lambda(k)} \frac{\lambda^{(N+1)}}{\lambda(N) - \lambda} \left[N + 1 + \frac{\lambda}{\lambda(N) - \lambda} \right] \right\}.$$

Отсюда среднее время реакции рабочей станции определяется на основании формулы Литтла: $v_1 = n_1 / \lambda$.

Рассмотрим пример, демонстрирующий основные особенности расчетных схем рассматриваемой модели.

Пусть необходимо оценить пропускную способность ВК, схематически изображенного на рис. 2, где показана его структура: процессор (узел 1) и три внешних запоминающих устройства (ВЗУ) (узлы 2, 3, 4). Каналы обмена данными в ВК не являются узким местом, так что очереди образуются лишь к процессору и ВЗУ. Объем оперативной памяти ВК такой, что возможна одновременная реализация лишь двух ИП (работа в двухпрограммном режиме). Параметры соответствующих ИП заданы; время одного обслуживания любым из устройств можно считать распределенным по экспоненциальному закону с интенсивностями $\mu_1 = 5, \mu_j = 1, j = \overline{2, 4}$; после завершения обслуживания процессором с равной вероятностью 0,3 выдаются запросы на устройства 2, 3, 4 и с вероятностью 0,1 ИП завершается.

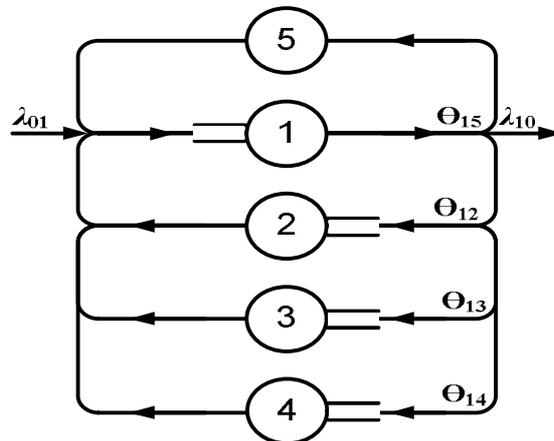


Рис. 2. Пример структуры ВК



При сравнении оценок основных характеристик реализации, полученных на основе агрегативного подхода, будем полагать, что пропускная способность рабочей станции определяется как $\lambda(N)$. Для простоты будем полагать, что предельный уровень мультипрограммирования $N=2$. Результаты расчетов представлены в табл. 1, где приведены значения среднего числа ИП в системе (N_1) и среднего времени реакции рабочей станции (U_1).

Таблица 1

Результаты вычислительного эксперимента с различными типами моделей

λ	Замкнутая модель		Агрегативная модель		Открытая модель	
	N_1	U_1	N_1	U_1	N_1	U_1
0,05	0,5	10	0,68	13,4	0,65	13
0,1	2,0	20	2,3	23	0,69	6,9
0,14	13	93	14,3	102	2,6	18
0,2	-	-	-	-	4,8	24
0,33	-	-	-	-	299	906

Анализируя эти данные, можно обратить внимание, что при малой загрузке рабочей станции (малых значениях входящего потока заявок на ИВР) результаты расчетов, полученные на основе агрегативного подхода и в результате анализа открытой сети МО, весьма близки. Это вполне естественно, поскольку в данном случае система работает с недогрузкой и ограничение на уровень мультипрограммирования почти не проявляется. При большой загрузке системы ($\lambda \approx 0,14$) оценки, полученные на основе агрегативного подхода, близки к соответствующим оценкам, полученным путем анализа замкнутой сети МО $[\gamma_i = \lambda(N), i = \overline{1, N}]$. Этот факт также легко объясняется: в условиях большой загрузки система почти постоянно работает при максимальном уровне мультипрограммирования, что соответствует наивысшей пропускной способности. Таким образом, использование агрегативного подхода в данном случае позволяет наиболее полно отобразить реальные условия функционирования ВК.

Аналогичным образом агрегатирование применимо и при исследовании замкнутых систем. Рассмотрим в качестве примера ВК, обслуживающий некоторое число M абонентов, каждый из которых обращается к рабочей станции с заявками на ИВР.

Время между последовательными обращениями абонента к рабочей станции распределено по экспоненциальному закону с параметром λ . Причем новая заявка может быть сформирована абонентом лишь после получения ответа на предыдущую. Следовательно, при наличии в рабочей станции (на обслуживании и в очереди) i заявок поток заявок от всех абонентов будет иметь пуассоновский характер с интенсивностью

$$\lambda_i = (M - i)\lambda, i = \overline{0, M}. \tag{5}$$

Пусть наибольший уровень мультипрограммирования для рабочей станции есть N .

Теперь, рассматривая рабочую станцию как замкнутую сеть МО, можем, как и ранее, определить величины $\lambda(n)$ в соответствии с (1). Предположим, что при работе рабочей станции с уровнем мультипрограммирования i поток завершения решений задач является пуассоновским с параметром γ_i , причем γ_i определяется в соответствии с (3).

В этом случае из анализа замкнутой сети с двумя приборами с переменной интенсивностью обслуживания, определяемой выражениями (3) и (5), получаем соотношения для расчета вероятностей P_i (числа заявок в ВК):

$$P_i = \lambda^i M! P_0 / \left[(M - i)! \prod_{k=1}^i \gamma_k \right], i = \overline{1, M}; \tag{6}$$

$$P_0 = \left[1 + \sum_{r=1}^M \frac{\lambda^r M!}{\prod_{k=1}^r \gamma_k (M-r)!} \right]^{-1}.$$

Отсюда среднее время реакции системы выразится как

$$v_1 = M \left[\sum_{i=1}^N P_i \lambda(i) + \lambda(N) \sum_{i=N+1}^M P_i \right]^{-1} - 1/\lambda. \quad (7)$$

Расчеты с использованием (7) показывают, что учет ограничений на уровень мультипрограммирования при применении агрегативного подхода позволяет более точно оценивать основные характеристики функционирования мультипрограммных ВК при конечном числе источников заявок. Особенно это относится к условиям большой загрузки ВК, когда при игнорировании указанных ограничений пропускная способность рабочей станции может быть завышена в несколько раз [1].

В [2] приведена одна из наиболее общих методик реализации агрегативного подхода, основанная на итеративном вычислении характеристик замкнутых сетей МО с общим распределением времени обслуживания заявок в порядке их поступления.

Методика дает точные результаты для сетей, удовлетворяющих условиям локального баланса.

Рассмотрим замкнутую сеть МО, в которой циркулирует I заявок и имеется J узлов обслуживания. Время обслуживания в узле $j, j = \overline{1, J}$ распределено по некоторому закону $B_j(t)$ со средним b_j . Заданы вероятности θ_{jk} перехода в узел k после обслуживания в узле j . Это позволяет рассчитать величины e_j путем решения системы уравнений вида:

$$e_j = \sum_{k=1}^J e_k \theta_{kj}.$$

Такая сеть именуется сетью A . В соответствии с методикой [2] строится некоторая последовательность сетей A_0, A_1, \dots аппроксимирующих сеть A . При этом сеть A_0 получается из A путем перехода к экспоненциальному распределению времени обслуживания в узлах с тем же значением среднего времени обслуживания. Сети $A_r, r=1, 2, \dots$ отличаются от A_0 лишь значением среднего времени обслуживания в узлах $1/\gamma_j(r)$.

Расчетная процедура методики, определяющая переход от A_k к A_{k+1} и условия завершения вычислений, заключается в реализации некоторой последовательности шагов, на первом шаге которой средние значения времени обслуживания в узлах сети A_0 полагаются равными соответствующим величинам в сети A , а распределение времени обслуживания – экспоненциальным. Далее для каждого узла j строится подсистема, содержащая все узлы сети, кроме j . При этом исследуется сеть с двумя узлами, так что выделенная подсистема отображается некоторой СМО, эквивалентной ей в смысле воздействия на узел j (для простоты анализа производится перенумерация: узел j нумеруется как 1, а эквивалентный прибор – как 2).

Для сети с двумя узлами известными методами рассчитываются значения среднего числа заявок в узле $j(n_{j1})$ и среднего числа заявок, обслуживаемых в единицу времени (λ_j), а также приведенное среднее число заявок, обслуживаемых в узле j в единицу времени, $\gamma_j = \lambda_j / e_j$. После этого осуществляется проверка сходимости величин n_{j1} и γ_j на основе соотношений

$$\left| I - \sum_{j=1}^J n_{j1} \right| \leq \varepsilon I; \quad \left| \gamma_j - \sum_{k=1}^J \gamma_k / J \right| \leq \varepsilon \gamma_j. \quad (8)$$



Здесь ε – малая величина, задающая порядок допустимой ошибки (например, $\varepsilon=0,01$). Выполнение условий (8) на некоторой итерации указывает на приемлемость полученного отображения A с помощью A_0 . В противном случае осуществляется корректировка средних значений времени обслуживания в узлах с помощью выражений (в зависимости от выполнения того или иного из условий (8) и производится следующая $(\gamma+1)$ -я итерация.

Описанная методика использовалась для расчета характеристик некоторых замкнутых сетей, а результаты расчетов сравнивались с результатами имитационного моделирования на ряде экспериментов [2]. В одном из этих экспериментов рассматривалась сеть, в которой число узлов варьировалось от 2 до 5, число заявок – также от 2 до 5, а времена обслуживания полагались распределенными по законам: экспоненциальному, Эрланга 2-го порядка и гиперэкспоненциальному. Результаты эксперимента показали, что расхождение в вычислении значения загрузки узлов не превышает 0,05, а в вычислении среднего числа заявок в узле в основном находится в пределах 0,05. В другом эксперименте исследовались сети из 6-7 узлов, число заявок в которых варьировалось от 2 до 12, а времена обслуживания полагались распределенными по законам гиперэкспоненциальному, Эрланга порядка 2, 4 и 6, экспоненциальному, а также считались постоянными. В результате наблюдалось хорошее согласование вычисляемых характеристик (в пределах 0,05 для загрузки узлов и 0,05 для среднего числа заявок в узле). Расчет тех же сетей точными методами в предположении экспоненциального времени обслуживания давал более значительные расхождения в оценке указанных характеристик.

Литература

1. Привалов А.Н., Моделирование информационных процессов в вычислительной подсистеме тренажёрных систем специального назначения [Текст] / А.Н. Привалов. – Тула: Изд-во ТулГУ, 2009. – 215 с.
2. Балыбердин, В.А., Оптимизация информационных процессов в распределённых системах обработки данных [Текст] / В.А. Балыбердин, А.М. Белевцев, О.А. Степанов. – М.: Технология, 2002. – 280 с.

PERFORMANCE EVALUATION OF VIRTUAL TRAINING SYSTEM WITH DISTRIBUTED DATA PROCESSING

A. N. PRIVALOV
V. L. KULESHOV

*Tula artillery
engineering institute*

*e-mail:
alexandr_prv@rambler.ru*

Application is considered aggregate models for modeling of information processes in computing subsystem, designed on basis of systems with distributed data processing. Experimental results are resulted.

Key words: Aggregate model, queuing network, information processes.

РАЗВИТИЕ МЕТОДОВ БЫСТРОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ С ПОМОЩЬЮ ФИЛЬТРОВ ДОБЕШИ

Н. И. ЧЕРВЯКОВ
Ю. В. КОНДРАШОВ

*Ставропольский
военный институт
связи ракетных войск*

e-mail: yrii-fifa83@mail.ru

В работе показана возможность построения высокоэффективных вейвлет – преобразований ортогональными вейвлет – фильтрами Добеши по алгоритму Малла на базе свертки, вычисленной в системе остаточных классов.

Ключевые слова: вейвлет – преобразование, вейвлет – фильтр, вейвлет – функция, алгоритм Малла, свертка, цифровая обработка сигналов, система остаточных классов.

В настоящее время вейвлет – преобразование широко применяется в задачах обработки и кодирования сигналов и изображений самой различной природы (речь, спутниковые изображения, рентгенограммы внутренних органов и др.), распознавания образов, при изучении свойств поверхностей кристаллов и нанобъектов и во многих других случаях.

Появление в 1988 году ортогональных вейвлет-фильтров Добеши или так называемых фильтров с компактным носителем в значительной мере усилило интерес к вейвлет-анализу, т.к. открылись новые возможности не только для теоретического, но и для практического применения вейвлет-преобразования.

Важно отметить то, что вейвлет-фильтры Добеши строятся, исходя из критерия длины фильтров и, следовательно, являются фильтрами с конечным числом коэффициентов [1]. Вейвлет-функции $\psi(t)$ фильтров Добеши принято обозначать литерой D с добавлением цифры, соответствующей длине вейвлет-фильтра Добеши, т.е. $D2$, $D4$, $D6$ и т.д.

Пусть даны два фильтра h и g с ненулевыми элементами [2]:

$$\begin{aligned} h_0 &= \frac{1 + \sqrt{3}}{8}, & h_1 &= \frac{3 + \sqrt{3}}{8}, & h_2 &= \frac{3 - \sqrt{3}}{8}, & h_3 &= \frac{1 - \sqrt{3}}{8}; \\ g_0 &= \frac{1 - \sqrt{3}}{8}, & g_1 &= -\frac{3 - \sqrt{3}}{8}, & g_2 &= \frac{3 + \sqrt{3}}{8}, & g_3 &= -\frac{1 + \sqrt{3}}{8}. \end{aligned} \quad (1)$$

Отметим соотношения между коэффициентами этих фильтров [2]:

$$\begin{aligned} h_0 + h_1 + h_2 + h_3 &= 1; \\ g_0 + g_1 + g_2 + g_3 &= 0; \\ g_0 &= h_3, \quad g_1 = -h_2, \quad g_2 = h_1, \quad g_3 = -h_0; \\ h_0 + h_2 &= h_1 + h_3 = 1/2, \quad 2h_2 = h_1 + 3h_3. \end{aligned} \quad (2)$$

Найдём передаточные функции $H(z)$ и $G(z)$ в z -представлении:

$$H(z) = h_0 + h_1 z + h_2 z^2 + h_3 z^3; \quad (3)$$

$$\begin{aligned} G(z) &= g_0 + g_1 z + g_2 z^2 + g_3 z^3 = h_3 - h_2 z + h_1 z^2 - h_0 z^3 = \\ &= -z^3 (-h_3 z^{-3} + h_2 z^{-2} - h_1 z^{-1} + h_0) = -z^3 H(-z^{-1}). \end{aligned} \quad (4)$$

Таким образом, мы получили

$$G(z) = -z^3 H(-z^{-1}). \quad (5)$$



Для восстановления сигнала требуются дополнительные фильтры $\tilde{H}(z)$ и $\tilde{G}(z)$. Определим их как сопряжённые квадратурные фильтры по формулам:

$$\tilde{H}(z) = H(z^{-1}), \quad \tilde{G}(z) = G(z^{-1}) = -z^{-3}H(-z) \quad (6)$$

Тогда второе соотношение $\tilde{H}(z)H(-z) + \tilde{G}(z)G(-z) = 0$ из (3-4) выполняется. Первое соотношение $\tilde{H}(z)H(z) + \tilde{G}(z)G(z) = 1$ принимает вид:

$$H(z^{-1})H(z) + G(z^{-1})G(z) = 1. \quad (7)$$

Вернёмся к частотной переменной $z = e^{-i\omega}$. Поскольку коэффициенты $\{h_n\}$ – вещественные, то $H(z^{-1}) = \overline{H(\omega)}$. Поэтому последнее соотношение принимает вид $|H(\omega)|^2 + |G(\omega)|^2 = 1$.

Найдём коэффициенты фильтров восстановления $\tilde{H}(z)$ и $\tilde{G}(z)$ из их определения $\tilde{H}(z) = H(z^{-1})$, $\tilde{G}(z) = G(z^{-1})$:

$$\begin{aligned} \tilde{h}_{-3} &= \frac{1-\sqrt{3}}{8}, & \tilde{h}_{-2} &= \frac{3-\sqrt{3}}{8}, & \tilde{h}_{-1} &= \frac{3+\sqrt{3}}{8}, & \tilde{h}_0 &= \frac{1+\sqrt{3}}{8}; \\ \tilde{g}_{-3} &= -\frac{1+\sqrt{3}}{8}, & \tilde{g}_{-2} &= \frac{3+\sqrt{3}}{8}, & \tilde{g}_{-1} &= -\frac{3-\sqrt{3}}{8}, & \tilde{g}_0 &= \frac{1-\sqrt{3}}{8}. \end{aligned} \quad (8)$$

Таким образом, метод одномерного дискретного вейвлет – преобразования (ДВП) N -го порядка последовательности x_n определяется следующими рекуррентными соотношениями:

$$\begin{aligned} a_n^{(i)} &= \sum_{k=0}^{N-1} g_k a_{2n-k}^{(i-1)} \quad i = 1, 2, \dots, J; \\ d_n^{(i)} &= \sum_{k=0}^{N-1} h_k a_{2n-k}^{(i-1)} \quad a_n^{(0)} \equiv x_n, \end{aligned} \quad (9)$$

где $a_n^{(i)}$ и $d_n^{(i)}$ являются аппроксимирующими и детализирующими коэффициентами i -го уровня, а g_k и h_k ($k = 0, 1, \dots, N-1$) – коэффициенты низкочастотного и высокочастотного анализирующих фильтров, соответственно.

С другой стороны, сигнал x_n может быть восстановлен по коэффициентам $\{a_n^{(J)}, d_n^{(J)}, d_n^{(J-1)}, \dots, d_n^{(1)}\}$ путём последовательной итерации по формулам:

$$\bar{a}_m^{(i-1)} = \begin{cases} \sum_{k=0}^{N/2-1} \bar{g}_{2k} \bar{a}_{\frac{m}{2}-k}^{(i)} + \sum_{k=0}^{N/2-1} \bar{h}_{2k} \bar{d}_{\frac{m}{2}-k}^{(i)}, & m \text{ чётное} \\ \sum_{k=0}^{N/2-1} \bar{g}_{2k+1} \bar{a}_{\frac{m-1}{2}-k}^{(i)} + \sum_{k=0}^{N/2-1} \bar{h}_{2k+1} \bar{d}_{\frac{m-1}{2}-k}^{(i)}, & m \text{ нечётное} \end{cases}, \quad (10)$$

где \bar{g}_k и \bar{h}_k являются коэффициентами низкочастотного и высокочастотного синтезирующих фильтров, соответственно.

Для того, чтобы восстановленный сигнал соответствовал исходному, должны быть соответствующим образом подобраны анализирующий (раскладывающий) и синтезирующий (собирающий) фильтры.

Для вейвлет-преобразования функции $f(x)$ необходимо вычислить серию коэффициентов $\{a_n, d_n, d_{n-1}, \dots, d_1\}$, где a_n – аппроксимация функции, d_i – детализирующие коэффициенты функции, $i = 1, \dots, n$. Каждый коэффициент находится интегрированием (11, 12):

$$a_{J-N,k} = (f, \varphi_{J-N,k}) = \int_R f(x) \overline{\varphi_{J-N,k}(x)} dx; \tag{11}$$

$$d_{J-m,k} = (f, \psi_{J-m,k}) = \int_R f(x) \overline{\psi_{J-m,k}(x)} dx, \quad m = 1, 2, \dots, N. \tag{12}$$

Возникает проблема вычисления большого количества интегралов с необходимой точностью. Следует также учитывать, что при высоком уровне разрешения J носители функций $\varphi_{J,k}(x)$ и $\psi_{J,k}(x)$ становятся малыми порядка $1/2^J$.

Быстрое вейвлет-преобразование, предложенное Мала позволяет решить эту проблему. Алгоритм Малла даёт возможность вычислять коэффициенты вейвлет-разложения без интегрирования, используя алгебраические операции на основе свёртки:

$$\begin{aligned} a_n^{(i)} &= \sum_{k=0}^{N-1} g_k a_{2n-k}^{(i-1)} \quad i = 1, 2, \dots, J; \\ d_n^{(i)} &= \sum_{k=0}^{N-1} h_k a_{2n-k}^{(i-1)} \quad a_n^{(0)} \equiv x_n, \end{aligned} \tag{13}$$

где $a_n^{(i)}$ и $d_n^{(i)}$ являются аппроксимирующими и детализирующими коэффициентами i -го уровня, а g_k и h_k ($k = 0, 1, \dots, N - 1$) – коэффициенты низкочастотного и высокочастотного анализирующих фильтров, соответственно; x_n – исходный сигнал; N – порядок фильтра.

Эти равенства обеспечивают быстрые алгоритмы вычисления вейвлет-коэффициентов (каскадные алгоритмы, алгоритмы Малла). Термин «быстрые» означает не только, что в (13) используются более быстрые алгебраические процедуры, но и то, что при каждом преобразовании общее число новых коэффициентов не увеличивается в два раза, а остаётся прежним.

Схема разложения сигнала по алгоритму Малла приведена на рис. 1.

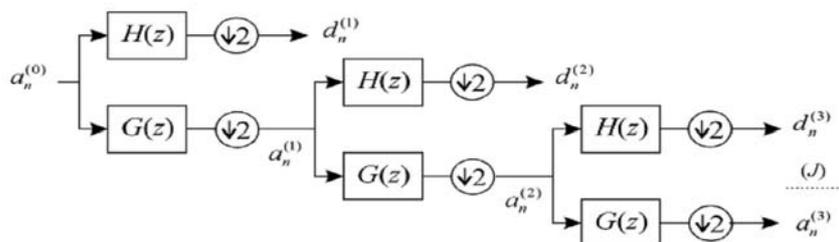


Рис. 1. Последовательность получения вейвлет-коэффициентов третьей октавы; $H(z)$ и $G(z)$, соответственно, высокочастотные и низкочастотные анализирующие фильтры в z -представлении

Единственное отличие фильтрации в алгоритме Малла от классического КИХ-фильтра, задаваемого уравнением $y(k) = \sum_{i=0}^m b_i x(k-i)$ [3], заключается в том, что значения фильтруемого ряда выбираются через один – индекс $2n - k$ в $a_{2n-k}^{(i-1)}$. Это и есть децимация $2 \downarrow$ – исключение из обработки каждого второго элемента.

Для двумерных сигналов – изображений – алгоритм разложения аналогичен тому, что применяется в одномерном случае (13). Пусть $\varphi(x)$ – масштабирующая вейвлет-функция и $\psi(x)$ – материнский вейвлет. Как известно, они порождают ба-

численные функции $\varphi_{J,n}(x)$ и $\psi_{J,k}(x)$. Двумерный сигнал $a(n_1, n_2)$ раскладывается по базисным в $L^2(\mathbf{R}^2)$ функциям $\varphi_{J,n}(x)\varphi_{J,m}(y)$, $\varphi_{J,n}(x)\psi_{J,m}(y)$, $\psi_{J,n}(x)\varphi_{J,m}(y)$ и $\psi_{J,n}(x)\psi_{J,m}(y)$. Соответствующие коэффициенты принято называть следующим образом.

Аппроксимирующие коэффициенты $a^{(J)}(n_1, n_2)$ получаются как коэффициенты разложения по вейвлет-базису $\varphi_{J,n}(x)\varphi_{J,m}(y)$. На рис. 2 (а) показано распределение пикселей после пошаговой обработки исходного изображения банком фильтров.

Горизонтальные детализирующие коэффициенты $d_2^{(J)}(n_1, n_2)$ получаются как коэффициенты разложения по вейвлет-базису $\varphi_{J,n}(x)\psi_{J,m}(y)$.

Вертикальные детализирующие коэффициенты $d_1^{(J)}(n_1, n_2)$ получаются как коэффициенты разложения по вейвлет-базису $\psi_{J,n}(x)\varphi_{J,m}(y)$.

Диагональные детализирующие коэффициенты $d_3^{(J)}(n_1, n_2)$ получаются как коэффициенты разложения по вейвлет-базису $\psi_{J,n}(x)\psi_{J,m}(y)$.

Схема разложения сигнала $a^0(n_1, n_2)$ изображена на рис. 2 (б).

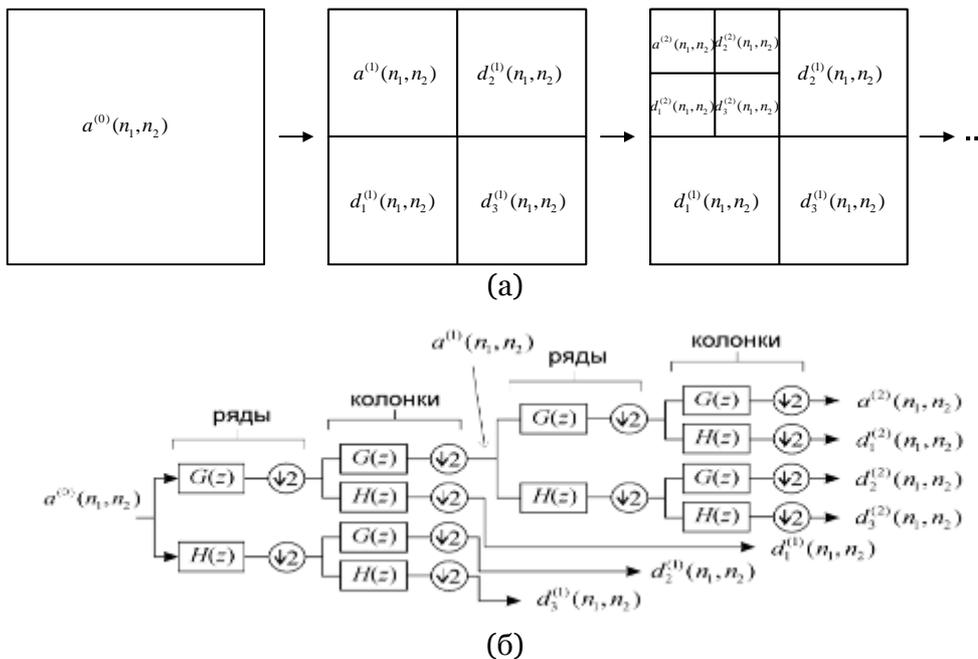


Рис. 2. Последовательность получения вейвлет-коэффициентов третьей октавы для двумерного сигнала: (а) –распределение пикселей после пошаговой обработки исходного изображения банком фильтров; (б) – в виде последовательности фильтров; $H(z)$ и $G(z)$, соответственно, высокочастотные и низкочастотные анализирующие фильтры в z -представлении

В аналитическом виде разложение двумерного сигнала фильтрами можно записать следующим образом:

$$\begin{aligned}
 a^{(i+1)}(n_1, n_2) &= \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} g(k_1) g(k_2) a^{(i)}(2n_1 - k_1, 2n_2 - k_2); \\
 d_1^{(i+1)}(n_1, n_2) &= \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} g(k_1) h(k_2) a^{(i)}(2n_1 - k_1, 2n_2 - k_2); \\
 d_2^{(i+1)}(n_1, n_2) &= \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} h(k_1) g(k_2) a^{(i)}(2n_1 - k_1, 2n_2 - k_2); \\
 d_3^{(i+1)}(n_1, n_2) &= \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} h(k_1) h(k_2) a^{(i)}(2n_1 - k_1, 2n_2 - k_2).
 \end{aligned} \tag{14}$$

В качестве собственно фильтров могут использоваться фильтры Добеши D4 четвёртого порядка. Вейвлеты Добеши являются вейвлетами с компактным носителем, что обеспечивает хорошие свойства приближения вейвлет-разложений. Они не имеют эксплицитного (явного) выражения, а задаются коэффициентами фильтрации. Анализирующие (разлагающие) высокочастотные (h) и низкочастотные (g) коэффициенты фильтра Добеши D4 задаются следующими коэффициентами [2]:

$$\begin{aligned}
 h_0 &= \frac{1 + \sqrt{3}}{4\sqrt{2}}, & h_1 &= \frac{3 + \sqrt{3}}{4\sqrt{2}}, & h_2 &= \frac{3 - \sqrt{3}}{4\sqrt{2}}, & h_3 &= \frac{1 - \sqrt{3}}{4\sqrt{2}}; \\
 g_0 &= \frac{1 - \sqrt{3}}{4\sqrt{2}}, & g_1 &= -\frac{3 - \sqrt{3}}{4\sqrt{2}}, & g_2 &= \frac{3 + \sqrt{3}}{4\sqrt{2}}, & g_3 &= -\frac{1 + \sqrt{3}}{4\sqrt{2}}.
 \end{aligned} \tag{15}$$

Графики вейвлетов Добеши D4 (db4) в среде MATLAB можно увидеть следующим образом (рис. 3):

```

[phi,psi,x]=wavfun('db4',10);
subplot(121); plot(x,phi);
title('y=\phi(x)'); axis square; grid on;
subplot(122); plot(x,psi);
title('y=\psi(x)'); axis square; grid on;

```

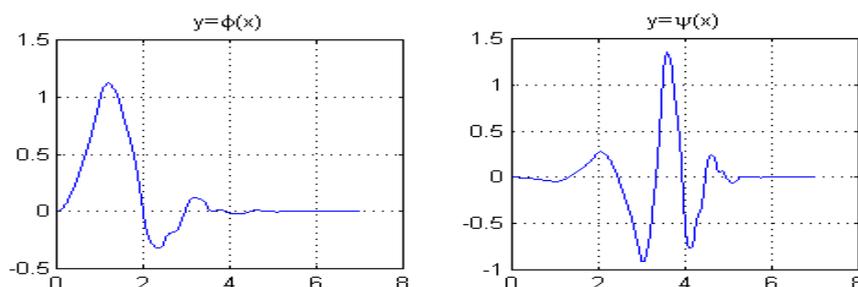


Рис. 3. Масштабирующая вейвлет-функция и материнский вейвлет Добеши D4

С целью повышения скорости вычисления свертки (13) предлагается её вычислять в системе остаточных классов, тогда выбирая модуль p_j свертка может быть выражена как:

$$\begin{aligned}
 a_n^{(i)} &= \left| \sum_{k=0}^{N-1} g_k a_{2n-k}^{(i-1)} \right|_{p_j} \Big|_{p_i} \quad i=1, 2, \dots, J, \\
 d_n^{(i)} &= \left| \sum_{k=0}^{N-1} h_k a_{2n-k}^{(i-1)} \right|_{p_j} \Big|_{p_i} \quad a_n^{(0)} \equiv x_n.
 \end{aligned} \tag{16}$$



Система остаточных классов и модулярные вычисления являются практически идеальным инструментом реализации линейной свертки, поскольку операции сложения, вычитания и умножения выполняются очень просто, а именно, если даны два числа A и B , представленные в системе остаточных классов (с набором взаимно простых оснований m_1, m_2, \dots, m_L) следующим образом:

$$\begin{aligned} A &= (a_1, a_2, \dots, a_L) : A \equiv a_1 \pmod{m_1}, A \equiv a_2 \pmod{m_2}, \dots, A \equiv a_L \pmod{m_L}, \\ B &= (b_1, b_2, \dots, b_L) : B \equiv b_1 \pmod{m_1}, B \equiv b_2 \pmod{m_2}, \dots, B \equiv b_L \pmod{m_L} \end{aligned} \quad (17)$$

то

$$A \pm B = (a_1, a_2, \dots, a_L) \pm (b_1, b_2, \dots, b_L) = (|a_1 \pm b_1|_{m_1}), (|a_2 \pm b_2|_{m_2}), \dots, (|a_L \pm b_L|_{m_L}) \quad (18)$$

и

$$A \cdot B = (a_1, a_2, \dots, a_L) \cdot (b_1, b_2, \dots, b_L) = (|a_1 \cdot b_1|_{m_1}), (|a_2 \cdot b_2|_{m_2}), \dots, (|a_L \cdot b_L|_{m_L}) \quad (19)$$

Математические модели (17-19) вычисляются на основе использования нейронных сетей конечного кольца [3], число которых определяется рядом каналов по числу оснований, работающих независимо друг от друга и параллельно во времени. Если каждую нейронную сеть конечного кольца отождествить с отдельным основанием системы остаточных классов, то образованная совокупность каналов будет представлять собой арифметическое устройство выполняющее с большой эффективностью вейвлет – преобразование сигналов.

Итак, система остаточных классов является наиболее подходящей технологией для реализации высокоэффективного вейвлет – преобразования для задач цифровой обработки сигналов.

Для эффективной реализации операций вейвлет – преобразования по алгоритму Малла на основе свертки предлагается использовать математическую модель вычислительного объекта, оперирующую числами, представленными в системе остаточных классов.

Литература

1. Добеши И. Десять лекций по вейвлетам. – М.: Ижевск: РХД, 2001.
2. Daubechies I. The Wavelet Transform, Time-Frequency Localization and Signal Analysis // IEEE Trans. Inform. Theory, 1990, № 5. P. 961-1005.
3. Червяков Н. И., Сахнюк П. А., Шапошников А. В., Макоха А. Н. Под редакцией А. И. Галущкина. Учебное пособие для ВВУЗов. – М.: Радиотехника, 2003. – 272 с.
4. Сергиенко А.Б. Цифровая обработка сигналов. – СПб.: Питер, 2003. – 604 с.
5. Астафьева Н.М. Вейвлет-анализ: основы теории и некоторые приложения // Успехи физических наук, 1996, № 11. С. 1145-1170.
6. Goswami J.C., Chan A.K. Fundamentals of Wavelets. Theory, Algorithms, and Applications. Wiley, 2000. – 306 p.
7. Прокис Дж. Цифровая связь. – М.: Радио и связь, 2000. – 800 с.
8. Желудев В.А. О вейвлетах на базе периодических сплайнов // Докл. РАН, 1994, № 1. С. 9-13.
9. Новиков Л.В. Основы вейвлет-анализа сигналов. Учебное пособие. – СПб., 1999. – 152 с.

DEVELOPMENT OF METHODS FAST WAVELET-TRANSFORMATION BY MEANS OF FILTERS DOBESHI

N. I. CHERVYKOV
Y. V. KONDRASHOV

*Stavropol military
institute of communication
of rockets armies*

e-mail: yrii-fifa83@mail.ru

In work possibility of construction highly effective wavelet – transformations orthogonal wavelet – filters Dobeshi on algorithm of Mull on the basis of the convolution calculated in system of residual classes is shown.

Key words: wavelet – transformation, wavelet – the filter, wavelet – functions, algorithm of Mull, convolution, system of residual classes, digital processing of signals.

НОВЫЙ ПОДХОД К АВТОМАТИЗИРОВАННОМУ ВЫЯВЛЕНИЮ ШАБЛОНОВ В ТЕЛЕМЕТРИЧЕСКИХ СИГНАЛАХ НА ОСНОВЕ ДЕКОМПОЗИЦИИ НА ЭМПИРИЧЕСКИЕ МОДЫ

Д. М. Клионский¹

Н. И. Орешко²

В. В. Геппенер¹

¹⁾ Санкт-Петербургский
государственный
электротехнический
университет “ЛЭТИ”
им. В.И. Ульянова-Ленина

²⁾ Научно-инженерный
центр Санкт-
Петербургского
электротехнического
университета

e-mail: klio2003@list.ru;

reshko@nic.spb.ru;

geppener@mail.ru

В статье представлен новый метод обработки нестационарных сигналов сложной внутренней структуры – декомпозиция на эмпирические моды. В силу наличия важнейших для практики свойств – локальности и адаптивности – появляется возможность высокоточного решения многочисленных задач, которые можно условно разделить на три группы: предварительная обработка, анализ внутренней структуры и интеллектуальный анализ сигналов. Рассматривается подход к анализу внутренней структуры сигналов с использованием теории регрессии и методов дискриминантного анализа. Важнейшим преимуществом является возможность автоматизации нового подхода. Рассматриваемая задача включает в себя выделение шумовой компоненты, характерных сигнальных компонент (шаблонов) и тренда. Описан способ классификации извлекаемых из сигналов эмпирических мод. Эффективность подхода проиллюстрирована на примере широко используемых в телеметрии мультигармонических сигналов.

Ключевые слова: адаптивность, эмпирическая мода (ЭМ), декомпозиция на эмпирические моды (ДЭМ), ЭМ-шаблон, шумовая ЭМ, трендовая ЭМ, классификационная статистика, регрессионный подход, мультигармонический сигнал.

Актуальность технологии обработки сигналов на основе ДЭМ

Большинство реальных сигналов, рассматриваемых в самых разных областях знаний и в рамках различных научных направлений (обработка речи, телеметрии, биомедицинских, акустических, метеорологических, сейсмических и др. типов сигналов), чаще всего являются нестационарными по своей природе, что проявляется в непостоянстве их отдельных характеристик во времени. У многих из на сегодняшний день существующих подходов при практическом использовании наблюдается ряд существенных недостатков. Например, широко распространенный классический анализ Фурье из-за своей относительной простоты вычислений (формулы Парсеваля, Планшереля, наличие быстрых вычислительных алгоритмов) практически сразу после своего появления стал доминировать над всеми другими методами анализа сигналов и применяться ко всем типам сигналов. Несмотря на то, что преобразование Фурье [4] применяется при условиях, обычно выполняемых на практике (условия Дирихле, абсолютной интегрируемости), существует несколько существенных ограничений, накладываемых на сигналы, для которых вычисляется это преобразование. Суть этих ограничений состоит в том, что сигналы должны быть *строго периодическими* функциями или содержать т.н. *повторяющиеся паттерны* (фрагменты одной общей природы, при этом не обязательно являющиеся идентичными), т.к. в противном случае анализ в частотной области даст неверную интерпретацию результатов. Также необходимо, чтобы сигналы обладали свойством стационарности по отношению к ряду характеристик (среднее значение и пр.).

Однако анализ Фурье неэффективен при исследовании сигналов с изменяющимся частотным содержанием (например, ЛЧМ-сигнал, из Фурье-спектра которого нельзя сделать вывода о линейном законе изменения частоты), т.к. тригонометрический базис содержит функции с не изменяющейся во времени частотой, в то время как сигналы могут включать в себя компоненты, занимающие разные частотные полосы. Кроме того, возникают сложности при необходимости определения основных



диапазонов концентрации энергии (на основе классической Фурье-периодограммы и ее модификаций), которая рассеивается по достаточно широкому диапазону частот, создавая порой ложную картину распределения энергии по частотам. Функции тригонометрического базиса (и некоторых других часто используемых базисов) не обладают временной локализацией (не обеспечивают качественного временного разрешения), что весьма существенно для нестационарных сигналов, т.к. требуется локализация особенностей, моментов резких изменений.

Получившие в последнее время широкое распространение частотно-временные методы (класс Козна и пр.) [4] дают заметные преимущества по сравнению с классическим анализом Фурье, однако с их помощью невозможно напрямую, без привлечения дополнительных средств, получить мультимасштабное представление сигнала (в виде совокупности компонент с разным разрешением), которое требуется для понимания его внутренней структуры. Для повышения точности и достоверности анализа нестационарных сигналов, включая решение важнейшей и первичной задачи *предварительной обработки*, необходим специальный подход, обладающий свойством *адаптивности* к каждому конкретному рассматриваемому сигналу. Из известных на сегодняшний день подходов дискретное вейвлет-преобразование на основе схемы диадического банка фильтров [4] частично удовлетворяет данному требованию, а также многим другим, представляющим исключительную практическую ценность, а именно: хорошо разработанный математический аппарат, наличие быстрых вычислительных алгоритмов (для повышения оперативности вычислений), широкий класс решаемых задач. В силу особенностей конструирования вейвлет-базиса (на основе масштабных растяжений и сдвигов вдоль временной оси материнской функции-вейвлета) появляется возможность *адаптивно* обрабатывать сигналы путем довольно точного учета локальных временных особенностей. Но *главная проблема*, порой затрудняющая их эффективное практическое применение, – наличие большого многообразия и *неочевидность выбора* вейвлета для решения конкретной задачи. Как один из выходов можно предложить перебор вейвлет-функций (если нет других эвристических критериев выбора базиса), однако эта процедура может оказаться весьма затратной по времени и количеству вычислений. В целом следует признать, что вейвлет-преобразование в настоящее время играет одну из ведущих ролей в обработке сигналов в силу наличия большого многообразия специально разработанных базисов и применимости к различным важным практическим задачам [4]. Подчеркнем еще раз, что основная трудность применения вейвлет-преобразования – необходимость наличия априорных сведений (о виде базиса, его особенностях и свойствах и пр.), что зачастую вызывает трудности.

В настоящей статье будет подробно рассмотрена новая высоко адаптивная технология анализа и обработки сигналов – **декомпозиция на эмпирические моды (ДЭМ)** [1-3]. ДЭМ, обладая практически всеми известными достоинствами вейвлет-разложения (наиболее существенный недостаток ДЭМ – отсутствие полной теоретической базы – проблема, над которой сейчас ведется активная работа), при этом свободен от данного недостатка. Другими словами, разложение по этой системе функций с целью последующего анализа проводится с учетом *локальных особенностей* (таких как экстремумы и нули сигнала) и *внутренней структуры* (наличия трех основных типов компонент – шумовых, трендовых и сигнальных) каждого конкретного сигнала. Сами функции формируются, а правильнее сказать, **извлекаются** непосредственно из исходного сигнала, следовательно, такой базис всегда **уникален, апостериорен и высоко адаптивен**. Термин “базис” здесь использован не совсем корректно, поскольку, согласно определению, базис – линейное независимое множество функций, линейная оболочка которого образует все линейное пространство, в котором этот базис используется. Здесь, однако, линейная независимость не является строго доказанной, поэтому такой базис является эмпирическим, аппроксимативным и составляет своего рода “строительные блоки” для представления сигналов. Далее под “базисом” применительно к ДЭМ будет как раз пониматься именно

такой аппроксимативный базис. Технология обработки сигналов на основе ДЭМ в настоящее время получила широкое распространение при решении многочисленных задач предварительной обработки сигналов, основные из которых следующие: очистка сигналов от шума, выделение тренда, анализ степени хаотичности на основе оценивания показателей Херста, Гельдера и фрактальной размерности, проведение мультиразрешающего анализа (исследование тонкой структуры сигналов и их глобального поведения в целом; анализ свойств отдельных компонент) и мультиполосного анализа (исследование сигналов в различных частотных полосах), выделение шаблонов в сигналах и их классификация.

Процесс формирования базисных функций на основе специальным образом разработанного алгоритма не исчерпывает всю технологию, а является лишь первым ее начальным этапом. Далее по найденным базисным функциям формируется частотно-временное представление сигналов – т.н. *спектр Гильберта-Гуанга* [1-3]. Такое представление на частотно-временной плоскости получается путем вычисления мгновенных частот каждой базисной функции, их специальной нормировке и нанесения рассчитанных значений на итоговую трехмерную диаграмму. Такое частотно-временное распределение позволяет еще и получить информацию об амплитуде, которая отображается соответствующим цветом (поэтому данное распределение, хотя и отображается как двумерное на плоскости, на самом деле является трехмерным: “время-амплитуда-частота” или “время-энергия-частота”). Спектр Гильберта-Гуанга, с одной стороны, аналогичен широко известным в частотно-временном анализе распределениям, но он конструируется на основе адаптивно сформированного базиса, что позволяет рассчитывать на более точные результаты при дальнейшем анализе. Приведем основные задачи, решаемые путем исследования построенного спектра Гильберта-Гуанга: выявление скрытых в шумах модуляций сигнала, идентификация частотных и временных диапазонов концентрации энергии, классификация сигналов по спектру Гильберта-Гуанга, расчет на основе спектра Гильберта-Гуанга *маргинального спектра*, являющегося аналогом частотного спектра Фурье для нестационарных сигналов, расчет на основе спектра Гильберта-Гуанга *мгновенной плотности энергии*, позволяющей наглядно выявлять изменения энергетических соотношений в сигналах.

После проведения предварительной обработки и анализа самой структуры сигнала следует третий, завершающий, этап интеллектуального анализа, который осуществляется на основе специальной технологии Data Mining [3]. Интеллектуальный анализ ориентирован на добычу из “сырых” исходных данных (прошедших, однако, этап предварительной обработки) новых, нетривиальных, практически полезных и поддающихся интерпретации человеком-экспертом знаний. Кроме того, задачи интеллектуального анализа должны быть автоматизированы, чтобы, по возможности, исключить эксперта из непосредственного процесса извлечения знаний, а привлечь его лишь на стадии интерпретации результатов. Data Mining использует представление сигналов в виде некоторых абстрактных структур – *моделей* – которые, вместе с тем, должны обеспечивать необходимую точность и достоверность результатов и отражать основные закономерности, свойственные некоторому явлению или процессу. Задач интеллектуального анализа довольно много, при этом основные из них, решаемые с помощью технологии ДЭМ, следующие: выявление участков в сигналах, обладающих однородными свойствами (сегментов), упрощение описания исходного сигнала, заданного в виде набора пар “временной отсчет – сигнальный отсчет” и устранение избыточности путем объединения отсчетов в однородные сегменты, поиск и объединение в группы (кластеры) ранее выделенных сегментов, построение временных шаблонов с учетом их упорядоченности во времени.

Понятие эмпирической и характеристической моды. Алгоритм ДЭМ

Переходя к описанию самой технологии ДЭМ и ее практических приложений, первое, что необходимо сделать – дать строгое определение понятиям “эмпирическая мода” и “характеристическая мода”. Эмпирическая мода (английское на-

звание IMF – Intrinsic Mode Function) [1-3] – функция, заданная непрерывно на интервале существования сигнала или дискретно в виде вектора отсчетов, имеющая в общем случае произвольную форму и аналитическую запись (если таковая существует), которая, однако, должна непременно удовлетворять двум необходимым условиям:

1) Общее суммарное число максимумов и минимумов такой функции (т.е. общее число экстремумов) должно быть строго равно числу нулей функции (в дискретном варианте задания нули могут быть найдены с использованием различных алгоритмов интерполяции) либо отличаться от числа нулей не больше, чем на единицу:

$$N_{\max} + N_{\min} = N_{\text{zero}} \pm 1 \text{ или } N_{\max} + N_{\min} = N_{\text{zero}}, \quad (1)$$

где N_{\max} , N_{\min} , N_{zero} – соответственно, число максимумов, минимумов и нулей функции, не считая краевые отсчеты сигнала, которые в некоторых случаях могут оказаться единственными экстремумами сигнала (случай монотонно возрастающей или убывающей функции на всей ее области определения).

2) Локальное (мгновенное) среднее значение функции, определенное как полусумма двух огибающих, верхней, интерполирующей найденные локальные максимумы и нижней, интерполирующей найденные локальные минимумы, – должно быть меньше или равно заранее определенного порогового значения η , зависящего от машинной точности ε и погрешностей, связанных с получением, преобразованием и передачей сигнальной информации. В качестве средства интерполяции чаще всего используются *кубические сплайны*, причина предпочтения которых приведена ниже. Добиться точного равенства нулю локального среднего значения в каждый момент времени невозможно по ряду объективных причин. К их числу относятся вычислительные погрешности (связанные с особенностями машинной арифметики с плавающей запятой), плохая обусловленность систем уравнений, на основе которых осуществляется расчет коэффициентов сплайнов, а также особенности самого сигнала (например, краевые эффекты – сильные осцилляции интерполирующей функции на краях, из-за которых вблизи краев локальное среднее может отличаться от нуля). Аналитически данное условие записывается в виде:

$$0.5 \cdot [U(k) + L(k)] \leq \eta, \quad k = \overline{1, N}, \quad (2)$$

где $U(k)$ и $L(k)$ – значения верхней и нижней огибающих сигнала в k -й момент времени (k – номер отсчета сигнала), N – общее количество сигнальных отсчетов, η – некоторый порог, устанавливаемый обработчиком-экспертом и обычно принимающий близкие к нулю значения.

Равенство числа экстремумов и числа нулей с точностью до 1 необходимо для того, чтобы ЭМ была узкополосной функцией (это дает преимущества при частотной локализации), т.к. мера узкополосности связана с числом экстремумов и нулей [2].

Условие, касающееся равенства нулю полусуммы двух интерполированных огибающих, имеет свою физическую интерпретацию. Из него следует, что ЭМ является стационарной функцией относительно своего локального среднего значения, которое неизменно и равно нулю (точнее, меньше некоторого порога) в любой момент времени. Кроме того, для достижения этого условия ЭМ должна иметь положительные значения в точках максимумов и отрицательные значения в точках минимумов, т.к. иначе для отдельных моментов времени условие может не выполняться. Наконец, важно отметить, что ЭМ в общем случае обладает одновременно и амплитудной и частотной модуляциями. Закон амплитудной модуляции может быть установлен из огибающих, полученных интерполяцией экстремумов или на основе преобразования Гильберта, а закон частотной модуляции устанавливается на основании закона изменения мгновенной частоты. В качестве информативного параметра на спектре Гильберта выступает цвет (определяемый значением наносимой абсолютной величины), т.к. именно по его изменениям можно определить соответствующие зависимости (изменение амплитудных соотношений, частотные модуляции и пр.).

Понятие огибающей, введенное для определения ЭМ, всегда рассматривается с учетом выбранного вида интерполяции. Как было сказано, используется в основном кубическая сплайн-интерполяция. У сплайнов вообще и кубических в частности есть ряд преимуществ по сравнению с другими функциями, обладающими, как и они, высокой степенью гладкости. Во-первых, по сравнению с классом полиномиальных функций, у сплайнов нет эффекта “раскачивания” – сильных осцилляций интерполирующей функции между узлами интерполяции в случае, если фрагмент имеет заведомо не полиномиальную природу. Но самым важным является то, что кубический сплайн минимизирует колебательное поведение функции – т.е. из всех дважды дифференцируемых, непрерывных на некотором интервале функций, интерполирующих заданную совокупность точек кубический сплайн меньше всего осциллирует [6]. Помимо интерполяционных сплайнов возможно также применение сглаживающего сплайна, т.е. функции, которая не будет гарантированно проходить через все узлы интерполяции. Эта методика весьма эффективна в том случае, если данные заведомо содержат погрешности и ошибки, а операция сглаживания позволяет уменьшить их влияние.

ЭМ должна обладать некоторой симметрией относительно оси времени, которая подразумевает наличие чередующихся локальных максимумов и минимумов, а интенсивность чередования определяется конкретной функцией. Между локальным максимумом и минимумом, как правило, располагается хотя бы один нуль функции, который в дискретном случае рассчитывается с помощью алгоритмов интерполяции. Примерами ЭМ являются гармонический сигнал, ЛЧМ-сигнал, Гауссов радиопульс, периодическая последовательность прямоугольных импульсов. Описание самого алгоритма ДЭМ, позволяющего получить набор компонент из исходного сигнала, приведено в [1-3].

Классификация ЭМ на основе физической интерпретации ДЭМ

При решении любой задачи из числа указанных ранее, необходимо предварительное проведение классификации всех извлеченных из сигнала ЭМ с целью выяснения их физического смысла – задача, ранее не рассматривавшаяся отдельно в рамках данного метода либо решаемая лишь для узкого класса сигналов без необходимого математического обоснования. Далее сказанное будет проиллюстрировано соответствующими примерами с участием *мультигармонических сигналов* (представляющих собой сумму нескольких элементарных гармоник), широко используемых в современной телеметрии. Для абсолютно произвольного сигнала все ЭМ можно разбить на две большие категории:

- 1) *Основные ЭМ* (включая *шумовые ЭМ* и *ЭМ-шаблоны*);
- 2) *Трендовые ЭМ* (включая *компенсирующие ЭМ* и *истинные трендовые ЭМ*).

Основные ЭМ в разложении всегда имеют четкий физический смысл и отражают внутреннюю структуру и особенности, свойственные данному конкретному сигналу. К их числу относятся шумовые ЭМ и ЭМ-шаблоны. Появление в разложении первых объясняется наличием в исходном сигнале шума, а вторые связаны непосредственно с самим полезным сигналом и входящими в него компонентами. В качестве примера, поясняющего сказанное, можно привести следующий: при разложении мультигармонического сигнала с аддитивным шумом получается несколько шумовых ЭМ (они располагаются на начальных уровнях и их число определяется интенсивностью шума, его частотными свойствами и общим числом отсчетов) и ЭМ-шаблоны (в разложении следуют за шумовыми ЭМ), к числу которых относятся соответствующие элементарные гармоники, образующие незашумленный мультигармонический сигнал. Понятие “ЭМ-шаблон” подразумевает некоторую функцию, характерную для данного типа сигналов, например, любую из элементарных гармоник для мультигармонического сигнала или, например, ЛЧМ-сигнал и элементарную гармонику в случае наличия их аддитивной смеси с шумом.

Трендовые ЭМ являются медленно меняющимися функциями (обычно описываются полиномами невысокой степени или экспоненциальной функцией), характеризующими изменение во времени локального среднего значения сигнала. Среди них, согласно вышеприведенной классификации, выделяют *истинные трендовые ЭМ* (всегда имеющие физический смысл), описывающие истинную динамику среднего значения и т.н. *компенсирующие ЭМ*, которые могут возникать в разложении в том случае, если тренд в сигнале отсутствует. Истинные трендовые ЭМ появляются, например, при разложении суммы гармонического сигнала и полиномиального тренда. Компенсирующие (ложные) ЭМ – результат несовершенства самого алгоритма ДЭМ, критериев остановки процесса отсеивания, неточностей при вычислениях (ошибок округления и пр.). Их появление не связано с какими-либо физическими или математическими особенностями рассматриваемых сигналов, а объясняется только лишь несовершенством вычислительной процедуры. Компенсирующие ЭМ обычно создают избыточность в разложении, а их название объясняется тем, что в сумме они дают функцию, очень близкую к нулю, т.е., по сути, компенсируют друг друга. Одной из наиболее типичных причин появления компенсирующих ЭМ являются побочные эффекты интерполяции огибающих – такие как появление “всплесков” (*overshoots*), сильные осцилляции вблизи концов сигнала (краевые эффекты). Для частичного преодоления вышеперечисленных проблем существуют специальные методы [2]: подавление краевых эффектов с помощью зеркального отображения экстремумов, наиболее близко расположенных к границам сигнала, тестирование различных критериев остановки процесса отсеивания, уточнение местоположения экстремумов для построения огибающих и пр.

На рис. 1 показан пример декомпозиции мультигармонического сигнала, состоящего из двух элементарных гармоник (сам исходный сигнал показан первым) в смеси с аддитивным гауссовским шумом.

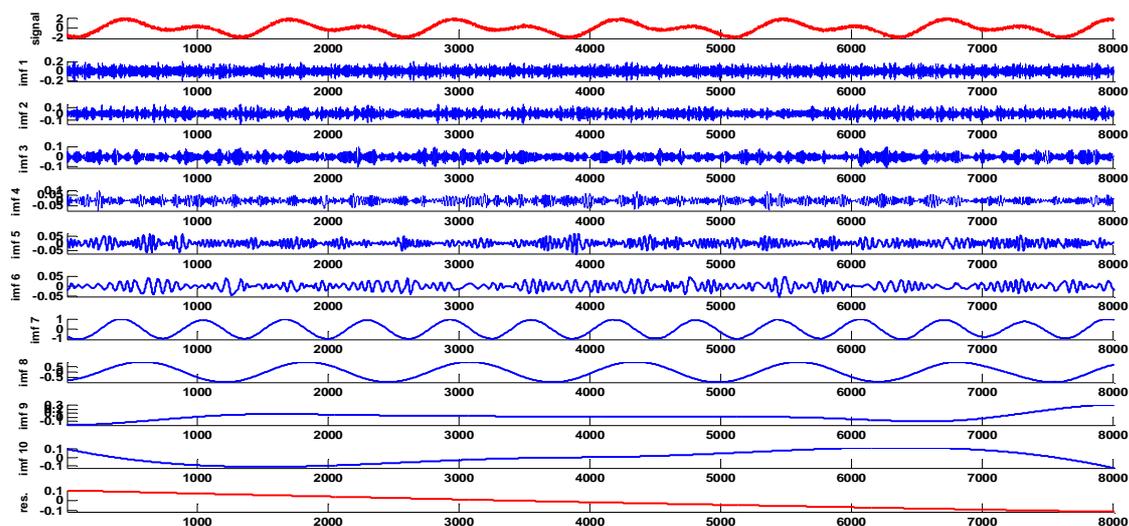


Рис. 1. Пример декомпозиции мультигармонического сигнала

Амплитуды гармоник одинаковы и равны 1, а частоты отличаются в 2 раза – такое соотношение взято для иллюстрации структуры диадического банка фильтров, которая присуща ДЭМ. Из картины разложения видно, что гармоники извлекаются друг за другом в порядке уменьшения их частоты (ЭМ № 7 и 8 соответственно). В общей сложности извлечено 10 ЭМ и результирующий остаток. Первые *шесть ЭМ* являются *основными*, или, более точно, *шумовыми*, их появление обусловлено наличием шума. Далее также идут *основные ЭМ*, относящиеся к категории *ЭМ-шаблонов*, – две элементарные гармоники. Наконец, две последние ЭМ и результирующий оста-

ток – компенсирующие ЭМ (в исходном сигнале отсутствует тренд). Такое название обусловлено еще и тем, что в сумме они дают функцию, очень близкую к нулю во всех точках, т.е., по сути, компенсируют друг друга. Компенсирующие ЭМ создают избыточность разложения.

Новый подход к выявлению шаблонов в сигналах и идентификации компонент на основе теории регрессии

Рассмотрим новый подход к выявлению всех рассмотренных типов ЭМ с целью идентификации шумовой составляющей, характерных сигнальных составляющих и тренда. Для этого модель исходного сигнала представим в следующем виде:

$$s = x\beta + e, \quad (3)$$

где s – исходный сигнал, x – матрица регрессоров, β – вектор неизвестных коэффициентов, определяющих удельный вес регрессоров при описании исходного сигнала и подлежащих оцениванию, e – белый гауссовский шум с параметрами $N(0; \sigma^2)$.

Теперь, используя результат разложения с помощью ДЭМ, запишем связь между всеми извлеченными ЭМ и исходным сигналом на основании свойства полноты разложения:

$$s(k) = \sum_{i=1}^{M-1} c_i(k) + r_M(k), \quad (4)$$

где $r_M(k)$ – результирующий остаток в разложении. Из первого слагаемого правой части можно выделить отдельно самую первую ЭМ, которая будет являться аппроксимацией шума, присутствующего в сигнале поскольку в силу наиболее сильно выраженного по сравнению со всеми остальными ЭМ высокочастотного характера первую ЭМ можно трактовать как приближение исходного шума:

$$s(k) = c_1(k) + \sum_{i=2}^{M-1} c_i(k) + r_M(k) = e(k) + \sum_{i=2}^{M-1} c_i(k) + r_M(k). \quad (5)$$

Теперь сделаем еще одно уточнение. Поскольку точная сходимость суммы всех ЭМ и результирующего остатка к исходному сигналу математически строго не доказана (эта сходимость рассматривается в инженерном смысле, т.е. абсолютная разность между исходным сигналом и суммой всех извлеченных ЭМ есть малое число, обычно порядка $10^{-16} \dots 10^{-10}$), введем некоторые весовые коэффициенты для каждой ЭМ (по умолчанию все веса единичные), которые будем определять в соответствии с МНК и которые должны улучшить точность восстановления. Тем самым должна улучшиться точность представления исходного сигнала полученными из него же компонентами:

$$s(k) = c_1(k) + \sum_{i=2}^{M-1} \beta_i c_i(k) + \beta_M c_M(k) = e(k) + \sum_{i=2}^{M-1} \beta_i c_i(k) + \beta_M c_M(k), \quad (6)$$

где β – вектор весовых коэффициентов. Выражение (6) можно переписать в более общем векторно-матричном виде:

$$s = C\beta + e, \quad (7)$$

где C – матрица, в столбцах которой хранятся отсчеты ЭМ (число столбцов на единицу меньше общего числа функций в разложении, т.к. первая ЭМ выступает как аппроксимация исходного шума и не включается в эту матрицу). МНК-оценка вектора коэффициентов β определяется формулой:

$$\hat{\beta} = (C^T C)^{-1} C^T s. \quad (8)$$

Введение данных весовых коэффициентов и их дальнейшее вычисление (например, по МНК) может использоваться не только для повышения точности при восстановлении сигналов, но и для придания ДЭМ смысла “базиса”. Ведь, как было сказано, базис из ЭМ является апостериорным, следовательно, для каждого сигнала он индивидуален и, будучи полученным однажды, он не может использоваться для представ-

ления другого сигнала. Однако при работе с однотипными сигналами (например, гармониками с разными амплитудами, но одинаковыми частотами) данные весовые коэффициенты могут применяться для получения нового базиса из старого путем умножения ранее извлеченных компонент на найденные веса. Для коэффициентов можно сформировать интервальную оценку, т.е. найти границы интервала, в который конкретный коэффициент регрессионной модели попадает с заданной вероятностью P . Можно показать, что границы доверительного интервала (ДИ) определяются следующим образом:

$$\beta_i \in \left[\hat{\beta}_i - t_\gamma \hat{\sigma} \left\{ \sqrt{(CC^T)^{-1}} \right\}_{ii}; \hat{\beta}_i + t_\gamma \hat{\sigma} \left\{ \sqrt{(CC^T)^{-1}} \right\}_{ii} \right], \quad (9)$$

где $\beta_i, \hat{\beta}_i$ - истинное значение коэффициента и его оценка соответственно.

В классической задаче регрессии после нахождения соответствующих коэффициентов модели иногда осуществляется их проверка на значимость, по результатам которой конкретный элемент модели с проверяемым коэффициентом либо сохраняется, либо исчезает. Аналогично эту задачу можно решить для всех найденных весовых коэффициентов. Для этого рассмотрим две альтернативные гипотезы, одна из которых предполагает незначимость данного коэффициента модели (нулевая гипотеза H_0), а вторая, соответственно, его значимость (ненулевая гипотеза H_1), т.е. его отличие от нуля. В данной ситуации нулевая гипотеза отвергается, поскольку все коэффициенты являются априорно значимыми, отличными от нуля, т.к. все ЭМ с теми или иными весами участвуют при восстановлении исходного сигнала. Однако нижеприведенную статистику, изначально предназначенную именно для проверки значимости коэффициентов модели, можно рассматривать как *классификационную* (т.е. как некоторую функцию, связанную с исходными данными), а рассчитываемые с ее помощью значения распределять по группам (каждая группа будет соответствовать одному из типов ЭМ) на основе некоторой *дискриминирующей процедуры*, например, кластер-анализа. Сама статистика имеет вид:

$$T_{li} = |\hat{\beta}_i| / \left\{ \sqrt{\{V^{-1}\}_{ii}} \cdot \hat{\sigma}_e \right\}, \quad (10)$$

где V – матрица ковариаций ЭМ, определяемая как $V = CC^T$ (при вычислении статистики используются диагональные элементы матрицы, обратной к ковариационной), σ_e – средноквадратическое отклонение (СКО) шума, которое можно определить по первой ЭМ с использованием, например, робастной медианной оценки [1,4], устойчивой к наличию погрешностей и аномальных ошибок в данных. Ниже (табл. 1) проиллюстрирован пример расчетов для мультигармонического сигнала, показанного на рис. 2 (частоты гармоник образуют геометрическую прогрессию со знаменателем 2; При разложении сигнала было получено 12 компонент). Медианная оценка СКО шума по первой ЭМ составляет 0.2876.

Исходя из полученных значений классификационной статистики можно сделать вывод, что 6,7,8 ЭМ (соответствующие трем элементарным гармоникам, выделенным в порядке уменьшения их частоты) являются *значимыми*, т.е. относятся к *ЭМ-шаблонам*, в то время как остальные являются незначимыми и относятся либо к шумовым, либо к компенсирующим (это видно по сильно отличающимся значениям статистики).

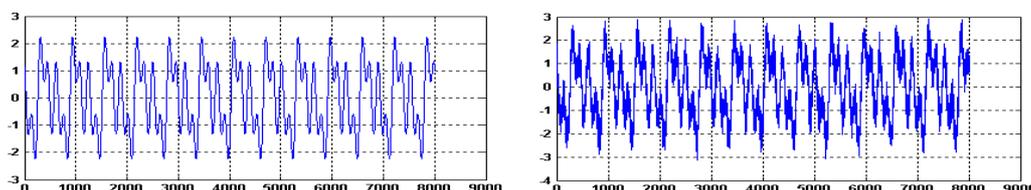


Рис. 2. Мультигармонический сигнал, состоящий из трех элементарных гармоник (слева) и его копия с аддитивным гауссовским шумом $N(0; 0.3^2)$ (справа)

Таблица 1

Значения статистики T_1 для всех выделенных из сигнала ЭМ

Номер ЭМ	Коэффициенты β , вычисленные по МНК	$\{v^{-1}\}_{ii}$	Нижняя граница ДИ	Верхняя граница ДИ	Длина ДИ	$T_1 \cdot 10^{-4}$
2	0.9557	0.0067	0.9448	0.9667	0.0219	0.1658
3	0.9075	0.0126	0.8869	0.9282	0.0413	0.0792
4	0.8944	0.0205	0.8606	0.9282	0.0676	0.0471
5	0.9867	0.0046	0.9792	0.9943	0.0151	0.2564
6	0.9982	0.0002	0.9978	0.9986	0.0008	4.9331
7	1.0022	0.0003	1.0017	1.0026	0.0009	4.6289
8	0.9998	0.0003	0.9994	1.0003	0.0009	4.4683
9	1.0003	0.0206	0.9665	1.0341	0.0676	0.0588
10	1.0290	0.0332	0.9744	1.0836	0.1092	0.0385
11	1.0487	0.1121	0.8643	1.2331	0.3688	0.0119
12	1.0540	0.0618	0.9523	1.1556	0.2033	0.0217

Для **автоматического** проведения данной классификации можно воспользоваться одним из алгоритмов кластер-анализа, например, алгоритмом К-средних (k-means) с числом кластеров, равным 2. Число 2 выбрано для того, чтобы в итоге получить две независимых группы компонент: ЭМ-шаблоны, выделенные в табл. 1 жирным шрифтом, и остальные компоненты (трендовые и шумовые). Также можно использовать специальные алгоритмы (например, основанные на критериях качества разбиения), позволяющие определить число кластеров вместо того, чтобы задавать его самим. Такая процедура может быть необходима для более тщательного анализа компонент, например, если требуется более тонкое разделение внутри группы ЭМ-шаблонов. Результаты кластеризации на основе значений статистики соответствуют высказанным предположениям: компоненты с 6-й по 8-ю относятся к одному кластеру, а все остальные – к другому. В некоторых случаях, для более детальной классификации, можно задать число кластеров, равное 3 с целью выделить три типа ЭМ: шумовые, ЭМ-шаблоны и компенсирующие ЭМ.

Ниже, на рис. 3, показан результат восстановления незашумленного мультигармонического сигнала, полученного в результате суммирования 6-й, 7-й и 8-й компонент разложения. На основании данного графика и его сравнения с графиком на рис. 2 можно сделать вывод о высокой точности восстановления полезного сигнала (прослеживается лишь незначительное влияние краевых эффектов).

Помимо значений статистики T_1 , можно назвать еще один способ классификации ЭМ, а именно длина ДИ для соответствующих регрессионных коэффициентов, которая, как видно из табл. 1, наименьшая именно для ЭМ-шаблонов, что выделяет последние из общего набора компонент. Таким образом, длина ДИ может служить еще одним описательным признаком, используемым при кластер-анализе.

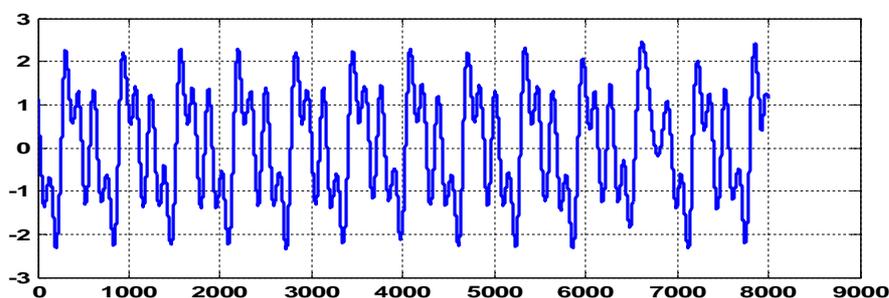


Рис. 3. Восстановленный мультигармонический сигнал без шума

В случае необходимости различения шумовых и трендовых компонент (если значения классификационной статистики оказываются для них близкими и кластер-анализ не позволяет их различить), возможно использование показателя Херста H [1,5], характеризующего степень регулярности функции и ее склонность к хаотическому поведению. Данная величина заключена в пределах $[0;1]$ и имеет два характерных диапазона значений. При $0 \leq H < 0.5$ компоненты являются *антиперсистентными* (с *неустойчивым трендом, с кратковременной памятью*), т.е. наблюдается т.н. тенденция возврата к среднему. Значения из данного диапазона имеют все шумовые ЭМ. При $0.5 < H \leq 1$ компоненты являются *персистентными* (с *устойчивым трендом, с долговременной памятью*), т.е. если имелась тенденцию к возрастанию (убыванию), то она будет сохранена и в дальнейшем на некотором промежутке. Чем ближе значение H к 1, тем сильнее выражена данная тенденция и тем дольше она сохраняется. Трендовые компоненты имеют показатель Херста, заключенный в этом диапазоне. Наконец, еще одним средством различения компонент является энтропия Шеннона, которая для шумовых компонент, как для более хаотичных и близких к случайным, дает значения, большие, чем для трендовых.

Особенности современной телеметрии

Существует несколько основных способов классификации телеметрических (ТМ) сигналов [7], однако наиболее часто используемой является их классификация по скорости изменения во времени, в соответствие с которой они делятся на *медленно* и *быстро меняющиеся сигналы* (ММС и БМС). Основное количественное отличие между ними состоит в разных *эффективной ширине спектра*. При этом качественного отличия может не наблюдаться, т.к. они могут описывать работу одного и того же *сложного динамического объекта* (СДО) [7].

ММС являются в телеметрии наиболее многочисленными и характеризуются спектром частот, группирующимся в диапазоне от 0 до нескольких десятков Гц, т.е. спектр сосредоточен в области низких частот. Во временной области такие сигналы, как правило, представляют собой набор фрагментов с определенным монотонным поведением, которое обычно достаточно точно описывается полиномом невысокой степени (3-5) без учета влияния присутствующего в сигнале шума. Такие сигналы являются *персистентными* (с *устойчивым трендом*), т.е. если сигнал имел тенденцию к возрастанию (или убыванию), то наверняка он ее сохранит и в дальнейшем. *Показатель Херста* таких сигналов лежит в пределах $0.5 < H \leq 1$. Чем ближе значение H к 1, тем сильнее выражена общая монотонная тенденция и тем дольше она сохраняется. К данной категории сигналов относятся температурные параметры, давления, скорости потоков жидкостей и газов, линейные и угловые перемещения частей некоторого объекта со временем, скорости и ускорения объектов.

БМС имеют частотный спектр, более широкий, чем ММС. Верхние граничные частоты могут достигать значений нескольких десятков и сотен кГц. Во временной области такие сигналы являются сильно осциллирующими и резко меняющимися, часто имея шумоподобный вид. С точки зрения анализа регулярности, данные процессы являются *антиперсистентными* (с *неустойчивым трендом*), т.е. наблюдается т.н. *повсеместная тенденция возврата к среднему*. Если, например, сигнал является монотонно возрастающим в течение некоторого промежутка времени, то затем неизбежно наступит монотонное убывание сигнала – это как раз будет гарантировать примерное равенство среднего значения в любом сечении сигнала. Значения показателя Херста для них лежат в пределах $0 \leq H < 0.5$. БМС составляют значительно меньшую группу, чем ММС. Они отражают сложные осциллирующие процессы, протекающие в различных системах. В качестве примеров можно привести вибрации конструкций корпуса летательных аппаратов, пульсации давления в камере сгорания двигателей, акустические и аэродинамические нагрузки и пр.

Регрессионный подход, направленный на автоматическое выделение шаблонов в сигналах, а также на идентификацию трех ранее представленных типов компонент, может применяться и к БМС, и к ММС для решения этих задач. Однако в каждом конкретном случае шаблоны являются различными по своей природе и соответственно получаемые результаты должны быть отдельно проанализированы.

Заключение

В данной статье представлен новый подход к анализу нестационарных сигналов сложной внутренней структуры – декомпозиция на эмпирические моды. Описаны идеи, заложенные в его основу, а также основные преимущества по сравнению с существующими методами анализа нестационарных сигналов. Рассмотрена задача классификации компонент, получаемых при применении и ДЭМ к сигналу, которая, во-первых, позволяет понять внутреннюю структуру самих сигналов, а во-вторых, идентифицировать различные типы компонент. Идентификация проводится на основе теории регрессии и дискриминантного анализа, что позволяет добиться ее автоматизации. Кроме того, значения специальной классификационной статистики, показателя Херста и энтропии Шеннона могут использоваться в качестве вспомогательных признаков для различения компонент.

Поддержка

Настоящая работа поддержана грантом 2009 года для студентов, аспирантов вузов и академических институтов, расположенных на территории Санкт-Петербурга, выделенным Комитетом по науке и высшей школе Санкт-Петербурга.

Литература

1. Клионский, Д.М. Декомпозиция на эмпирические моды и ее использование при анализе дробного броуновского движения [Текст] / Д.М. Клионский, Н.И. Орешко, В.В. Геппенер // Цифровая обработка сигналов. – 2008. – № 3. – с. 37-45.
2. Huang, N.E. Hilbert-Huang Transform and Its Applications [Text] / N.E. Huang, S.S. Shen // World Scientific. – 2005.
3. Klionsky, D.M Empirical mode decomposition in segmentation of slowly and fast changing non-stationary signals [Text] / D.M. Klionsky, V.V. Geppener // Proceedings of PRIA Conference, Yoshkar-Ola, Russia. – 2007.
4. Малла, С. Вейвлеты в обработке сигналов / С. Малла. // изд. М.: МИР. – 2005. – 671 с.
5. Петерс, Э. Э. Фрактальный анализ финансовых рынков / Э.Э. Петерс. // изд. М.: Интернет-Трейдинг, 2004. – 304 с.
6. Мэтьюз, Д. Г. Численные методы. Использование Matlab / Д.Г. Мэтьюз, К.Д. Финк // изд. дом “Вильямс”, Москва-Санкт-Петербург-Киев. – 2001.
7. Назаров, А.В. Современная телеметрия в теории и на практике / А.В. Назаров, Г. И. Козырев, И. В. Шитов // СПб.: Наука и Техника. – 2007. – 672 с.



NEW APPROACH TO AUTOMATIC PATTERN EXTRACTION IN TELEMETRIC SIGNALS ON THE BASIS OF THE EMPIRICAL MODE DECOMPOSITION

D. M. KLIONSKIY¹

N. I. ORESHKO²

V. V. GEPPENER³

¹⁾ *Saint-Petersburg
Electrotechnical
University "LETI"
by V.I. Ulyanov-Lenin*

²⁾ *Research and
Engineering Center
of Saint-Petersburg
Electrotechnical
University*

e-mail: klio2003@list.ru;

oreshko@nic.spb.ru;

geppener@mail.ru

The present paper deals with the new method of non-stationary signal processing – empirical mode decomposition (EMD), which is principally intended for signals with complex structure. Due to a number of important properties – locality and adaptivity – this method allows us to solve many practical tasks with high accuracy and reliability. The area of EMD applications can be divided into three main groups: preprocessing tasks, signal structure analysis and intellectual investigation of signals. A new approach to signal structure analysis with the help of the regression theory has been suggested and considered in the paper. This approach is capable of extracting noise-like components, the most typical signal components (signal patterns) and trend-like components (genuine and spurious ones). Also a new way of classifying all extracted IMFs on the basis of their physical interpretation is described. The procedure can be done automatically by using discrimination analysis (cluster-analysis). An example with a multiharmonic signal embedded in additive Gaussian noise is introduced and the numerical results are provided and explained. Finally, we outline some most widely used types of telemetric signals along with the way of applying the regression approach to them.

Key words: adaptivity, intrinsic mode function, empirical mode decomposition, IMF-pattern, noise-like IMF, trend-like IMF, compensating IMF, classifying statistic, regression approach, discriminating procedure, multiharmonic signal.

СИСТЕМНЫЙ АНАЛИЗ И УПРАВЛЕНИЕ

УДК 621.316.7

НЕКОРРЕКТНЫЕ ЗАДАЧИ И НЕПАРАМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ СИСТЕМ УПРАВЛЕНИЯ

В. И. КАПАЛИН

*Московский
государственный
институт электроники
и математики*

e-mail: v_kapalin@mail.ru

Проблема идентификации связана с проблематикой теории некорректных задач. В статье рассматриваются некоторые основные в этой области проблемы и обсуждаются пути преодоления некоторых сложностей при применении непараметрических методов идентификации.

Ключевые слова: идентификация, априорная информация, «черный ящик», нейросетевой подход.

Теория управления нередко имеет дело с объектами, математическое описание которых известно неточно, либо отсутствует полностью. Для успешного управления такими объектами их необходимо идентифицировать. Что же понимается под задачей идентификации? Термин «идентификация» стал использоваться в отечественной науке начиная с 60-х годов прошлого века. Этот термин употребляется для задач, в которых математическую модель объекта управления требуется построить по информации о реакциях объекта на известные внешние воздействия. Термин идентификация – это транслитерация английского слова «identification», которое однако в английском языке не является сугубо научным термином. Так, англо-русский словарь В.К. Мюллера дает перевод слова «identification» просто как «опознание», «установление личности». Задачи идентификации, т.е. опознавания моделей «живых организмов и машин», рассматривались как одна из задач кибернетики Норбертом Винером. Именно этой задаче в ее наиболее общей непараметрической и нелинейной постановке была посвящена последняя публикация Винера «Нелинейные задачи теории случайных процессов», вышедшая на русском языке в 1961 году [1]. Эта задача обсуждалась и в ряде работ других авторов [2]. Частично тема винеровской теории идентификации затронута и в настоящем докладе.

В постановке задачи идентификации существует известная свобода. Традиционно методы идентификации разделяют на две большие группы: непараметрические и параметрические [3]. Непараметрические методы ориентированы на случай, когда априорная информация о структуре модели объекта отсутствует или игнорируется, т.е. когда объект рассматривается как «черный ящик» бихевиоризма и кибернетики (рис. 1).

В этом случае отыскиваются некоторые функциональные характеристики модели – импульсная переходная функция, частотные характеристики, ядра Вольтерра и Винера или их изображения.



Рис. 1

Если имеется априорная информация об уравнениях модели объекта, заданных с точностью до неизвестных параметров, то задача идентификации сводится к оценке этих параметров. Это случай параметрической идентификации. Общих рекомендаций, когда следует использовать методы параметрической, а когда – непараметрической идентификации не существует – все определяется конкретной задачей исследования. Однако во всех случаях идентификации приходится считаться с неточностями в задании модели, неточностями в измерениях сигналов шумами и вычислительными погрешностями. Как результат малые погрешности в эмпирических данных и в задании модели могут привести к значительным ошибкам в результатах идентификации. Это типичная ситуация проблематики некорректных задач [4] и настоящий доклад связан с обсуждением этой проблематики для задач непараметрической идентификации. Для линейных стационарных систем, для которых описание типа «вход-выход» задается интегралом свертки. Задача непараметрической идентификации в этом случае сводится к решению интегрального уравнения Вольтерра I рода относительно неизвестной импульсной переходной функции, т.е. ядра оператора Вольтерра.

$$y(t) = \int_0^t h(\tau)x(t-\tau)d\tau, \quad t \in [0, T]. \quad (1)$$

Это – задача о решении интегрального уравнения I-го рода. Она относится к числу некорректных и требует, вообще говоря, применения методов регуляризации. Однако, в отличие от задач математической физики в задачах идентификации применение общего метода решения некорректных задач – метода линеаризации сглаживающего функционала А.Н. Тихонова практически исключается по следующим основным причинам.

В методе сглаживающего функционала для получения гладкого решения вместо решения достаточно простого уравнения Вольтерра нужно решать некоторое интегро-дифференциальное уравнение, что требует применения несравненно более сложных численных методов. Попытки заменить решение этого интегро-дифференциального уравнения решением интегрального уравнения Фредгольма II рода за счет применения слабой регуляризации оказывается безрезультатным. Слабая регуляризация не обеспечивает требуемой гладкости решений и сходимости семейства приближенных решений к точному решению даже в пространстве непрерывных функций [5]. Наконец отыскание в методе сглаживающего функционала параметра регуляризации по невязке тоже вызывает в задачах идентификации значительные вычислительные сложности.

Возможно, однако, и другое решение задач непараметрической идентификации с учетом их некорректности, не использующие метод сглаживающего функционала и гораздо более простой с вычислительной точки зрения. В основе этого решения лежат методы регуляризации М.М. Лаврентьева и А.С. Апарцина [5], в которых не требуется перехода к интегро-дифференциальным уравнениям, а регуляризованное решение находится из интегрального уравнения Вольтерра. Как следствие необ-

ходимые алгоритмы регуляризации оказываются легко реализуемыми в нынешнем негласном университетском стандарте – пакете MATLAB.

Суть этих методов заключается в замене исходной некорректной задачи на, в некотором смысле, близкую к ней, но корректную задачу. С этой целью можно использовать три подхода.

Первый из них заключается в выборе шага при решении уравнения Вольтерра с помощью одной из квадратурных формул. Как было доказано, в этом случае регуляризацию обеспечивает правильный выбор шага дискретизации при условии, что решение ищется с помощью одной из формул прямоугольников – левых, правых или средних. Применение более точных квадратурных формул, таких как формула Симпсона или формула Грегори приводит к расходящимся методам. Решение задачи идентификации при использовании формулы средних прямоугольников может быть найдено по рекуррентной формуле.

$$\tilde{h}\left[\left(i - \frac{1}{2}\right)\Delta\right] = \frac{1}{x\left[\frac{\Delta}{2}\right]} \left(\frac{y[i\Delta]}{\Delta} - \sum_{j=1}^{i-1} \tilde{h}\left[\left(j - \frac{1}{2}\right)\Delta\right] x\left[\left(i - j + \frac{1}{2}\right)\Delta\right] \right). \quad (2)$$

Второй подход связан с регуляризацией по-Лаврентьеву и заключается в замене исходного уравнения первого рода на уравнение второго рода с параметром регуляризации. В этом случае решение может быть найдено с помощью следующей модификации рекуррентной формулы. Здесь α – параметр регуляризации.

$$\tilde{h}\left[\left(i - \frac{1}{2}\right)\Delta\right] = \frac{1}{\alpha + \Delta x\left(\frac{\Delta}{2}\right)} \left\{ y[i\Delta] - \sum_{j=1}^{i-1} \tilde{h}\left[\left(j - \frac{1}{2}\right)\Delta\right] x\left[\left(i - j + \frac{1}{2}\right)\Delta\right] \right\}. \quad (3)$$

Наконец, третий подход к решению уравнения Вольтерра первого рода с учетом его некорректности связан с применением метода наименьших квадратов. Здесь возможны два варианта решения.

При первом – модель объекта записывается в непрерывном времени и для решения используется интегральное уравнение Фредгольма I-го рода, задающее необходимое условие минимума квадратичного функционала, тоже записанного для непрерывного времени. Искомое решение находится путем дискретизации интегрального уравнения, задающего условие минимума квадратичного функционала и решения полученной системы линейных алгебраических уравнений.

Второй, значительно более простой с вычислительной точки зрения подход заключается в использовании с самого начала дискретного аналога линейного оператора Вольтерра и минимизации квадратичного функционала, записанного для дискретного времени. Здесь необходимое условие минимума сразу дает систему линейных алгебраических уравнений относительно дискретных значений ядра оператора Вольтерра и проблема решения уравнений Фредгольма первого рода не возникает. Регуляризацию в случае применения метода наименьших квадратов можно осуществить либо по-Лаврентьеву

$$\left(\Delta \tilde{A}^T \tilde{A} + \alpha I \right) \tilde{H} = \tilde{A}^T Y, \quad (4)$$

либо применяя для решения полученной СЛАУ процедуру сингулярного разложения матриц – SVD

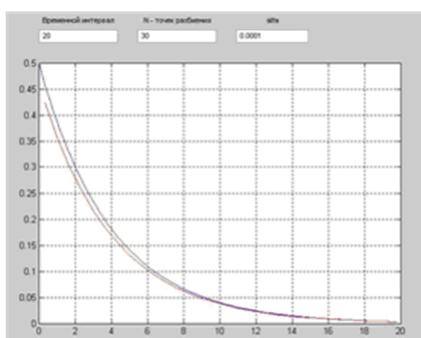
$$A = U \Sigma V^T. \quad (5)$$

Охарактеризовав проблему и возможные пути ее решения, перейдем к практическим результатам исследования, т.е. к результатам вычислительных экспериментов.

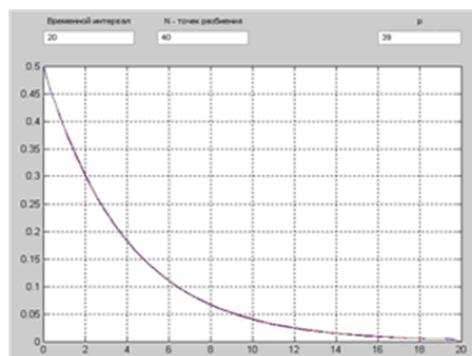
Для проведения экспериментов использовался пакет MATLAB 7.3. Для вывода результатов в графической форме применялся MATLAB Compiler 4. Были рассмотрены случаи точных измерений выходного сигнала и измерений выходного сигнала в присутствии аддитивной помехи. В качестве объектов идентификации были выбраны стандартные передаточные функции морского дизеля, что определялось требовани-

ем заказчика работы – Вьетнамского морского университета, г. Хайфон. В качестве входного сигнала использовались единичная ступенька и полигармонический сигнал. Аддитивная помеха тоже задавалась полигармоническим сигналом или белым шумом. В общей сложности на разработанном автономном программном обеспечении было проведено 72 эксперимента с различными видами передаточных функций морского дизеля и различными возмущениями. Рассмотрим результаты проведенных экспериментов.

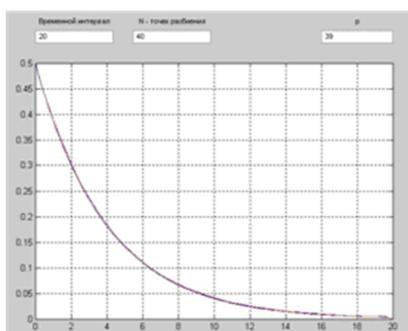
Случай 1. Точные измерения. Целью здесь было проверить работоспособность рассмотренных методов идентификации в присутствии только вычислительных погрешностей. Из полученных результатов на рис. 2-3 следует, что все три группы методов дают практически одинаковые результаты. Поэтому в этом случае предпочтительней самый простой алгоритм – рекуррентный без регуляризации.



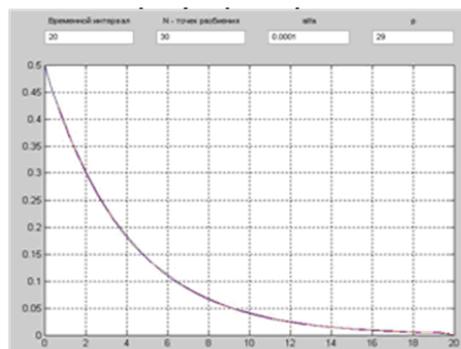
Рекуррентное решение



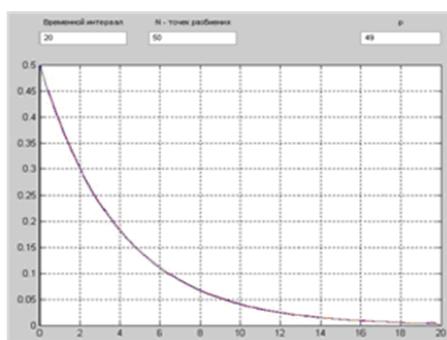
Рекуррентное решение с регуляризацией



МНК

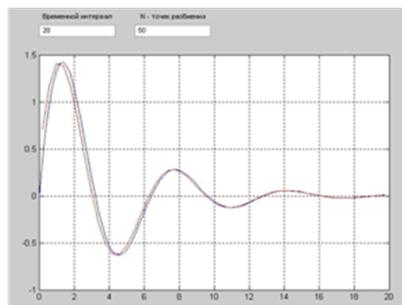


МНК с регуляризацией

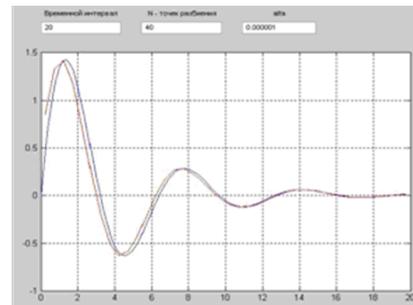


SVD

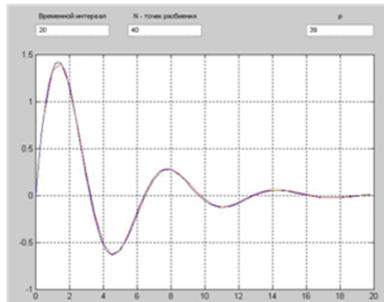
Рис. 2



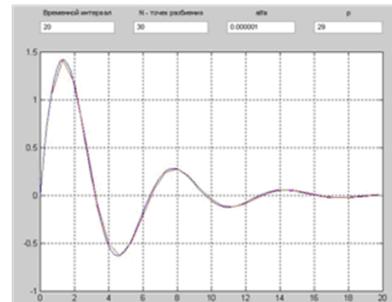
Рекуррентное решение



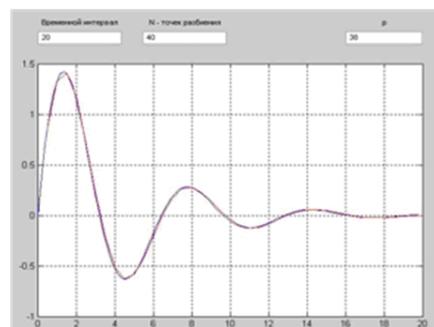
Рекуррентное решение с регуляризацией



МНК



МНК с регуляризацией

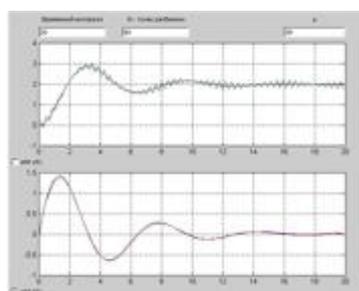


SVD

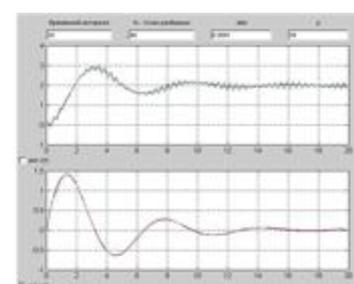
Рис. 3

Случай 2. Гладкие помехи. Следующая группа экспериментов было проведена для случая гладких помех. Здесь применение регуляризации по-Лаврентьеву ощутимо – как в случае рекуррентного метода решения, так и в случае применения метода наименьших квадратов (рис. 4). Полученные этими методами результаты примерно такие же, как дает по точности применение процедуры SVD.

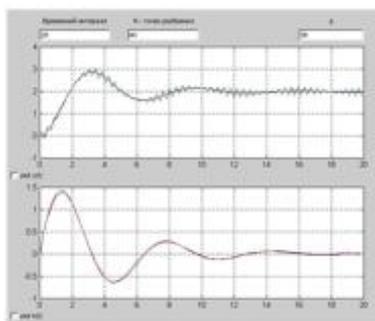
Случай 3. Помехи в виде белого шума. Это самый опасный вид помех. В этом случае рекуррентные алгоритмы и процедура SVD не дали вообще никакого результата. Приемлемые результаты удалось получить только методом наименьших квадратов с применением регуляризации по-Лаврентьеву (рис. 5).



МНК

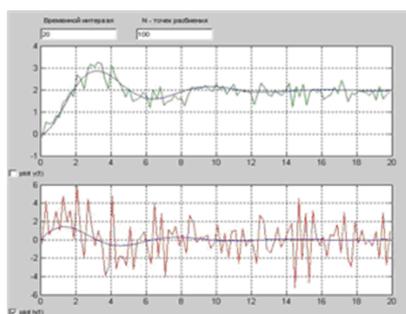


МНК с регуляризацией

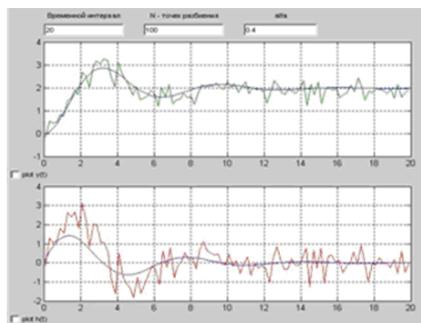


SVD

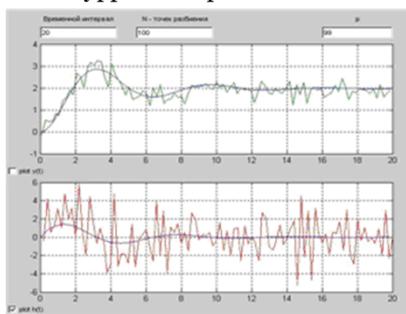
Рис. 4



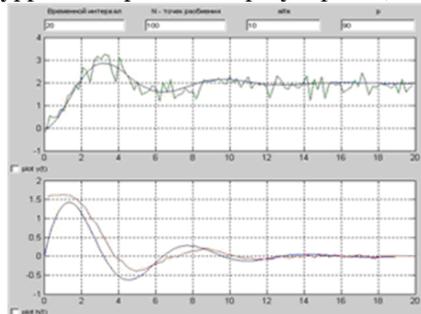
Рекуррентное решение



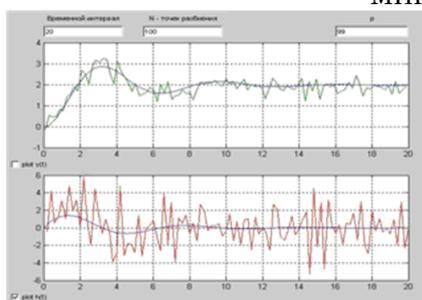
Рекуррентное решение с регуляризацией



МНК



МНК с регуляризацией



SVD

Рис. 5

Рассмотренные здесь методы относятся к классическим методам непараметрической идентификации линейных стационарных систем. Из них на класс нелинейных систем обобщается только метод наименьших квадратов. В теории нелинейных систем Вольтерра-Винера модель «черного ящика» задается полиномом Вольтерра

$$y(t) = \sum_{i=1}^M \int_0^t \dots \int_0^t h_i(\tau_1, \dots, \tau_i) u(t-\tau_1) \dots u(t-\tau_i) d\tau_1 \dots d\tau_i, \quad (6)$$

который с использованием формулы прямоугольников можно заменить дискретной моделью. Минимизация квадратического критерия качества здесь, как и в линейном случае, дает систему линейных алгебраических уравнений. Однако получающаяся размерность задачи оказывается весьма велика, что существенно затрудняет получение практических результатов в общем нелинейном случае.

На практике в качестве нелинейной модели используется обычно не общая модель Вольтерра-Винера, а ее частные случаи – модели Гаммерштейна и Винера-Гаммерштейна, алгоритмы настройки которых включены в расширение System Identification Toolbox пакета MATLAB.

Рассмотренные ранее методы непараметрической идентификации используют аппарат интегральных уравнений. Альтернативным путем решения задачи построения модели «черного ящика» ставшим возможным совсем недавно стал нейросетевой подход. Его практическое использование в учебном процессе стало возможным после включения в MATLAB 6.5 расширения Neural Network Toolbox. Это расширение включает более 150 различных функций для создания, обучения и использования различных нейронных сетей. Для настройки моделей можно использовать специальный блок пакета Simulink. Построение нейросетевого регулятора может также быть осуществимо в библиотеке Neural Network Block Set пакета Simulink.

Если применение интегральных уравнений для решения задач непараметрической идентификации основывается на известных теоремах функционального анализа о линейных и нелинейных функционалах, то применение нейронных сетей основывается на теоремах о полноте. Фактически эта теорема означает, что с помощью нейронных сетей можно моделировать любую нелинейную зависимость при условии правильного выбора архитектуры сети и ее правильного обучения.

Нейросетевой подход использовался здесь для построения нейросетевого регулятора для рассмотренных моделей морского дизеля – как в линейном, так и в нелинейном случае – для модели Гаммерштейна. Практика настройки нелинейных моделей показала, что это гораздо более сложная процедура, чем настройка линейных моделей. Для нейросетевой модели морского дизеля FODEN FD7 на средних скоростях

$$W(p) = \frac{-0.36p + 19.06}{s^2 + 15.58s + 13.04} \quad (7)$$

был реализован в Simulink NARMA-L2 регулятор. Результат верификации показал работоспособность синтезированного нейросетевого регулятора.

Позвольте теперь мне перейти к заключению по докладу.

Рассмотренная в докладе непараметрическая идентификация в отличие от параметрической идентификации не привязана жестко к техническим объектам и может использоваться и использовалась в различных задачах кибернетики. Теория непараметрической идентификации напрямую связана с методами решения некорректных задач и нейросетевыми технологиями. Как показали проведенные исследования, соответствующие алгоритмы идентификации могут быть практически реализованы в расширениях System Identification Toolbox и Neural Network Toolbox пакета MATLAB.

По-прежнему остается открытой проблема идентификации непараметрических нелинейных моделей Вольтерра-Винера. Однако значительное практическое продвижение в этой области следует ожидать, по-видимому, только после включения соответствующих алгоритмов или даже целых расширений в пакет MATLAB.

Литература

1. Винер Н. Нелинейные задачи в теории случайных процессов. – М.: ИЛ, 1961.
2. Пупков К.А., Капалин В.И., Ющенко А.С. Функциональные ряды в теории нелинейных систем. – М.: Наука, 1976.
3. Льюнг Л. Идентификация систем. – М.: Наука, 1991.



4. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. – М.: Наука, 1979.

5. Верлань А.Д., Сизиков В.С. Интегральные уравнения: методы, алгоритмы, программы. – Киев: Наукова думка, 1986.

ILL-POSED PROBLEMS AND NONPARAMETRIC IDENTIFICATION OF CONTROL SYSTEMS

V. I. KAPALIN

*Moscow State
Institute of Electronics
and Mathematics*

e-mail: v_kapalin@mail.ru

Problem of identification is connected with the problem of the solution of ill-posed problems. The current paper gives some essential features in that area and discusses ways of overcoming some difficulties in non-parametric methods of identification.

Key words: identification, the aprioristic information, «black box», нейросетевой the approach.

ИСПОЛЬЗОВАНИЕ МОДЕЛЕЙ ПАРЕТО ПРИ ФОРМИРОВАНИИ ИННОВАЦИОННОЙ ДЕЯТЕЛЬНОСТИ ТЕЛЕКОММУНИКАЦИОННОЙ КОМПАНИИ

А. П. ИГНАТЬЕВ
З. В. СМЕРНОВА

ЗАО «РК-ТЕЛЕКОМ»

e-mail :
aignatiev@rktelecom.ru
smzoya@rambler.ru

В статье рассмотрены вопросы использования моделей Парето в деятельности телекоммуникационной компании в современных условиях при анализе распределения инфокоммуникационных технологий для формирования инфокоммуникационных услуг с учетом прогнозов.

Ключевые слова: распределение, инновации, Парето, развитие, формирование пакета инфокоммуникационных услуг с учетом прогноза, ценообразование телекоммуникационных услуг.

Системный подход в исследованиях управления инновационной деятельностью интегрированных компаний, в том числе телекоммуникационной компанией «РК-ТЕЛЕКОМ» предусматривает выявление факторов, влияющих на этот процесс, связей и зависимостей, формирующих процесс управления, специфику и обязательные условия осуществления управляющих воздействий.

Инновации выполняют особую функцию в системе воспроизводства – функцию порождения изменений, являются источником саморазвития и самоорганизации предпринимательских систем как важнейший внутренний процесс и как структурообразующий элемент. Инновации составляют основу преобразований в научно-технических интеллектуальных системах, прежде всего, воздействуют на их структуру.

Для количественного определения действия системы воспроизводства, в терминах существования входных воздействий, необходимо ввести целевую функцию системы :

$$F : X \times \Omega \times Y \rightarrow W, \quad (1)$$

где W – множество действительных чисел. Если W имеет более одной компоненты

$$W = W^r = x \{W_i, i \in I_r\}, \quad (2)$$

где $I_r = 1, 2, \dots, N$ – число компонент (соответствующую систему называют многокритериальной). Иногда удобнее представлять целевую функцию в виде двух функций: выходной $P : X \times \Omega \rightarrow Y$ и функции (1), тогда

$$\Phi(x, w) = F(x, w, P(x, w)), w \in \Omega. \quad (3)$$

Функционал (3), описывающий действие всей системы, является функционалом эффективности. Реальные системы воспроизводства, как правило, имеют несколько целей и состоят из совокупности подсистем. Определим частные целевые функции (ЦФ) подсистем как

$$f_i : X_i * Y_i \rightarrow W, i \in I. \quad (4)$$

Тогда функционал (3) можно записать в виде

$$\Phi(x, w) = \Phi(f(x), w), \quad (5)$$

где $f(x) = \{f_i(x), i \in I\}$ – показатели качества подсистем.

В процессе управления инновационной деятельностью компании неизбежно сталкиваются с рядом затруднений, вызванных так называемой неопределенностью. Эта неопределенность связана с отсутствием точных и достоверных сведений, необходимых для принятия управленческих решений.

Неопределенности являются принципиальной неотъемлемой составляющей инновационного процесса, поскольку инновации неразрывно связаны с исследованиями и поиском нового и неизведанного. В условиях неопределенности выбор рациональных значений параметров систем может осуществляться как задача нахождения удовлетворительных решений: требуется найти такое $\hat{x} \in X^o$, что $\forall w \in \Omega$,

$$\Phi(\hat{x}, w) \geq \tau(w), \quad (6)$$

где $\tau(w)$ – функция, определяющая минимально допустимое значение целевой функции. Из-за высокого уровня абстракции множество Ω охватывает как параметрические, так и структурные неопределенности, т.е. фактически является множеством всех факторов, влияющих на решение задачи (6). Заметим, что наиболее важным вопросом в задаче (6) является нахождение функции $\tau(w)$, которая определяет минимальное или допустимое качество системы при любых проявлениях неопределенности $w \in \Omega$. Вид $\tau(w)$ зависит как от свойств функции $\Phi(x, w)$, так и от типа неопределенности, имеющей место на ранних этапах проектирования сложной системы.

Однако их можно свести к трем основным типам:

$\Omega^{(1)}$ – множество неопределенностей, обусловленных средой. Этот тип носит статистический характер и может быть описан методами теории вероятностей;

$\Omega^{(2)}$ – множество неопределенностей, обусловленных целенаправленным противодействием (вызваны незнанием поведения конкурирующих фирм), обычно учитывается методами теории игр;

$\Omega^{(3)}$ – неопределенности, связанные с теми неточностями описания, которые в принципе не могут быть оценены статистически. Например, на этапе внешнего проектирования имеет место нечеткое знание состояния и возможностей внутреннего проектирования и т.п.

Последний вид неопределенностей в настоящее время все чаще связывается с нечеткими множествами, т.е. классами объектов, в которых нельзя указать резкую границу, отделяющую объекты, принадлежащие к данному классу от объектов, не принадлежащих к нему. В задаче управления инновационным процессом, формализуемой в виде модели многокритериальной оптимизации (6), с учетом соотношений (2) и (5), первым и наиболее важным считается выделение области компромиссов, оптимальных по Парето (по имени швейцарского экономиста Вильфредо Парето).

Диаграммы Парето можно использовать при решении любых проблем, возникающих в деятельности телекоммуникационной компании (далее – ТК), таких как: трудности с оборотом кредитных сумм, с освоением новых правил принятия заказов, удлинению времени от получения заказа от клиентов до его исполнения, наличия невостребованных услуг, на которые отвлекаются средства компании, поступление рекламаций, количество которых не уменьшается невзирая на старания повысить качество услуг.

Диаграмма Парето используется и когда опыт управляющего центра распространяется на всю структуру телекоммуникационного холдинга, и в противоположном случае, когда положительный опыт отдельных подразделений хотят внедрить на всем предприятии. С помощью диаграммы Парето выявляют основные причины успехов и широко пропагандируют эффективные методы работы.

При использовании диаграммы Парето для контроля важнейших факторов наиболее распространенным методом анализа является так называемый ABC-анализ.

Схематически анализ Парето можно изобразить так (рис.1):

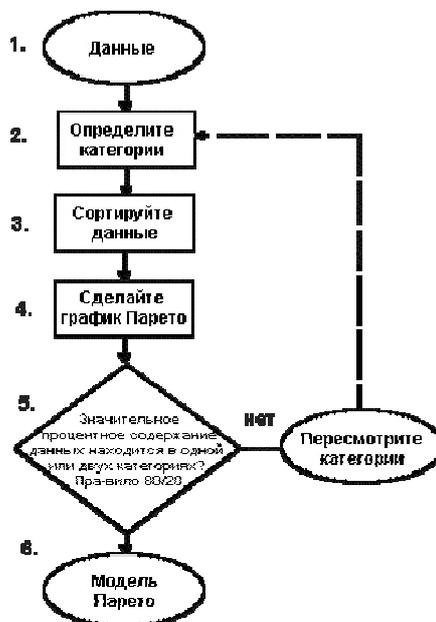


Рис. 1. Схема анализа Парето

Анализ Парето – это способ организации данных, позволяющий выявить наиболее существенные факторы, характеризующие анализируемый объект.

График Парето – это тип графика, в котором строятся полосы в нисходящем порядке, начиная слева. Основой графика Парето является правило "80-20"; 80% проблем являются результатом 20% причин.

Телекоммуникационный (ТК) холдинг с точки зрения корпоративного контура (одна компания владеет акциями/паями/долями других) может быть монокомпанией с точки зрения управленческого контура. И наоборот, холдинг в смысле управленческого контура может быть одним юридическим лицом, поскольку выделяемые в управленческом учете бизнес-направления могут и не быть отдельными юридическими лицами.

Но любой холдинг – это не обособленное образование вне какой-либо предметной области, главной конечной целью которого является достижение внутреннего управленческого баланса между управляющим центром и структурными единицами. Цель любого хозяйствующего субъекта – оптимальное развитие с наращиванием доли рынка существования.

По классическим законам развития ТК утверждается, что объем производственной информации, созданный в стране за год, пропорционален валовому национальному продукту (ВВП) и имеет линейную зависимость. Информационно-экономический закон подтверждается многочисленными реальными примерами развития и расширения телефонной сети общего пользования в мире.

Информационно-экономический закон требует соблюдения принципа пропорционально-опережающего развития ТК страны, т.к. отрасль заблаговременно должна быть готова к растущему обмену информации для последующего обеспечения экономического роста страны.

Наиболее характерной закономерностью развития средств и услуг ТК является логистический закон, который характеризуется тремя уровнями развития ТК:

- начальный этап (этап линейного развития);
- этап быстрого развития, который характеризуется экспоненциальным ростом;
- этап насыщения (когда спрос на услуги ТК полностью удовлетворяется).

Таким образом, постепенно старая услуга заменяется новой, более современной, что сопровождается сменой технологий. Ярким примером вышеуказанного является процесс развития коммутационной техники от ручных, электромеханических и координатных до современных цифровых систем коммутации (ЦСК).

Рассмотрим общие закономерности распределения доходов и услуг (учрежденного государством):

- Закон распределения больших доходов (закон Парето);
- Закон распределения 20/80;
- Закон распределения доходов и услуг.

Первый эмпирический закон Парето (1895 год), по имени швейцарского экономиста Вильфредо Парето, определяет, что вероятность доходов больше величины A и равна $1/A^\alpha$, где $\alpha > 1$ – некоторый параметр, полагаемый $\alpha \approx 1,5$.

Второй – практическое правило, утверждающее, что 20% населения обладают 80% доходами. Правило относится ко многим сферам деятельности человека, в том числе производительности труда, научным исследованиям и т.д., поэтому его называют правилом 20/80. Многие экономисты и прогнозисты в мире отождествляют указанные выше Закон Парето и правило 20/80 {правило $P/(1-P)$ }.

Третий закон доказан российскими учеными-связистами, которые на основе математических методов установили асимптотическую сходимость правила 20/80 к Закону Парето.

Здесь интегральное распределение дохода Q среди населения P описывается ниже следующим соотношением:

$$Q(P) = 1 - (1 - P)^\alpha, \text{ где } \alpha = (\alpha - 1) / \alpha (1)$$

Данная формула более наглядно изображается кривыми интегрального распределения, которые по характеру очень близки к кривым Лоренца. Установлено, что чем ближе α к единице, тем более неравномерным становится распределение дохода, с уменьшением доли населения, обладающей максимальным доходом страны.

Распределение различных услуг ТК (например, обычные (ОТА) и мобильные телефоны (МТ), персональные компьютеры (ПК) и Интернет (Инт.) в зависимости от объема ВВП представлено на рис.2.

Как видно, новые и дорогостоящие технологии имеют неравномерное распределение по сравнению с ранними, традиционными технологиями. Это объясняется тем, что новые технологии и услуги имеют более высокие цены, и поэтому ими пользуются состоятельные члены общества.

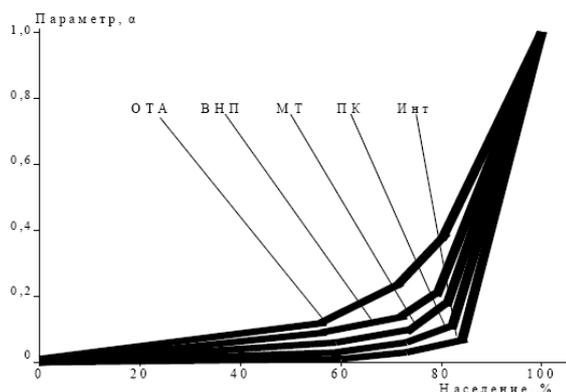


Рис. 2. Распределение инфокоммуникационных технологий среди населения мира (по данным ИТУ – Международного Союза Телекоммуникаций)

По истечению некоторого времени данная технология улучшается, удешевляется, и цена на услуги связи уменьшается, ее распространение становится менее не-

равномерным, и параметр-альфа растет, что продолжается до появления новой технологии и т.д.

Таким образом, для определения объема ТК рынка необходимо знать два параметра – уровень развития экономики (ВВП) и распределение доходов среди населения страны, т.е. иметь соответствующую кривую Лоренца.

Однако во многих развивающихся странах мира с переходной экономикой существует «черный рынок», не контролируемый государством, который обеспечивает практический спрос на ТК рынке. При наличии данных о «черном рынке», например β , он достаточно просто может быть учтен.

Например, пусть предполагаемый черный рынок равен β ВВП, где $\beta < 1$, и этот коэффициент прогнозируется правительством или прессой. Тогда в расчетах и прогнозах необходимо учитывать суммарный доход, т.е. $(1+\beta)$ ВВП, следовательно, параметр α^* равен:

$$\alpha^* = (1 + \beta) \alpha / [1 + (1 - \beta) \alpha] \quad (2)$$

Определив указанные параметры, можно приступить к определению трех периодов прогноза:

- краткосрочный (от 3-5 лет) – охватом до 20% населения;
- среднесрочный (от 7-10 лет) – охватом до 50% населения;
- долгосрочный (от 10-25 лет) – для всего населения.

Правило Парето также применимо при ценообразовании телекоммуникационных услуг при их производстве.

На телекоммуникационном рынке ярко проявляется «Правило 20/80» — универсальный принцип, сформулированный итальянским экономистом и социологом Вильфредо Парето в 1897 году. Позднее этот принцип исследовал англичанин Ричард Кох. Принцип Парето или принцип 20/80 означает, что 20% усилий дают 80% результата. Применительно к телекоммуникациям, исходя из «правила 20/80», можно сделать такие выводы:

- 20% конечных потребителей приносит 80% доходов;
- 20% конечных потребителей создают 80% нагрузки;
- для 80% потребителей достаточно 20% из набора возможных услуг.

Эти, на первый взгляд, парадоксальные результаты весьма полезны как для маркетинговой стратегии участников рынка, так и для специалистов, обеспечивающих создание и развитие телекоммуникационных сетей.

В апреле 2006 года американская исследовательская организация Pew Internet and American Life Project совместно с Associated Press и Интернет-компанией AOL в результате массового опроса пользователей мобильной связи получила интересные результаты, подтверждающие принцип Парето. Так, количество респондентов, использующих или желающих использовать в своем телефоне подвижной связи возможности мобильного контента, не превышает 20%. При исследовании возможностей Raylay-шлюзов выяснилось, что для представления 80 % услуг требуется лишь 20 % возможностей шлюза.

Важной тенденцией телекоммуникационного рынка является ежегодное снижение тарифов на услуги в среднем на 10 процентов.

Ещё одно знаменательное наблюдение последних лет: удельные доходы от передачи одного байта информации ежегодно существенно снижаются.

Безусловно, в конкретных проявлениях этого принципа можно наблюдать отклонения от абсолютного соотношения 20/80.

Заключение

Представлен научный материал по системно-экономическому подходу использованию моделей Парето для формирования инновационной деятельности телекоммуникационной компании. На примерах описаны различные методы использования Правила Парето при анализе вектора инновационной деятельности теле-



коммуникационной компании: для определения вектора её направления на основании анализа ретроспективных данных, для ценообразования телекоммуникационных услуг при их производстве.

Литература

1. Игнатьев А.П. Совершенствование системного менеджмента для телекоммуникационных компаний в условиях инновационного производства. Изд. Белгород, 2009, 248стр.
2. Модели оценки потребителем эффективности вложений в информационные технологии: www.ibir.ru
3. Шатраков А.Ю., Журавлева Э.М., Парфенова М.Я. Принятия решения экономической среде Изд. Москва, ГОУ «МАРТИТ» 2004г.
4. Дёмин В.К., Тютин Н.Н., Храмешин Г.К., Чудинов С.М. Региональные информационные системы методы их структурной и функциональной оценки. Изд. Белгород, 2008, 320стр.
5. Смирнова З.В., Меркулова А.В., Методы экономического механизма обеспечения инвестиционной деятельности предприятий телекоммуникационного комплекса М.: МАРТИТ, «Региональная экономика» № 11 (33), 2006 г.
6. Злобин В.П., Методы и инструменты измерения и повышения качества в строительстве и проектировании: <http://www.iso9001.ru/publik-27.html>.
7. Ребрин Ю.И., Управление качеством, учебное пособие. Таганрог: Изд-во ТРТУ, 2004. http://www.aup.ru/books/m93/4_4.htm.

USE OF MODELS OF PARETO AT FORMATION OF INNOVATIVE ACTIVITY OF THE TELECOMMUNICATION COMPANY

A. P. IGNATEV
Z. V. SMIRNOVA

JSC «RK-TELECOM»

e-mail:
aignatiev@rktelecom.ru
smzoya@rambler.ru

Questions use of models of Pareto in activity of the telecommunication company in modern are considered at the distribution analysis information technologies for formation infocommunication services taking into account forecasts.

Keywords: distribution, innovations, Pareto, development, package formation infocommunication services taking into account the forecast, pricing of telecommunication services.

ОСНОВЫ И ПЕРСПЕКТИВЫ ПРИМЕНЕНИЯ ИНФОРМАЦИОННЫХ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ДЛЯ ПРОЦЕССОВ ЖИЗНЕННОГО ЦИКЛА НАУКОЕМКИХ ИЗДЕЛИЙ

А. Е. БРОМ
В. А. ШУТЕЕВ

*Московский
государственный
технический
университет
им. Н.Э. Баумана*

*e-mail:
aspirant.mgtu@gmail.com*

В статье описаны задачи, стоящие на первых двух этапах жизненного цикла изделия. Предлагается автоматизировать связанные с ними процессы путем внедрения мультиагентов. Также приведены примерные схемы мультиагентных систем с подробным описанием каждой из них.

Ключевые слова: мультиагентные системы, жизненный цикл, наукоемкая продукция.

Введение

Мультиагентные системы (МАС) – технология, появившаяся относительно недавно и еще не получившая широкого распространения ни в научных работах, ни в производстве, но представляющая перспективной благодаря возможности вести работу децентрализованно. Эта тема нашла свое отражение в таких трудах, как “Интеллектуальные информационные системы” [1], “От многоагентных систем к интеллектуальным организациям: философия, психология, информатика” [2], “Введение в мультиагентные системы” [3]. В работе “Практические и промышленные применения мультиагентных систем” [4] описываются разрабатываемые МАС для практического применения на производстве, в то время как другие труды останавливаются в основном на теоретической базе таких систем, уделяя основное внимание темам координации агентов, механизмов переговоров, онтологии знаний агентов.

Ни в одном из исследований не затрагивалась тема постановки задач для поддержки жизненного цикла наукоемкой продукции, выполнение которых можно было бы поручить МАС. Цель данной статьи – описать задачи, стоящие на начальных этапах жизненного цикла, с тем, чтобы можно было распределить их между различными агентами для более эффективной информационной и организационной поддержки жизненного цикла. Но прежде, чем переходить к постановке задач, разберем значение некоторых понятий.

Жизненный цикл изделия – это совокупность этапов или последовательность бизнес-процессов, через которые изделие проходит за время своего существования.

Под конечным изделием понимается комбинация материалов, предметов, программных и иных компонентов, готовых к использованию по назначению. Компоненты конечного изделия тоже являются изделиями. Данные об изделии составляют основной объем информации в интегрированной информационной системе. Международные стандарты ИСО 10303 и ИСО 15384 регламентируют технологию представления данных об изделии и его компонентах на стадии проектирования и подготовки производства, стандарты ИЛП [DEF STAN 0060] – представление данных об изделии в контексте обеспечения эффективной эксплуатации, стандарты серии ИСО 9000 рассматривают данные о качестве изделий.

В системе международных стандартов ISO 9004 определены следующие этапы жизненного цикла:

1. Маркетинг (поиск, изучение и анализ рынка).

2. Исследования, проектирование и разработка технических требований к создаваемому продукту.

3. Материально-техническое снабжение (закупка материалов и комплектующих).

4. Подготовка и разработка технологических процессов.

5. Производство (предоставление услуг).

6. Контроль, проведение испытаний и обследований.

7. Упаковка и хранение.

8. Реализация и хранение.

9. Монтаж (установка и ввод в эксплуатацию) и эксплуатация.

10. Послепродажное техническое обслуживание.

11. Утилизация.

Информационная логистическая система управления жизненным циклом наукоемкой продукции, включающая информационную поддержку всех вышеперечисленных этапов, разрабатывается на основе концепции МАС.

МАС можно считать сетевую совокупность в разной степени самостоятельных или автономных объектов, называемых агентами, способных получать, хранить, обрабатывать и передавать информацию в интересах решения как собственных, так и общих для группы объектов задач анализа и синтеза информации. [1]

Агентом такой системы может быть аппаратная, программная или аппаратно-программная сущность, способная действовать как в собственных, так и в общесистемных интересах. [1]

Главная особенность системы и наполняющих ее агентов состоит в их автономности, которая трактуется как способность действовать без вмешательства извне и осуществлять объективный контроль своего состояния и результатов своих действий.

Именно это требуется от систем поддержки жизненного цикла наукоемкой продукции. Все этапы жизненного цикла могут обслуживаться агентами с целью более эффективного планирования и мониторинга процессов.

Ниже приведем процессы на первых двух этапах жизненного цикла, которые могут подвергнуться автоматизации и внедрению интеллектуальных агентов.

Этап маркетинга

Целью данного этапа является выявление потребностей потенциальных потребителей (потребности рынка), определение их характеристик (возраст, пол, экономическая состоятельность, социальное положение и др.), определение того товара, который сможет удовлетворить потребность потенциальных потребителей (в соответствии с концепцией маркетинга) и его основных характеристик. Также данный продукт должен быть конкурентоспособным, чтобы иметь успех у потребителя.

Можно выделить следующие задачи:

Необходимо использовать открытые источники для анализа потребностей потребителей. В виду все большей доступности текстов статей журналов в сети интернет, а также того, что большое количество конференций проходит в интернет-формате, можно выделить задачу сбора информации в сети. Информация должна быть собрана, сохранена в должном и удобном для применения виде, проанализирована и отфильтрована. Для этого целесообразно использовать базу данных (БД). Необходимо реализовать интеллектуальный поиск по БД. Желательно также иметь каталог изделий с возможностью поиска по выполняемым функциям.

Нужна БД отзывов, предложений, корректив, вносимых постоянными клиентами, которые используют или использовали данную продукцию прежде, или использовали продукцию со схожей функциональностью.

Следует держать руку на пульсе новых технологий, которые могут увеличить успешность продукции на рынке. Нужно задействовать процессы, которые могут оп-

ределять, идентифицировать и интерпретировать новые технические веяния для эволюции используемых технологий.

Этап исследования, проектирования и разработки технических требований к создаваемому продукту

На данном этапе определяются функциональные характеристики изделия, создается его спецификация. Изделие проектируется.

Часто коллектив проектировщиков состоит из нескольких групп, территориально удаленных друг от друга, которые совместно работают над разными узлами в рамках одного проекта (или разработки). Встает задача их координации для создания виртуальной группы разработчиков. Подразумевается, что необходимо составлять рабочее расписание, следить за сроками разработки, мотивировать членов команды, адресуя необходимые сентенции в адрес работников. Поэтому нужно решать задачу напоминания о сроках удаленно; проводить координацию работы.

В каждой отдельно взятой группе могут быть использованы свои собственные наборы средств разработки (CAD). В то же время, для совместной работы групп разработчиков необходима синхронизация и совместимость используемых ими форматов данных без применения сторонних средств и потери времени на конвертацию. Возможен выбор определенного промежуточного формата, который будет поддерживаться агентами всех разработчиков.

При разработке определенного узла перед специалистом стоит задача подбора функциональной единицы под заданные параметры: габаритный размер, вес, функциональность (мощность, скорость и др.), стоимость, доступность. Для оптимального выбора помощь интеллектуальных агентов будет весьма кстати.

Может так случиться, что две группы разработчиков (или два разработчика) будут иметь различное мнение о применении в проектируемом изделии того или иного узла. Свою помощь в разрешении споров могут предоставить интеллектуальные агенты через систему торгов. Тогда разработчики, обладая ресурсом, могут вложить его в одно из альтернативных решений. В результате будет выбрана самая востребованная альтернатива.

В случае необходимости принятия очень важных решений все участники проекта извещаются о факте их принятия. При этом они могут среагировать негативно и предложить свою альтернативу. Таким образом, каждое важное решение будет под наблюдением целой группы.

Все решения о внедрении того или иного агрегата, изменения в дизайне; идеи, возникающие во время проектирования, должны быть отмечены и сохранены в БД для будущих потребностей в документации процессов, поиска альтернатив для замены и др. История принятия всех решений будет использована в будущем.

Интеллектуальные агенты могут быть использованы для помощи в создании документации посредством анализа истории принятия решений, использования принятых решений о выборе функциональных узлов, нахождения их спецификаций и другой аналитической и синтетической деятельности.

Для задач проектирования изделия, в которых можно четко указать состав его функциональных узлов и требуется подобрать на их место конкретные реализации (альтернативы), интеллектуальные агенты могут составить несколько готовых альтернатив и, определив их плюсы и минусы, проранжировать по требуемому критерию.

При проведении аудио- и видеоконференций между территориально удаленными группами разработчиков требуется решить организационные вопросы: получение всеми сторонами расходных материалов, определение очередности выступлений и сохранение регламента, обеспечение бесперебойной трансляции аудио- или видеосигнала всем заинтересованным сторонам. Это задачи для децентрализованной МАС.

Индивидуальные агенты проектировщиков могут понимать предпочтения их владельца, учиться обеспечивать комфорт для конкретного человека, делая вирту-

альное рабочее пространство максимально эффективным, а также следить за теми ограничениями, которые предъявлены для данной задачи. Агент может подсчитывать стоимость, вес и другие ограничения у каждого дизайнера – и сообщать, если что-то не так.

Во многих научных работах по процессу параллельного проектирования подчеркивается проблема непонимания разработчиками границ своей ответственности в рамках большого проекта. Т.е. часто непонятна их конкретная цель в большом проекте. Можно предложить создание подробного графического и текстового описания целей всех уровней (что невозможно изобразить на бумаге и представить на совещании из-за опускания «мелочей») с предоставлением каждому из членов команды (разработчиков, лидеров разработки, управленцев) тех, которые поставлены именно перед ним. В каждом случае будут выделены основные задачи, что позволит не потерять цель из вида. Таким образом, будет запущен процесс визуализации задач.

Проблема распределения задач для решения в конкретной группе разработчиков может решаться интеллектуальными агентами на основе уже решенных задач, оценки их результатов и, конечно, предпочтений.

Журналирование этапа проектирования и разработки необходимо для последующего применения на других стадиях жизненного цикла. Задачей распределения и сортировки документов, созданных на этапе проектирования, может заниматься МАС.

Примеры возможного использования МАС для решения задач, стоящих на первых двух этапах ЖЦ

МАС может обеспечивать поддержку поиска информации для маркетингового исследования, актуальной для данной отрасли. При этом она должна взять на себя все рутинные операции: поиск ценной информации, сохранение, фильтрацию и предоставление готового отчета пользователю.

Ниже приведена схема с 6 типами агентов, задействованных в такой МАС, и принципиальными взаимодействиями между ними (рис.1).

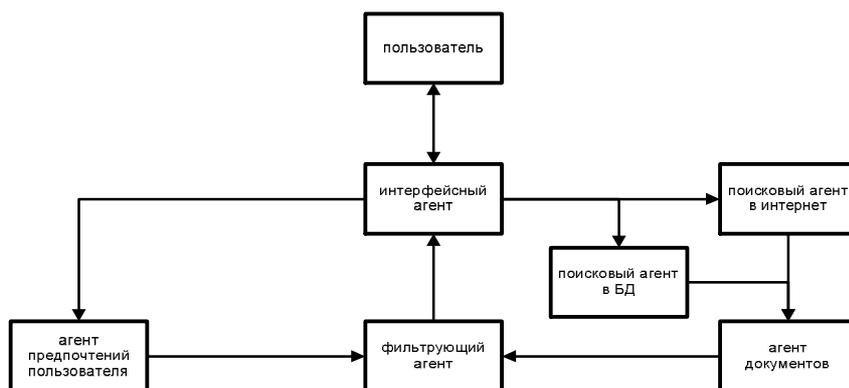


Рис. 1. Схема мультиагентной системы поиска информации

Каждый из агентов выполняет определенную функцию. Опишем их необходимость и выполняемые функции, а также цели их взаимодействия.

Интерфейсный агент – агент, напрямую взаимодействующий с пользователем через графический интерфейс. Пользователь вводит запрос на интересующую тему, выбирает из списка нужное, расставляет приоритеты – так или иначе, передает агенту входные данные. Результаты выдаются в виде множества документов, таблиц, графиков, рисунков. Можно сказать, что это агент, представляющий интересы пользователя в данной информационной МАС. Зачастую, только через него идет коммуникация между пользователем и мультиагентной средой.

В данной схеме МАС интерфейсный агент предоставляет интерфейс идентификации пользователя, форму запроса и выбора предпочтений, а также отображение найденной информации на экране.

Прежде всего, пользователь должен идентифицировать себя для более эффективного поиска (и активации агента предпочтений, считывающего информацию по данному пользователю из БД), после чего ему необходимо ввести текстовую строку, по которой следует проводить поиск в Интернете и БД. Дополнительной возможностью может быть функция определения широты поиска, используемых источников информации, временных ограничений. После получения данных такого типа от пользователя интерфейсным агентом, происходит передача их агенту предпочтений пользователя и поисковым агентам (в БД и Интернет).

Агент предпочтений пользователя отвечает за метаинформацию о пользователе. Данный агент содержит информацию о пользователе, предпочитаемый стиль выдачи результата, анализ ранее произведенных поисковых запросов и тех результатов, которые оказались наиболее востребованными для данного пользователя. Это личный агент пользователя, делающий его работу наиболее комфортной, а значит – эффективной. От интерфейсного агента он получает текущие формализованные предпочтения пользователя. На основании совокупности этих данных агент может сформировать требования к выдаваемой пользователю информации. Эта информация передается фильтрующему агенту для дальнейшей работы.

Поисковый агент в Интернет, получив запрос и ограничения на поиск от интерфейсного агента, отправляет ряд запросов на наиболее релевантные поисковые сайты по Интернет (из списка определенных пользователем). Получив множество подходящих по мнению поискового робота результатов, этот агент проводит их тщательный анализ. Он осуществляет его посредством «чтения» содержания документов, на которые были получены ссылки, для определения релевантности данной информации. Документы, прошедшие проверку, сохраняются и по окончании всего процесса передаются агенту документов.

Поисковый агент в БД обладает схожими функциями с поисковым агентом в Интернет. При этом поиск производится в существующей БД, наполненной структурированной информацией, по патентам, функциональным описаниям и другим специфичным, и поэтому необходимым на данном этапе документам. Агент сам выстраивает запросы к СУБД и получает документы, которые так же, как и в случае с поисковым агентом в Интернет, необходимо проверить на релевантность заданному запросу.

Таким образом, поисковые агенты в Интернет и БД проводят первую фильтрацию документов. Это необходимо для передачи уже релевантных документов агенту документов.

Агент документов получает документы от поисковых агентов в их собственных форматах представления данных и представляет их в едином формате, удобном для передачи и чтения фильтрующему агенту.

Фильтрующий агент получает документы, прошедшие первичное определение релевантности поисковыми агентами, от агента документов, и предпочтения пользователя от агента предпочтений. Результат фильтрации документов с использованием предпочтений пользователя передается интерфейсному агенту, который отображает его пользователю.

Пользователь применяет полученную информацию, оценивает ее, и эти данные отправляются агенту предпочтений пользователя для обновления предпочтений в целях дальнейшей эффективной работы.

Для задач журналирования стадии разработки необходима МАС, которая будет заниматься оценкой важности документов, их распределением по соответствующим тематике категориям для быстрого и легкого доступа, сохранением их изменений для возможности возврата к предыдущим версиям.

Посредством интерфейсного агента пользователь задает название или содержание документа, который ему необходимо получить из БД. Данные для сохранения в БД он передает также через этого агента – с возможностью ввести название, характеризующие его метки и описание.

Агент документооборота создает и удаляет агенты по мере их необходимости, организует коммуникацию между ними.

Агент сравнения – агент, который сравнивает документ, полученный от интерфейсного агента, с его более ранней версией, предоставленной агентом БД. Информацию об изменениях он передает агенту документооборота для анализа и последующего сохранения в БД (через агента БД).

Агент модели данных – агент, представляющий ту или иную категорию, в которую будет отнесен документ. Он обладает ключевыми данными для определения его тематики и сохранения в соответствующем разделе в необходимом формате.

Агент БД получает запросы от других агентов, после чего перенаправляет их к БД для получения данных. Затем он отправляет результаты обратившимся к нему агентам для дальнейшей обработки.

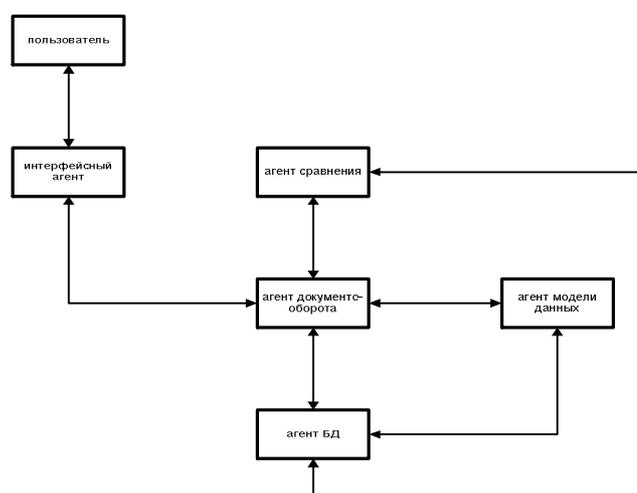


Рис. 2. Схема мультиагентной системы поддержки документооборота

Выводы

Выше представлены лишь некоторые примеры использования МАС для решения задач управления жизненным циклом наукоемкой продукции; приведены выделенные задачи из числа процессов, проходящих на первых этапах; и предложены схемы МАС для решения некоторых задач. В настоящее время выделяются задачи на остальных этапах жизненного цикла наукоемкой продукции. В дальнейшем будут предложены механизмы коммуникации для успешной реализации приведенных схем.

Литература

1. Дубровин, А.Д. Интеллектуальные информационные системы: Учебное пособие для студентов факультета МИСИТ Московского Государственного университета культуры и искусств, обучающихся по специальности 080801 – прикладная информатика (в менеджменте)/ А.Д. Дубровин, науч. ред. О.В. Шлыкова. – 1-ое изд. – М.: МГУКИ, 2008. – 232 с.
2. Тарасов, В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика / В.Б. Тарасов. – 1-ое изд. – М.: Эдиториал УРСС, 2002. – 352 с.
3. Wooldridge, M. An Introduction to Multi-Agent Systems / M. Wooldridge. – 1-st ed. – Chichester: John Wiley and Sons, 2002. – 366 p.
4. Parunak, H. Van Dyke. Practical and Industrial Applications of Agent-Based Systems / H. Van Dyke Parunak.- <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.34.9519>, 1998.



BASICS AND PERSPECTIVES OF USING INFORMATION INTELLIGENT SYSTEMS TO SUPPORT LIFE CYCLE PROCESSES OF HIGH TECHNOLOGY PRODUCTS

A. E. BROM
V. A. SHUTEEV

*Moscow state
technical university
named after N.E. Bauman*

*e-mail:
aspirant.mgtu@gmail.com*

The article considers the tasks of the first two stages of a life cycle. strategy of budget management of a region and the budget policy of a region. An idea of automatization the cocurrent processes by implementing multiagents is raised. Exemplary schemes of multiagent systems with detailed descriptions are provided.

Key words: multiagent systems, life cycle, high technology products.

ГРУППОВОЕ МНОГОКРИТЕРИАЛЬНОЕ ПРИНЯТИЕ РЕШЕНИЙ С НЕСОВПАДАЮЩИМИ ПРЕДПОЧТЕНИЯМИ*

А. Б. ПЕТРОВСКИЙ

*Институт
системного
анализа
Российской
академии наук*

e-mail: pab@isa.ru

Рассматривается новый подход к групповому многокритериальному принятию решений, где предпочтения нескольких лиц, принимающих решения, не совпадают и, возможно, противоречивы, а варианты решения оцениваются по многим разнородным критериям. Подход основан на методологии вербального анализа решений и теории мультимножеств. Приведены примеры применения подхода для решения практических задач классификации и упорядочения многопризнаковых объектов.

Ключевые слова: групповое принятие решений, теория мультимножеств, классификация, упорядочение, многопризнаковые объекты.

Введение

Типичными задачами теории принятия решений является выделение одного или нескольких лучших объектов (вариантов, альтернатив, кандидатов, и т.п.), упорядочение или ранжирование объектов по их свойствам, классификация или сортировка объектов по заданным категориям. Свойства рассматриваемых объектов характеризуются многими разнородными признаками (атрибутами) или оценками по многим критериям, имеющим количественные и/или качественные шкалы.

Решение задач индивидуального выбора осуществляется на основе предпочтений единственного лица, принимающего решение (ЛПР). Поэтому требуется, чтобы его субъективные предпочтения были согласованными. При наличии нескольких независимых ЛПР коллективный выбор зависит от различных предпочтений многих участников, которые не обязаны совпадать. Несогласованность индивидуальных суждений обуславливается разнообразием и несовпадением целей и интересов разных ЛПР, собственным пониманием решаемой задачи, специфичностью знаний самих ЛПР, и многими другими обстоятельствами. Еще одним источником несогласованности информации в задачах группового выбора служит неоднозначность представления объектов, когда один и тот же объект может существовать в нескольких версиях (экземплярах), имеющих, в частности, и противоречивые описания. При этом сами объекты должны рассматриваться и анализироваться как единое целое, а свертка значений признаков или невозможна, или математически некорректна. Множественность представления многопризнаковых объектов возникает, например, когда объекты оцениваются несколькими экспертами по многим критериям.

Известно достаточно много методов индивидуального выбора объектов, описываемых многими количественными и/или качественными признаками [3, 6, 8-11, 15-18]. Сравнительно немного методов группового упорядочения многопризнаковых объектов [6, 12] и практически отсутствуют методы групповой классификации. Главные трудности при коллективном принятии решений связаны с необходимостью учета несогласованных и противоречивых предпочтений нескольких ЛПР, а также обработкой больших объемов числовых и вербальных данных, не прибегая к дополнительным преобразованиям типа усреднения, смешивания, взвешивания, которые могут привести к необоснованным и необратимым искажениям исходной информации.

* Работа поддержана программами фундаментальных исследований президиума РАН «Интеллектуальные информационные технологии, математическое моделирование, системный анализ и автоматизация» и ОНИТ РАН «Информационные технологии и методы анализа сложных систем», Российским фондом фундаментальных исследований (проекты 08-01-00247, 08-07-13532, 09-07-00009, 09-07-12111).

Преодолеть указанные выше трудности оказалось возможным, воспользовавшись аппаратом теории множеств для представления многопризнаковых объектов и обработки данных. В работе рассмотрены методы групповой классификации и упорядочения объектов, существующих в нескольких версиях, которые описываются многими разнородными и несовпадающими признаками. Методы основаны на теории метрических пространств множеств и использованы в практических задачах группового многокритериального принятия решений.

1. Методы многокритериального принятия решений

Предпочтения ЛПР являются ключевым фактором рациональности выбора. ЛПР формализует свои предпочтения, задавая характеристики анализируемой проблемы и свойства рассматриваемых объектов, сравнивая варианты решения, оценивая качество сделанного выбора. Предпочтения могут быть заданы бинарными отношениями, функциями (ценности, полезности, целевыми), решающими правилами, которые имеют логическую, математическую и вербальную форму. В то же время, решая задачу, человек может выражать свои предпочтения непоследовательно, совершать ошибки, допускать противоречия.

В случае индивидуального выбора постулируется согласованность субъективных предпочтений отдельного ЛПР, которая обычно состоит в транзитивности его суждений. Для выявления и устранения возможных несоответствий в суждениях единственного ЛПР в методах принятия решений применяют специальные процедуры.

Коллективный выбор, осуществляемый группой независимых ЛПР, принципиально отличается от индивидуального выбора, так как связан с необходимостью агрегирования многих предпочтений. Каждое ЛПР может иметь свои персональные ценности, собственные цели, интересы и источники информации. Как следствие, индивидуальные субъективные предпочтения нескольких персон могут быть схожими, различными и несовпадающими. В известных методах группового принятия решений обычно стремятся исключить несогласованность и противоречивость индивидуальных суждений, либо заменяют совокупность многих мнений одним общим мнением, выражающим некоторую «усредненную» точку зрения или наиболее согласованную со всеми суждениями. Вместе с тем, индивидуальные предпочтения не всегда допустимо усреднять или согласовывать. Поэтому существует потребность в новых подходах, обеспечивающих возможность рационального выбора, основываясь на противоречивой информации.

Перечислим кратко некоторые известные методы принятия решений [6]. Выделение лучших вариантов проводится с помощью разнообразных методов оптимального выбора, основанных на поиске экстремума одной или многих функций, характеризующих эффективность или качество решения. В методах многокритериальной оптимизации и многомерной полезности обобщенный критерий обычно задается сверткой многих частных числовых критериев в виде «взвешенной суммы». Однако определение весов исходных показателей является достаточно серьезной проблемой. При большом числе частных критериев построение обобщенной функции полезности сопряжено с большими трудозатратами ЛПР.

Для упорядочения объектов в целом или по многим критериям часто используются методы, основанные на парных сравнениях объектов. При наличии многих критериев и/или нескольких ЛПР результирующее упорядочение объектов строится в методе TOPSIS [6, 12] на основе парных сравнений векторов оценок, представляющих объекты. В методах аналитической иерархии [6, 8] варианты ранжируются по их приоритетности, которая последовательно вычисляется путем попарного сравнения вариантов, критериев их оценки и участников по отношению к глобальной цели решаемой задачи. Полное упорядочение объектов получается, если сравнимы все пары вариантов и транзитивны предпочтения единственного ЛПР. Если некоторые из вариантов окажутся несравнимыми, то упорядочение будет частичным. Если же все варианты несравнимы, то упорядочить их не удаётся. При увеличении числа объектов,



критериев и ЛПР число возможных сопоставлений резко возрастает. В таких случаях построение итогового упорядочения объектов существенно затрудняется из-за возможных неточностей и противоречий в суждениях ЛПР.

Прямая сортировка объектов по заданным классам – один из наиболее популярных методов классификации вследствие простоты применения. Каждый объект, оцененный по одному числовому критерию, немедленно попадает в один из указанных классов. Достаточно часто для упорядочения и классификации объектов применяются функции ценности или полезности. Весьма распространены методы, основанные на взвешенной свертке критериев, например, методы многомерной полезности MAUT [6]. Порядковая сортировка многопризнаковых объектов в методе TOMASO [15] проводится с помощью интеграла Шоке, который агрегирует семейства дискриминантных функций. В интерактивной процедуре классификации [13] предпочтения единственного ЛПР описываются линейной функцией полезности, являющейся взвешенной суммой многих скалярных критериев. Однако эти и подобные им методы не позволяют объяснить полученные результаты, поскольку невозможно восстановить исходные данные по агрегированным показателям. Кроме того, назначение ЛПР весов критериев является субъективной и не имеющей строгих обоснований процедурой, а построение функции полезности при большом числе критериев сопряжено с большими трудозатратами ЛПР.

В семействе методов ELECTRE [16-18] объекты оцениваются по многим критериям, имеющим балльные шкалы и разные веса, которые задаются ЛПР. Объекты сравниваются попарно по так называемому отношению ограниченной предпочтительности (outranking relation), которое определяется специальными индексами согласия (конкорданса) и несогласия (дискорданса). Значения этих индексов вычисляются в процессе решения задачи и используются для упорядочения объектов и построения границ классов. Классификация многокритериальная альтернатив с применением огрубленных множеств предложена в [9, 11]. Предпочтения ЛПР выражаются с помощью наборов решающих правил, которые с разной степенью определенности относят альтернативы в заданные классы. Метод оперирует с достаточно большим числом решающих правил классификации, трудным для непосредственного анализа ЛПР, и требует обучения на специально выделенных массивах данных.

Иной подход к упорядочению и порядковой классификации многопризнаковых объектов предлагает методология вербального анализа решений, который ориентирован на слабо структурируемые проблемы выбора, описываемые качественными и количественными признаками. Отличительные особенности вербального анализа решений: получение информации от ЛПР и экспертов на привычном для их профессиональной деятельности языке; использование нечисловой информации без каких-либо ее преобразований в числовую на всех этапах анализа и решения задачи выбора; проверка согласованности суждений ЛПР и устранение выявленных противоречий; возможность объяснения полученного решения [3, 10]. Оценка и сравнение могут проводиться как для всех гипотетически возможных, так и для конкретных альтернатив. Для упорядочения вариантов в семействе методов ЗАПРОС используется специальная так называемая единая порядковая шкала, которая конструируется из комбинаций оценок на шкалах критериев. Методы ОРКЛАСС, СТЕПКЛАСС и ЦИКЛ позволяют строить полную и непротиворечивую классификацию многокритериальных альтернатив.

2. Групповой вербальный анализ решений

Групповой вербальный анализ решений – новый методологический подход в теории многокритериального принятия решений. Объектами группового вербального анализа служат слабо структурируемые проблемы выбора, где предпочтения нескольких ЛПР могут быть несогласованными, а варианты описываются многими разнородными количественными и качественными признаками и могут существовать в нескольких версиях (экземплярах).

Сформулируем основные задачи группового многокритериального выбора. Задана совокупность объектов (вариантов, альтернатив) $A = \{A_1, \dots, A_n\}$, каждый из которых оценивается k экспертами по m критериям Q_1, \dots, Q_m , характеризующим особенности объекта. Каждый критерий (содержательный признак) Q_s , $s=1, \dots, m$ имеет свою собственную порядковую или номинальную шкалу $X_s = \{x_s^1, \dots, x_s^{h_s}\}$, числовую, символьную или вербальную. Предполагается, что разные критерии могут иметь различную относительную важность $w_s > 0$. Будем также считать, что каждый объект оценивается всеми k экспертами по всем m критериям, и что оценки экспертов независимы. Таким образом, многопризнаковый объект A_i , $i=1, \dots, n$ присутствует в k версиях (экземплярах) $A_i^{(f)}$, $f=1, \dots, k$. Основываясь на предпочтениях ЛПР, требуется упорядочить объекты, например, от лучших к худшим.

В задаче классификации с несколькими учителями каждый эксперт наряду с оценкой относит каждый объект A_i , а точнее его версию $A_i^{(f)}$, к одному из классов C_1, \dots, C_g , которые могут быть упорядочены по предпочтительности, либо нет. Принадлежность объекта A_i к классу C_l , $l=1, \dots, g$ описывается сортирующим атрибутом R со шкалой значений $R = \{r_1, \dots, r_g\}$, который можно считать еще одним качественным признаком. Таким образом, имеется k версий (экземпляров) каждого объекта A_i , различающихся оценками по критериям Q_1, \dots, Q_m , и k индивидуальных экспертных правил сортировки этого объекта

IF (условия), THEN (решение), (1)

обычно не согласованных между собой. Здесь терм «условия» включает совокупность значений признаков $x_{i1}^{e_1(f)}, \dots, x_{im}^{e_m(f)}$; $r_{il}^{(f)}$, $x_{is}^{e_s(f)} \in X_s$, $r_{il}^{(f)} \in R$; терм «решение» указывает на принадлежность версии $A_i^{(f)}$ объекта A_i к классу C_l . Требуется найти одно или несколько достаточно простых обобщенных правил групповой классификации вида (1), которые в наибольшей степени совпадают с несогласованными индивидуальными правилами экспертной классификации объектов, позволяют отнести объекты к заданным классам, не отвергая возможную противоречивость оценок, и выявить противоречиво классифицированные объекты.

Многопризнаковый объект A_i обычно принято представлять в пространстве $X = X_1 \times \dots \times X_m$ как вектор или кортеж $x_i = (x_{i1}^{e_1}, \dots, x_{im}^{e_m})$, составленный из количественных (числовых) или качественных (символьных, вербальных) значений признаков [3, 6, 9-11]. Качественные переменные нередко и не всегда обоснованно трансформируются в числовые, например, используя лексикографические шкалы или функции нечеткой принадлежности [12, 19]. Ситуация существенно образом усложняется, если одному и тому же объекту A_i будет соответствовать не один, а несколько m -мерных кортежей. В таких случаях объект A_i представляется в пространстве X группой, состоящей из k кортежей $\{x_i^{(1)}, \dots, x_i^{(k)}\}$ вида $x_i^{(f)} = (x_{i1}^{e_1(f)}, \dots, x_{im}^{e_m(f)})$, $f=1, \dots, k$, которая должна рассматриваться как единое целое. Значения признаков могут быть похожими, различающимися и даже противоречивыми, что может приводить к несравнимости m -мерных кортежей $x_i^{(f)} = (x_{i1}^{e_1(f)}, \dots, x_{im}^{e_m(f)})$, характеризующих один и тот же объект A_i . Более того, если разные объекты или разные версии одного и того же объекта имеют одинаковые значения признаков, но при их классификации они отнесены экспертами к различным классам, то в пространстве X невозможно даже построить разделяющие классы гиперповерхности. Поэтому анализ совокупностей таких многомерных объектов в пространстве X достаточно труден. Эти трудности можно преодолеть, если воспользоваться иным представлением многопризнаковых объектов.

Введем вместо прямого произведения m шкал признаков $X = X_1 \times \dots \times X_m$ множество $X = X_1 \cup \dots \cup X_m$ – обобщенную шкалу признаков, состоящую из m групп признаков, и представим объект в таком символическом виде:

$$A_i = \{k_{A_i}(x_1^1) \circ x_1^1, \dots, k_{A_i}(x_1^{h_1}) \circ x_1^{h_1}, \dots, k_{A_i}(x_m^1) \circ x_m^1, \dots, k_{A_i}(x_m^{h_m}) \circ x_m^{h_m}\}, \quad (2)$$

где число $k_{A_i}(x_s^{e_s})$ указывает, сколько раз признак $x_s^{e_s} \in X_s$ встречается в описании объекта A_i , а знак \circ обозначает кратность признака $x_s^{e_s}$. Множество X характеризует свойства совокупности объектов $A = \{A_1, \dots, A_n\}$. Такая запись объекта A_i представляет его как



мультимножество A_i , порожденное множеством X , которое задано функцией кратности $k_A: X \rightarrow Z_+ = \{0, 1, 2, 3, \dots\}$. Мультимножество A_i конечно, когда конечны все $k_{A_i}(x_s^{e_s})$. Мультимножества A и B равны ($A=B$), если $k_A(x_s^{e_s}) = k_B(x_s^{e_s})$. Мультимножество B включено в мультимножество A ($B \subseteq A$), если $k_B(x_s^{e_s}) \leq k_A(x_s^{e_s}), \forall x_s^{e_s} \in X$ [4].

Введены следующие операции над мультимножествами: объединение $A \cup B$, пересечение $A \cap B$, сложение $A + B$, вычитание $A - B$, симметрическая разность $A \Delta B$, умножение на число $b \cdot A, b_i > 0$, умножение $A \cdot B, n$ -ая степень A^n , прямое произведение $A \times B$, прямая n -ая степень $(\times A)^n$. Новые типы операций над мультимножествами открывают новые возможности для агрегирования многопризнаковых объектов. Например, группа X_a объектов может быть получена как сумма $X_a = \sum_i A_i$, объединение $X_a = \cup_i A_i$ или пересечение $X_a = \cap_i A_i$ мультимножеств A_i , описывающих объекты A_i , либо как линейная комбинация мультимножеств вида $X_a = \sum_i b_i \cdot A_i, X_a = \cup_i b_i \cdot A_i$ или $X_a = \cap_i b_i \cdot A_i$.

На семействе мультимножеств $A = \{A_1, \dots, A_n\}$ можно определить новые классы метрических пространств мультимножеств (A, d) , которые задаются метриками:

$$d_{1p}(A, B) = [m(A \Delta B)]^{1/p}; \quad d_{2p}(A, B) = [m(A \Delta B) / m(Z)]^{1/p}; \quad d_{3p}(A, B) = [m(A \Delta B) / m(A \cup B)]^{1/p}, \quad (3)$$

где p – целое число, m – мера мультимножества, заданная на алгебре мультимножеств $L(Z)$, Z – максимальное мультимножество с $k_Z(x) = \max_{A \in A} k_A(x)$. Метрика $d_{1p}(A, B)$ называется основной и является метрикой типа Хемминга, традиционно используемой во многих приложениях. Метрика $d_{2p}(A, B)$ называется полностью усредненной и характеризует различие между двумя мультимножествами A и B , отнесенное к расстоянию, максимально возможному в исходном пространстве. Метрика $d_{3p}(A, B)$ называется локально усредненной и задает различие, отнесенное к максимально возможной «общей части» $A \cup B$ только этих двух мультимножеств в исходном пространстве. При $p=1$ в случае обычных множеств $d_{11}(A, B) = m(A \Delta B)$ называют расстоянием Фреше, а $d_{31}(A, B) = m(A \Delta B) / m(A \cup B)$ – расстоянием Штейнхауса [1].

Меру мультимножества можно определить различными способами, например, как линейную комбинацию функций кратности: $m(A) = \sum_s w_s k_A(x_s^{e_s}), w_i > 0$. В этом случае, например, основная метрика имеет при $p=1$ следующий вид

$$d_{11}(A, B) = m(A \Delta B) = \sum_{s=1}^m w_s \sum_{e_s=1}^{h_s} |k_A(x_s^{e_s}) - k_B(x_s^{e_s})|.$$

Различные свойства мультимножеств и метрических пространств мультимножеств рассматриваются и обсуждаются в [4].

3. Групповое упорядочение многопризнаковых объектов

Опишем метод группового упорядочения многопризнаковых объектов АРАМИС (Агрегирование и Ранжирование Альтернатив относительно Многопризнаковых Идеальных Ситуаций). Метод предназначен для ранжирования объектов, оценённых несколькими экспертами по многим качественным критериям, без построения индивидуальных ранжировок объектов [6].

Будем считать многопризнаковые объекты точками некоторого метрического пространства мультимножеств (A, d) , например, с основной метрикой d_{11} . В пространстве (A, d) можно выделить два объекта (возможно, гипотетических) – абсолютно лучший и абсолютно худший, которые имеют соответственно самые лучшие и самые худшие оценки по всем критериям. Этим объектам соответствуют мультимножества

$$A_{\max} = \{k \circ x_1^1, 0, \dots, 0, k \circ x_2^1, 0, \dots, 0, \dots, k \circ x_m^1, 0, \dots, 0\},$$

$$A_{\min} = \{0, \dots, 0, k \circ x_1^{h_1}, 0, \dots, 0, k \circ x_2^{h_2}, \dots, 0, \dots, 0, k \circ x_m^{h_m}\},$$

которые принято называть идеальным и антиидеальным решениями. Таким образом, задача упорядочения многопризнаковых объектов сводится к задаче упорядочения мультимножеств.

Упорядочим все объекты по величине их расстояния от идеального решения A_{\max} . Если $d_{11}(A_{\max}, A_i) < d_{11}(A_{\max}, A_j)$, то объект A_i предпочтительнее объекта A_j ($A_i \succ A_j$). Если для некоторых объектов $d_{11}(A_{\max}, A_i) = d_{11}(A_{\max}, A_j)$, то объекты A_i и A_j будут или эк-

вивалентными, или несравнимыми. Тем самым получаемое ранжирование объектов окажется нестрогим.

Так как каждый объект A_i оценивается k экспертами по всем m критериям, то выражение для расстояния от идеального решения A_{\max} до мультимножества A_i можно записать как

$$d_{11}(A_{\max}, A_i) = 2 \sum_{s=1}^m w_s [k - k_{A_i}(x_s^1)].$$

Условие сравнения многопризнаковых объектов приобретает тогда следующую форму: объект A_i предпочтительнее объекта A_j ($A_i \succ A_j$), если

$$\sum_{s=1}^m w_s k_{A_i}(x_s^1) > \sum_{s=1}^m w_s k_{A_j}(x_s^1).$$

Тем самым правило упорядочения многопризнаковых объектов сводится к сравнению взвешенных сумм $S_{A_i^1} = \sum_s w_s k_{A_i}(x_s^1)$ первых (наилучших) оценок объектов по всем критериям Q_s . Лучшим будет тот объект A_i , у которого эта сумма $S_{A_i^1}$ будет больше. Если найдутся группы эквивалентных или несравнимых объектов A_{i_1}, \dots, A_{i_t} , имеющих одинаковые суммы $S_{A_i^1}$, то нужно вычислить для каждого объекта A_{i_r} , $r=1, \dots, t$ в соответствующей группе взвешенную сумму $S_{A_{i_r}^2} = \sum_s w_s k_{A_{i_r}}(x_s^2)$ всех вторых оценок по всем критериям Q_s и упорядочить объекты внутри каждой группы от лучшего к худшему по величинам $S_{A_{i_r}^2}$ сумм вторых оценок. Если останутся подгруппы эквивалентных или несравнимых объектов $A_{i_{r_1}}, \dots, A_{i_{r_v}}$, имеющих одинаковые суммы $S_{A_{i_r}^2}$, то вычислить для каждого объекта $A_{i_{r_p}}$ в соответствующей подгруппе взвешенную сумму $S_{A_{i_{r_p}}^3} = \sum_s w_s k_{A_{i_{r_p}}}(x_s^3)$ всех третьих оценок по всем критериям Q_s и упорядочить объекты внутри каждой подгруппы от лучшего к худшему по величинам сумм $S_{A_{i_{r_p}}^3}$ третьих оценок. Процедура продолжается до полного упорядочения всех объектов из совокупности $A = \{A_1, \dots, A_n\}$. Если число h_s значений оценок $x_s^{e_s}$ у некоторых критериев Q_s окажется меньше требуемого на данном b -ом шаге алгоритма, то следует считать $k_{A_{i_r \dots p}}(x_s^b) = 0$. В случае, когда все критерии считаются одинаково важными, все коэффициенты w_s полагаются равными 1.

Аналогично строится процедура упорядочения многопризнаковых объектов A_i по отношению к антиидеальному решению A_{\min} , считая, что объект A_i лучше объекта A_j ($A_i \succ A_j$), если мультимножество A_i находится дальше от антиидеального решения A_{\min} , то есть если выполняется условие $d_{11}(A_{\min}, A_i) > d_{11}(A_{\min}, A_j)$. Подчеркнем, что упорядочение совокупности многопризнаковых объектов $A = \{A_1, \dots, A_n\}$ по отношению к антиидеальному решению может не совпадать с упорядочением по отношению к идеальному решению.

4. Согласованная групповая классификация многопризнаковых объектов

Изложим теперь метод групповой классификации многопризнаковых объектов МАСКА (МногоАспектная Согласованная Классификация Альтернатив). Метод предназначен для построения согласованных правил классификации объектов, оценённых несколькими экспертами по многим качественным критериям, которые агрегируют противоречивые индивидуальные правила сортировки объектов и выявляют противоречиво классифицированные объекты [14].

Введем иную гипершкалу признаков – множество $X' = X_1 \cup \dots \cup X_m \cup R$, которое объединяет подмножества критериальных оценок X_s и подмножество сортирующих признаков R . Сопоставим многопризнаковому объекту A_i мультимножество:

$$A_i = \{k_{A_i}(x_1^1) \circ x_1^1, \dots, k_{A_i}(x_1^{h_1}) \circ x_1^{h_1}, \dots, k_{A_i}(x_m^1) \circ x_m^1, \dots, k_{A_i}(x_m^{h_m}) \circ x_m^{h_m}, k_{A_i}(r_1) \circ r_1, \dots, k_{A_i}(r_g) \circ r_g\}, \quad (4)$$

порожденное множеством X' . Значения функции кратности $k_{A_i}(x_s^{e_s})$ и $k_{A_i}(r_i)$ мультимножества A_i равны соответственно числам экспертной оценки $x_s^{e_s}$ объекта A_i по критерию Q_s и заключения r_i .

Мультимножество (4) можно интерпретировать как коллективное правило классификации вида (1), объединяющее k индивидуальных правил сортировки объекта A_i . В этом правиле терм (условия) включает различные комбинации оценок по



критериям $x_s^{e_s}$ и признаков r_l , а терм «решение» обозначает принадлежность объекта A_i к некоторому классу C_l . Например, объект A_i относится к классу C_l согласно одному из следующих правил «большинства голосов»: если $k_{A_i}(r_l) > k_{A_i}(r_p)$ для всех $p \neq l$, или если $k_{A_i}(r_l) > \sum_{p \neq l} k_{A_i}(r_p)$. Очевидно, что оценки и заключения многих экспертов могут быть схожими, различающимися и противоречивыми. Эти расхождения не являются случайными ошибками, а выражают субъективность предпочтений экспертов.

Для упрощения задачи предположим, что совокупность объектов $A = \{A_1, \dots, A_n\}$ нужно разделить только на два класса C_a (более предпочтительные) и C_b (менее предпочтительные). Разбиение именно на два класса не является принципиальным ограничением. Если необходимо рассортировать объекты на большее число классов, можно сначала разбить совокупность объектов на две группы, затем одну из них или обе группы – на подгруппы, и так далее. Например, если требуется выделить несколько групп конкурсных проектов, то сначала заявки можно разделить на принятые и отложенные, затем принятые заявки – на безусловно и условно поддержанные, а отложенные заявки – на возможно поддержанные и неподдержанные, и так далее.

Сопоставим каждому классу объектов мультимножество, которое образуем, суммируя мультимножества (4), которые представляют многопризнаковые объекты, принадлежащие к этому классу по принятому правилу большинства голосов. Тогда классу C_t , $t = a, b$ соответствует мультимножество:

$$Y_t = \{k_{Y_t}(x_1^{e_1}) \circ x_1^{e_1}, \dots, k_{Y_t}(x_1^{h_1}) \circ x_1^{h_1}, \dots, k_{Y_t}(x_m^{e_m}) \circ x_m^{e_m}, \dots, k_{Y_t}(x_m^{h_m}) \circ x_m^{h_m}, k_{Y_t}(r_1) \circ r_1, \dots, k_{Y_t}(r_g) \circ r_g\}, \quad (5)$$

где $k_{Y_t}(x_s^{e_s}) = \sum_{i \in I_t} k_{A_i}(x_s^{e_s})$, $k_{Y_t}(r_l) = \sum_{i \in I_t} k_{A_i}(r_l)$, $I_a \cup I_b = \{1, \dots, n\}$, $I_a \cap I_b = \emptyset$. Мультимножество (5) можно считать групповым правилом классификации объектов вида (1), объединяющим индивидуальные правила всех экспертов по сортировке совокупности объектов $A = \{A_1, \dots, A_n\}$ в класс C_t .

Сконструируем новые объекты, представленные мультимножествами

$$R_l = \{k_{R_l}(y_a) \circ y_a, k_{R_l}(y_b) \circ y_b\}, \quad Q_j = \{k_{Q_j}(y_a) \circ y_a, k_{Q_j}(y_b) \circ y_b\}$$

над множеством новых признаков $Y = \{y_a, y_b\}$, элементы которого характеризуют принадлежность объектов к классам C_a и C_b . Здесь $k_{R_l}(y_l) = k_{Y_t}(r_l)$, $l = 1, \dots, g$, $k_{Q_j}(y_l) = k_{Y_t}(x_j)$, $j = 1, \dots, h$, $h = h_1 + \dots + h_m$. Назовем мультимножества R_l категориальными, а мультимножества Q_j – содержательными.

Нетрудно убедиться, что согласно правилу большинства голосов категориальные мультимножества формируют два мультимножества $R_a = \sum_{l \in L_a} R_l$ и $R_b = \sum_{l \in L_b} R_l$, где $L_a \cup L_b = \{1, \dots, g\}$, $L_a \cap L_b = \emptyset$, расстояние $d^* = d(R_a, R_b)$ между которыми в метрическом пространстве (B, d) новых мультимножеств будет наибольшим расстоянием между объектами, входящими в разные классы. $B = \{R_1, \dots, R_g, Q_1, \dots, Q_h\}$, d – одна из метрик (3).

Категориальные мультимножества R_a и R_b определяют декомпозицию рассматриваемой совокупности объектов A на классы C_a и C_b , лучшую из всех возможных для заданных индивидуальных экспертных правил сортировки:

$$\begin{aligned} & \text{IF } \langle (\sum_{l \in L_a} k_{A_i}(r_l) > \sum_{l \in L_b} k_{A_i}(r_l)) \rangle, \text{ THEN } \langle \text{Объект } A_i \in C_a \rangle, \\ & \text{IF } \langle (\sum_{l \in L_a} k_{A_i}(r_l) < \sum_{l \in L_b} k_{A_i}(r_l)) \rangle, \text{ THEN } \langle \text{Объект } A_i \in C_b \rangle. \end{aligned} \quad (6)$$

При отсутствии противоречий между индивидуальными правилами экспертов по предварительной сортировке объектов, расстояние d^* будет равно соответственно $d_{11}^* = kn$, $d_{12}^* = 1/(h+g)$, $d_{13}^* = 1$.

Основная идея нахождения простого обобщенного правила групповой классификации многопризнаковых объектов, наилучшим образом агрегирующего несогласованные индивидуальные правила сортировки многих экспертов, формулируется так. Для каждой s -ой группы признаков Q_s нужно построить содержательные мультимножества Q_{sa}^* и Q_{sb}^* , которые будут определять s -ую лучшую декомпозицию совокупности объектов A на классы C_a и C_b .

По аналогии с категориальными мультимножествами R_a и R_b можно сформировать содержательные мультимножества Q_{sa} и Q_{sb} , которые образуются как суммы мультимножеств: $Q_{sa} = \sum_{j \in J_{sa}} Q_j$ и $Q_{sb} = \sum_{j \in J_{sb}} Q_j$, где $J_{sa} \cup J_{sb} = \{1, \dots, h_s\}$, $J_{sa} \cap J_{sb} = \emptyset$, и задают

некоторую декомпозицию объектов на классы C_a и C_b . Каждое мультимножество Q_{st} соответствует некоторой подгруппе значений признаков $Q_{st} = \bigcup_{j \in J_{st}} Q_j$, $t = a, b$.

Рассматривая содержательные мультимножества как точки в пространстве мультимножеств (B, d) , требуется построить для всех признаков такие пары мультимножеств Q_{sa}^* и Q_{sb}^* , которые находились бы на максимально возможном расстоянии. Мультимножества Q_{sa}^* и Q_{sb}^* являются решением следующей оптимизационной задачи:

$$d(Q_{sa}, Q_{sb}) \rightarrow \max d(Q_{sa}, Q_{sb}) = d(Q_{sa}^*, Q_{sb}^*). \quad (7)$$

Значения признаков $x_j \in Q_{st}^*$, получаемые из решения m оптимизационных задач (7) и характеризующие класс C_t , будем называть классифицирующими признаками для этого класса. Классифицирующие признаки можно упорядочить по величине расстояния $d(Q_{sa}^*, Q_{sb}^*)$ или точности аппроксимации $V_s = d(Q_{sa}^*, Q_{sb}^*) / d(R_a, R_b)$. Классифицирующий признак $x_j \in Q_{st}^*$, обеспечивающий приемлемую точность аппроксимации $V_s \geq V_0$, включается в обобщенное решающее правило для групповой сортировки многопризнаковых объектов. Значение точности аппроксимации V_s характеризует относительную значимость s -ой группы признаков Q_s в обобщенном правиле классификации.

Разные комбинации классифицирующих признаков дают следующие обобщенные решающие правила для групповой сортировки объектов на классы C_a и C_b :

$$\text{IF } \langle (x_j \in Q_{ua}^*) \text{ AND } (x_j \in Q_{va}^*) \text{ AND } (x_j \in Q_{wa}^*) \text{ AND } \dots \rangle, \text{ THEN } \langle \text{Объект } A_i \in C_a \rangle, \quad (8)$$

$$\text{IF } \langle (x_j \in Q_{ub}^*) \text{ AND } (x_j \in Q_{vb}^*) \text{ AND } (x_j \in Q_{wb}^*) \text{ AND } \dots \rangle, \text{ THEN } \langle \text{Объект } A_i \in C_b \rangle. \quad (9)$$

При одновременном учете экспертных правил сортировки (6) и обобщенного решающего правила (8) или (9) среди отобранных окажутся как правильно, так и неправильно классифицированные объекты. Необходимо найти такие классифицирующие признаки, которые обеспечат максимальные числа N_a и N_b правильно классифицированных объектов и минимальные числа N_{ac} и N_{bc} неправильно классифицированных объектов.

Начнем с обобщенного решающего правила (8). Найдем сначала критерий Q_u^* , обеспечивающий максимальную разность $N_a - N_{ac}$ между числами правильно и неправильно классифицированных объектов. Затем рассмотрим все комбинации критерия Q_u^* с остальными критериями и выберем пару критериев Q_u^* , Q_v^* , обеспечивающих максимальную разность $N_a - N_{ac}$. Далее, рассмотрев все комбинации пары критериев Q_u^* , Q_v^* с остальными критериями, выберем тройку Q_u^* , Q_v^* , Q_w^* , затем четверку и т.д. критериев, обеспечивающих максимальную разность $N_a - N_{ac}$. Последовательно шаг за шагом уменьшая разность $N_a - N_{ac}$ между числами правильно и неправильно классифицированных объектов, найдем критерии, которые включаются в согласованное решающее правило для построения уточненного класса $C_a \setminus C_{ac}$ безусловно предпочтительных объектов:

$$\text{IF } \langle (\sum_{x \in Q_{ua}^*} k_{Ai}(x_j) > \sum_{x \in Q_{ub}^*} k_{Ai}(x_j)) \text{ AND } (\sum_{x \in Q_{va}^*} k_{Ai}(x_j) > \sum_{x \in Q_{vb}^*} k_{Ai}(x_j)) \text{ AND } \dots \\ \text{AND } (\sum_{r \in R_a} k_{Ai}(r) > \sum_{r \in R_b} k_{Ai}(r)) \rangle, \text{ THEN } \langle \text{Объект } A_i \in C_a \setminus C_{ac} \rangle, \quad (10)$$

Аналогичным образом получается согласованное решающее правило для построения уточненного класса $C_b \setminus C_{bc}$ безусловно неpreferируемых объектов:

$$\text{IF } \langle (\sum_{x \in Q_{ua}^*} k_{Ai}(x_j) < \sum_{x \in Q_{ub}^*} k_{Ai}(x_j)) \text{ AND } (\sum_{x \in Q_{va}^*} k_{Ai}(x_j) < \sum_{x \in Q_{vb}^*} k_{Ai}(x_j)) \text{ AND } \dots \\ \text{AND } (\sum_{r \in R_a} k_{Ai}(r) < \sum_{r \in R_b} k_{Ai}(r)) \rangle, \text{ THEN } \langle \text{Объект } A_i \in C_b \setminus C_{bc} \rangle, \quad (11)$$

с помощью которого минимизируется разность $N_b - N_{bc}$ между числами правильно и неправильно классифицированных объектов.

Согласованные групповые правила (10) и (11) наилучшим образом аппроксимируют семейство индивидуальных правил сортировки многих экспертов и, вообще говоря, различны.

Одновременно формируется класс $C_c = C_{ac} \cup C_{bc}$ противоречиво классифицированных объектов, которые удовлетворяют следующему решающему правилу:

$$\text{IF } \langle [(\sum_{x \in Q_{ua}^*} k_{Ai}(x_j) > \sum_{x \in Q_{ub}^*} k_{Ai}(x_j)) \text{ AND } (\sum_{x \in Q_{va}^*} k_{Ai}(x_j) > \sum_{x \in Q_{vb}^*} k_{Ai}(x_j)) \text{ AND } \dots$$



$$\text{AND } (\sum_{r \in R_a} k_{Ai}(r) < \sum_{r \in R_b} k_{Ai}(r)) \quad \text{OR} \quad [(\sum_{x \in Q_{ua}} k_{Ai}(x) < \sum_{x \in Q_{ub}} k_{Ai}(x)) \quad \text{AND} \quad (\sum_{x \in Q_{va}} k_{Ai}(x) < \sum_{x \in Q_{vb}} k_{Ai}(x))] \quad \text{AND} \dots \text{AND} \quad (\sum_{r \in R_a} k_{Ai}(r) > \sum_{r \in R_b} k_{Ai}(r))] ,$$

$$\text{THEN } \langle \text{Объект } A_i \in C_c \rangle. \quad (12)$$

Эти объекты нуждаются в дополнительном анализе. Правило (12) помогает обнаружить возможные противоречия в индивидуальных экспертных правилах сортировки.

Заключение

Задачи групповой классификации и упорядочения объектов, которые характеризуются многими разнородными признаками, причем каждый из объектов может существовать в нескольких различающихся версиях, являются достаточно трудными для решения и до настоящего момента слабо изучены. Трудности имеют и содержательные основания (например, необходимость рассмотрения нескольких версий объектов), и формальные причины (например, некорректность процедуры «усреднения» качественных признаков). Использование аппарата теории мультимножеств при решении практических задач принятия решений, анализа данных, распознавания образов и в других областях позволяет ввести новые виды операций, создать новые методы обработки разнородной информации, не приводящие к потере или искажению данных.

Разработанные подходы были использованы для построения рейтинга российских компаний, работающих в секторе информационно-коммуникационных технологий [2], для оценки кредитоспособности держателей кредитных карт [5], при анализе конкурсов целевых фундаментальных исследований, выполняемых по грантам Российского фонда фундаментальных исследований [7]. Решающие правила (8)-(12) для групповой сортировки объектов можно легко записать на естественном языке, используя формулировки вербальных значений классифицирующих признаков. Например, согласованное групповое правило для конкурсного отбора проектов целевых фундаментальных исследований выглядит следующим образом [7]:

«Проект безусловно поддерживается, если он имеет исключительно высокую или значительную научную значимость; массовый или междисциплинарный масштабы применимости результатов исследований; предлагаемые в проекте исследования завершаются лабораторным образцом или ключевыми элементами разработки».

В работе изложены методы классификации и упорядочения многопризнаковых объектов, основанные на их представлении с помощью мультимножеств, которые применимы и в случае несогласованности и противоречивости описаний объектов. Использование мультимножеств позволяет значительно расширить круг рассматриваемых проблем и решать новые разнообразные задачи классификации, сортировки, ранжирования многопризнаковых объектов.

Литература

1. Деза М.М., Лоран М. Геометрия разрезов и метрик. – М.: МЦНМО, 2001.
2. Кто в России самый интеллектуальный? Рейтинг ведущих российских разработчиков высоких технологий. // Компания. – 2000. – №47(143) – С.38-39.
3. Ларичев О.И. Вербальный анализ решений. – М.: Наука, 2006.
4. Петровский А.Б. Пространства множеств и мультимножеств. – М: Едиториал УРСС, 2003.
5. Петровский А.Б. Модель оценки кредитоспособности владельцев кредитных карт по противоречивым данным. // Искусственный интеллект. – 2004. – №2. – С.155-161.
6. Петровский А.Б. Теория принятия решений. – М.: Издательский центр «Академия», 2009.
7. Петровский А.Б., Тихонов И.П., Балышев А.В., Комарова Н.А. Построение согласованных групповых правил для конкурсного отбора научных проектов // Труды третьей межд. конф. "Системный анализ и информационные технологии". – М.: ПолиПринтСервис, 2009. – С.337-348.
8. Т.Саати Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1993.



9. Doumpos M., Zopounidis C. Multicriteria Decision Aid Classification Methods. – Dordrecht: Kluwer Academic Publishers, 2002.
10. Furems E. Knowledge-based multi-attribute classification problems structuring. // Ruan D., Montero J., Lu J., Martinez L., D'hondt P., Kerre E. (eds.) Computational Intelligence in Decision And Control. – Singapore: World Scientific Publisher, 2008. – P.465-470.
11. Greco S., Matarazzo B., Slowinski R. Rough sets methodology for sorting problems in presence of multiple attributes and criteria. // European Journal of Operational Research. – 2002. – Vol.138. – N2. – P.247-259.
12. Hwang C.L., Lin M.J. Group Decision Making under Multiple Criteria. – Berlin: Springer-Verlag, 1987.
13. Koksalan M., Ulu C. An interactive approach for placing alternatives in preference classes. // European Journal of Operational Research. – 2003. – Vol.144. – N2. – P.429-439.
14. Petrovsky A.B. Multi-attribute sorting of qualitative objects in multiset spaces. // Multiple Criteria Decision Making in the New Millenium. – Berlin: Springer-Verlag, 2001. – P.124-131.
15. Roubens M. Ordinal multiattribute sorting and ordering in the presence of interacting points of view. // Bouyssou D., Jacquet-Lagrange E., Perny P., Slowinski R., Vanderpooten D., Vincke P. (eds.) Aiding Decisions with Multiple Criteria: Essays in Honor of Bernard Roy. – Dordrecht: Kluwer Academic Publishers, 2001. – P.229-246.
16. Roy B. Multicriteria Methodology for Decision Aiding. – Dordrecht: Kluwer Academic Publishers, 1996.
17. Roy B., Bouyssou D. Aide Multicritere a la Decision: Methodes et Cas. – Paris: Economica, 1993.
18. Vincke P. Multicriteria Decision Aid. – Chichester: Wiley, 1992.
19. Zadeh L.A. From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions. // IEEE Transactions on Circuits and Systems. – 1999. – Vol.45. – N1. – P.105-119.

GROUP MULTICRITERIAL DECISION MAKING WITH INCONSISTENT PREFERENCES

A. B. PETROVSKY

*Institute for System
Analysis, Russian
Academy of Sciences*

e-mail: pab@isa.ru

The paper considers a new approach to group multicriterial decision making, where preferences of several decision makers are inconsistent and possibly contradictory, and alternatives are estimated by manifold criteria. The approach is based on the methodology of verbal decision analysis and the theory of multisets. There are examples of the approach application to solve practical problems of classifying and ordering multi-attribute objects.

Key words: group decision making, theory of multisets, classifying, ordering, multi-attribute objects

СИНТЕЗ НЕЧЕТКИХ РЕШАЮЩИХ ПРАВИЛ ДЛЯ МЕДИКО-ЭКОЛОГИЧЕСКИХ ПРИЛОЖЕНИЙ НА ОСНОВЕ АНАЛИЗА СТРУКТУРЫ ДАННЫХ

Н. А. КОРЕНЕВСКИЙ¹⁾

С. А. ФИЛИСТ¹⁾

Г. В. ЧУРСИН²⁾

*¹⁾Курский
государственный
технический
университет*

e-mail:SFilist@gmail.com

*²⁾Курская
государственная
сельскохозяйственная
академия*

В работе предлагается метод синтеза нечетких решающих правил на основе объединения способов агрегации функций принадлежности, формы и параметры которых определяются при использовании данных о структуре признаков и классов, получаемых методами разведочного анализа. Показывается, что такой подход позволяет решать задачи прогнозирования, ранней и дифференциальной диагностики для различных типов медико-экологических задач с приемлемым для практики качеством.

Ключевые слова: нечеткая логика, прогнозирование, ранняя диагностика, дифференциальная диагностика, функция принадлежности, коэффициент уверенности, разведочный анализ.

Многочисленные исследования показывают, что при практическом применении методов теории распознавания образов наилучших результатов удается достичь, если при синтезе соответствующих решающих правил учитывается информация о структуре данных и других особенностях решаемых задач. Для медицинских и экологических приложений к таким особенностям следует относить следующие обстоятельства. Исследуемые структуры классов могут сильно пересекаться в пространстве информативных признаков, особенно если речь идет о прогнозировании и ранней (донозологической) диагностике; классы состояний могут менять свое положение в пространстве признаков; доступные для измерения признаки измеряются в различных шкалах (порядка, наименований, интервалов), имеют различную природу и фиксируются на различных этапах ведения пациентов (данные опросов, осмотров, инструментальных и лабораторных исследований); в реальных условиях не весь требуемый набор информативных признаков может быть зарегистрирован у конкретного обследуемого; как исходные признаки, так и принимаемые решения могут иметь нечеткий характер и т.д.

Анализ литературных данных и результаты собственных исследований позволяют сделать вывод о том, что в этих условиях предпочтение следует отдавать двум подходам, принятым в теории принятий решений: на основе теории нечеткой логики принятия решений [1, 3, 5, 8, 9] и на основе аппарата, обеспечивающего изучение структуры классов с выдвижением гипотез о наилучших классификаторах в ходе вычислительного эксперимента (диалоговые системы распознавания образов) [2, 4, 7]. Каждый из этих подходов обладает определенными достоинствами, но при решении практических задач они используются раздельно, что снижает потенциально достижимые возможности проектируемых классификаторов.

С целью повышения эффективности решения задач прогнозирования и медицинской диагностики заболеваний, в том числе и заболеваний порождаемых вредными экологическими факторами, нами предлагается объединить два этих подхода, реализуя задачи синтеза нечетких решающих правил в три этапа.

На первом этапе производится разведочный анализ, позволяющий изучить геометрическую структуру классов в пространстве информативных признаков. Под структурой имеется в виду взаимоположение объектов различных классов на обучающей выборке.

На втором этапе под известную структуру классов и типы признаков выбираются носители и параметры функций принадлежности. Эти функции решают задачи клас-

сификации по подпространствам и областям исходного пространства признаков. При этом выбор осуществляется с таким расчетом, чтобы при заданной сложности классификатора каждая функция принадлежности на каждом технологическом шаге принятия решений обеспечивала максимально возможную уверенность классификации или прогнозирования.

На третьем этапе функции принадлежности объединяются в частные (по группам однотипных признаков и подпространствам) и финальные решающие правила, обеспечивающие требуемую уверенность в принимаемых решениях.

Здесь функции принадлежности определяются так, как в основополагающей работе [5]. Агрегация функций принадлежности в нечеткие решающие правила может осуществляться с использованием операций, определенных в классической теории нечетких множеств: объединения, пересечения, отрицания, импликации и т.д.

В ряде приложений удобно частные и финальные решающие правила определять через коэффициенты уверенности так, как это предлагается в работе [5] с использованием нечетких правил вида:

$$КУ_{\omega_\ell} = МД_{\omega_\ell} - МНД_{\omega_\ell}; \quad (1)$$

$$МД_{\omega_\ell}(j+1) = МД_{\omega_\ell}(j) + МД_{\omega_\ell}^*(x_i) \cdot [1 - МД_{\omega_\ell}(j)]; \quad (2)$$

$$МНД_{\omega_\ell}(k+1) = МНД_{\omega_\ell}(k) + МНД_{\omega_\ell}^*(y_q) \cdot [1 - МНД_{\omega_\ell}(k)], \quad (3)$$

где $КУ_{\omega_\ell}$ – коэффициент уверенности в гипотезе ω_ℓ ; $МД_{\omega_\ell}$ и $МНД_{\omega_\ell}$ – меры доверия и недоверия к гипотезе ω_ℓ соответственно, j и k – номера итераций в расчетах $МД_{\omega_\ell}$ и $МНД_{\omega_\ell}$ соответственно; $МД_{\omega_\ell}^*(x_i)$ и $МНД_{\omega_\ell}^*(y_q)$ – меры доверия и недоверия к ω_ℓ от одного свидетельства (признака, фактора) x_i и y_q соответственно; i и q – номера свидетельств.

Если выбираемые свидетельства «работают» только на повышение уверенности в ω_ℓ , а используемые функции принадлежности выбираются как меры доверия к ω_ℓ , то агрегацию решающих правил можно проводить, объединяя два направления нечеткой логики принятия решений, развиваемых в работах Заде [8, 9] и Шортлифа [5]. Тогда коэффициент уверенности может быть определен итерационной формулой вида:

$$КУ_{\omega_\ell}(j+1) = КУ_{\omega_\ell}(j) + \mu_{\omega_\ell}(S_p) \cdot [1 - КУ_{\omega_\ell}(j)], \quad (4)$$

где $\mu_{\omega_\ell}(S_p)$ – функция принадлежности к ω_ℓ с носителем по шкале S_p , которая может определяться как по отдельно взятому свидетельству, так и по группе свидетельств, объединяемых по какому-либо закону в обобщающий интегральный показатель, свидетельствующий в пользу гипотезы ω_ℓ .

В общем случае, функции принадлежности могут характеризовать как $МД_{\omega_\ell}^*(x_i)$ так и $МНД_{\omega_\ell}^*(y_q)$, тогда коэффициенты уверенности будут определяться модифицированными формулами (1-3) с заменой $МД_{\omega_\ell}^*(x_i)$ на $\mu_{\omega_\ell}(x_i)$ и $МНД_{\omega_\ell}^*(y_q)$ на $\mu_{\omega_\ell}(y_q)$.

Для проведения разведочного анализа с целью изучения структуры исследуемых классов нами разработан специальный пакет прикладных программ, решающий следующие задачи: выделение характерных точек обучающей выборки (многомерных центров классов и выделяемых объектов, групп наиболее близких и наиболее далеких объектов между парами различных классов, казуистических и артефактных объектов); рас-

чет расстояний между характерными точками и между всеми заданными точками, как внутри своего класса, так и до точек чужого класса; различные методы отображения многомерных данных в двумерные пространства с сохранением выбираемых структурных свойств исследуемых объектов (сохранение близких расстояний, сохранение далеких расстояний, сохранение структур задаваемых ядер и т.д.); построение гистограмм распределений объектов исследуемых классов на координатах признаков (признаковые гистограммы); построение гистограмм распределения объектов исследуемых классов на шкалах, определяемых как меры близости до эталонных многомерных структур (точек, гиперплоскостей, гиперкубов, гиперсфер и т.д.) (дистальные гистограммы); определение исходных координат объектов по выбираемым участкам гистограмм и областям отображающих пространств; определение группировок объектов в многомерном пространстве признаков; определение областей пересечений различных классов в исходном пространстве с описанием структурных особенностей этих областей.

Знание различных характеристик структурных особенностей классов позволяет обоснованно, под структуру классов: выделить подпространства, где с точки зрения пользователя и достигаемых результатов удобно и целесообразно построить частные нечеткие решающие правила; выбрать тип носителя и характеристики частных функций принадлежности; из частных решающих правил построить агрегирующее решающее правило, обеспечивающее требуемое качество классификации [1,6].

Под различные типы структур классов разрабатываются рекомендации по выбору типов носителей и параметров функций принадлежности, обеспечивающих высокое качество классификации при хорошей интерпретируемости получаемых результатов и небольшой вычислительной сложности.

Рисунок 1 иллюстрирует вариант выбора носителя функций принадлежности для случая линейно-разделимых классов ω_1 и ω_2 в двумерном пространстве признаков $\{x_1, x_2\}$. Из этого рисунка хорошо видно, что признаковые гистограммы классов $h_{\omega_1}(x_1), h_{\omega_2}(x_1), h_{\omega_1}(x_2), h_{\omega_2}(x_2)$, сильно перекрываются. Также сильно перекрываются и частные функции принадлежности.

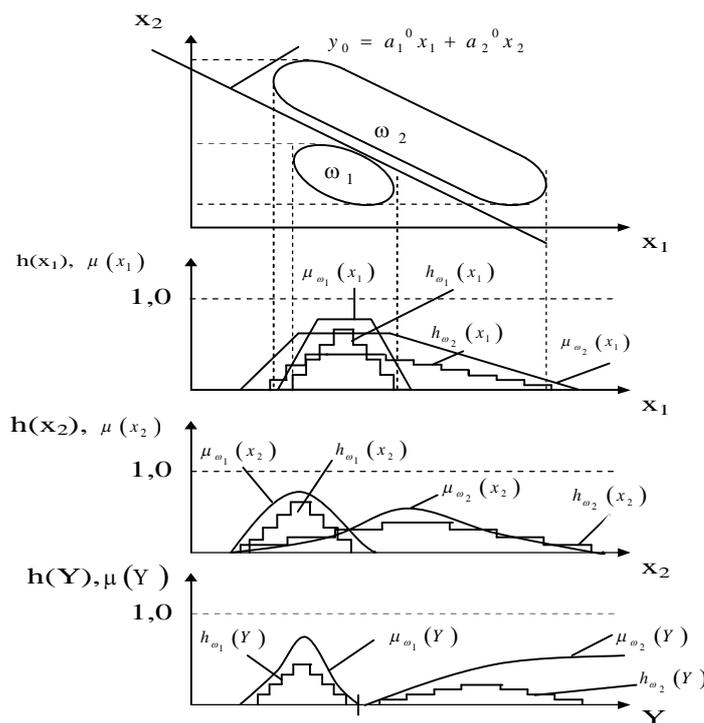


Рис. 1. Построение признаковых и дистальных гистограмм для линейно-разделимых классов

Анализ признаков функций принадлежности показывает, что по ним нельзя построить надежные решающие правила для классификации ω_1 и ω_2 . С другой стороны, имея по данным разведочного анализа информацию о линейной разделимости классов ω_1 и ω_2 , легко получить разделяющую линию $y_0 = a_1^0 x_1 + a_2^0 x_2$ и семейство взаимно-параллельных линий вида $Y = a_1 x_1 + a_2 x_2$, использование которых позволяет, с одной стороны, подтвердить линейную разделимость (дистальные гистограммы $h_{\omega_1}(Y)$, $h_{\omega_2}(Y)$), а с другой стороны, построить непересекающиеся функции принадлежности $\mu_{\omega_1}(Y)$, $\mu_{\omega_2}(Y)$, обеспечивающие построение надежного классификационного правила типа

$$\Omega = \max \{ \mu_{\omega_1}^j(Y), \mu_{\omega_2}^j(Y) \}, \quad \Omega = \{ \omega_1, \omega_2 \}. \quad (5)$$

В соответствии с этим правилом объект с номером j относится к тому из классов ω_1 или ω_2 , для которого функция принадлежности имеет большее значение.

Рисунок 2 иллюстрирует вариант построения решающего правила для случая пересекающихся классов ω_1 и ω_2 через функции принадлежности $\mu_{\omega_1}(Y)$, $\mu_{\omega_2}(Y)$, с носителем на шкале $Y = a_1 x_1 + a_2 x_2$.

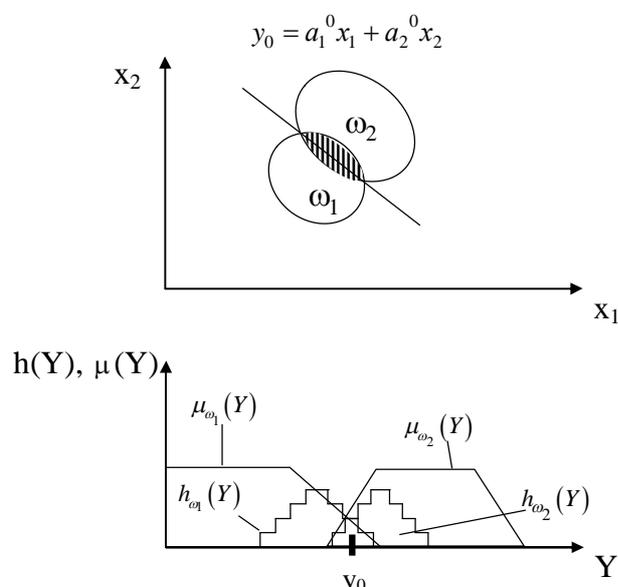


Рис. 2. Вариант выбора функций принадлежности для пересекающихся классов

В многомерном случае если для двух классов ω_1 и ω_2 в ходе разведочного анализа по исследуемым подпространствам или всему пространству информативных признаков определена целесообразность использования линейной или кусочно-линейной разделяющей поверхности, то в качестве носителя для соответствующих функций принадлежности удобно использовать выражение вида:

$$Y = \sum_{i=1}^n a_{ik} x_{ik}, \quad (6)$$

где x_{ik} – признак с номером i в подпространстве с номером k ($i=1, \dots, n$; $k=1, \dots, K$); a_{ik} – настраиваемые параметры, ориентирующие гиперплоскость (6) по критерию, минимизирующему ошибку классификации; Y – переменная величина, пропорциональная величине расстояния от начала координат до гиперплоскости (6).



Если зафиксировать положение гиперплоскости (6) в подпространстве (пространстве признаков), определив для минимального количества ошибок классификации $Y=y_0$, то по величине Y относительно y_0 можно судить о расстоянии от объекта классификации до линейной разделяющей гиперповерхности, заданной выражением

$$y_0 = \sum_{i=1}^n a_{ik}^0 x_{ik} .$$

В этом случае в качестве ориентира для выбора вида и параметров функций принадлежности удобно использовать дистальные гистограммы $h_{\omega_1}(Y)$ и $h_{\omega_2}(Y)$ распределения объектов классов ω_1 и ω_2 по шкале Y . По графикам гистограмм эксперты, решающие задачу определения функций принадлежности, могут выбрать: точки перехода на оси Y , в которых $\mu_{\omega_\ell}(Y) = 0$ и где $\mu_{\omega_\ell}(Y) > 0$ (переход между интервалом отказа от гипотезы ω_ℓ и интервалом, где гипотеза ω_ℓ принимается); участки (точки) где $\mu_{\omega_\ell}(Y)$ принимает максимальные значения; точки, где $\mu_{\omega_\ell}(Y)$ принимают половинные значения от своего максимума; форму наклонных участков и т.д.

При установлении факта пересечения объектов классов ω_1 и ω_2 в исходном пространстве строятся пересекающиеся участки функций принадлежности $\mu_{\omega_1}(Y)$, $\mu_{\omega_2}(Y)$ с учетом плотностей объектов в исходном пространстве, порождающих зону пересечения гистограмм $h_{\omega_1}(Y)$, $h_{\omega_2}(Y)$, и с учетом “запасов” на возможные прогнозируемые ошибки классификации в реальных условиях.

При выборе формы и параметров функций принадлежности следует иметь в виду, что если гистограммы отражают частоту появления объектов в исследуемых классах, то функции принадлежности отражают экспертную уверенность в диагностике.

Если методы разведочного анализа показывают, что интервал пересечения гистограмм $h_{\omega_1}(Y)$, $h_{\omega_2}(Y)$ образован объектами, не имеющим пересечения классов в исходном пространстве, то может быть рассмотрена возможность использования для классификации кусочно-линейных разделяющих поверхностей. В этом варианте в начале функции $\mu_{\omega_1}(Y)$, $\mu_{\omega_2}(Y)$ строятся только для объектов, не создающих интервал пересечения гистограмм $h_{\omega_1}(Y)$, $h_{\omega_2}(Y)$. Далее эти объекты из обучающей выборки исключаются, и по новым обучающим выборкам строится линейная разделяющая поверхность (6) с получением группы частных функций принадлежности $\mu_{\omega_1}(Y_{m,k})$, где m – номер частной функции принадлежности в подпространстве k .

Составляющие (6) кусочно-линейной гиперповерхности строятся до получения заданного или возможно достижимого качества классификации, а уверенность в классификации относительно полученной кусочно-линейной разделяющей поверхности определяется через агрегацию $\mu_{\omega_1}(Y_{m,k})$ в соответствии с выражением

$$\mu_{k,\omega_\ell}^o = \max_m \left[\mu_{\omega_\ell}(Y_{m,k}) \right] . \quad (7)$$

Учитывая, что в ряде нейросетевых структур первый слой реализует кусочно-линейную разделяющую поверхность, описанный механизм может быть использован для обучения этого слоя.

Если количество классов больше двух, то обучающая выборка разбивается на две: класс, относительно которого строится нечеткое решающее правило (базовый

класс ω_1), и все остальные классы, от которых отделяется класс ω_1 (противоположный класс ω_2). Для вновь созданных обучающих выборок определяются функции принадлежности к классу ω_1 . Далее класс ω_1 заменяется следующим и т.д. до тех пор, пока не будут построены функции принадлежности для всех ℓ классов ($\ell = 1, \dots, L$).

В ходе разведочного анализа может быть установлен факт «вложения» одного класса в другой, например, так, как это показано для двумерного пространства на рисунке 3а.

В таком варианте дистальную гистограмму предлагается строить относительно центра, «охватываемого» класса в соответствии с выражением вида:

$$Y = \sqrt{\sum_{i=1}^n (a_i - x_i)^2}, \quad (8)$$

где a_1, \dots, a_n – координаты центра «охватываемого» класса.

Визуальный анализ такой гистограммы позволяет увидеть возможность разделимости линейно-неразделимых классов (рис. 3б) и подобрать соответствующие функции принадлежности.

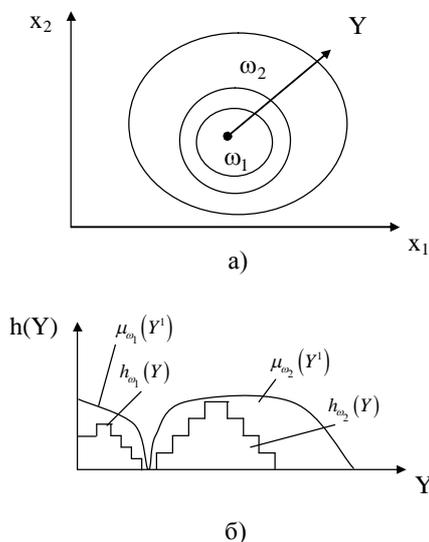


Рис. 3. Построение функций принадлежности для «вложенных» структур классов

В более общем случае если в ходе разведочного анализа выясняется целесообразность использования разделяющих гиперплоскостей с уравнением вида:

$$Y = F_{\omega_\ell}(A, X), \quad (9)$$

то целесообразно в качестве носителя для соответствующей функции принадлежности использовать расстояние между произвольной точкой $X^* = \{x_1^*, \dots, x_n^*\}$ и поверхностью (9).

Таким образом, частное или финальное решение о классификации будет приниматься на основании значения функции принадлежности $\mu_{\omega_\ell}(D)$, где

$D = F_D \left[X^*, F_{\omega_\ell}(A, X) \right]$; $A = a_1, \dots, a_n$ – вектор настраиваемых параметров, определяющий ориентацию разделяющей гиперповерхности в пространстве признаков $X = \{x_1, \dots, x_n\}$; F_{ω_ℓ} – функциональная зависимость, определяющая вид функции (9);

F_D – функциональная зависимость, определяющая способ расчета носителя D .

Если в ходе определения функций принадлежности достигается заданное качество классификации, то их можно считать мерой доверия к классу ω_ℓ , то есть отождест-



влять с мерой доверия к классу ω_ℓ в выражении (2), или, при отсутствии свидетельств, отвергающих гипотезу ω_ℓ , с соответствующим коэффициентом уверенности:

$$KU_{\omega_\ell} = \mu_{\omega_\ell}(Y).$$

В ряде случаев в задачах диагностики и прогнозирования состояния сложных систем геометрическая интерпретация нецелесообразна или неприемлема. Тогда принятие решения может быть осуществлено с использованием формул расчета коэффициентов уверенности в отнесении объекта к классу ω_ℓ - KU_{ω_ℓ} . В работе [5] для расчета предлагается использовать выражение типа (1). В более общем случае KU_{ω_ℓ} может быть вычислен с помощью других формул, форма и параметры которых выбираются в соответствии с типом решаемой задачи и ролью диагностических и (или) прогностических признаков x_i , входящих в общие правила принятия решений. В простейшем случае при отсутствии факторов недоверия к ω_ℓ (МНД=0) и при условии, что $\mu_{\omega_\ell}(Y)$ соответствует понятию меры доверия (МД), то коэффициент уверенности может совпадать с соответствующей функцией принадлежности, т.е. $KU_{\omega_\ell} = \mu_{\omega_\ell}(Y)$, где функция принадлежности может характеризовать уверенность отнесения объекта к классу ω_ℓ при наличии свидетельства, представляемого носителем Y , или может быть получена как частное решение о классификации, например, в подпространствах или пространствах признаков в геометрической интерпретации задачи распознавания. В частном случае $Y = x_j$.

При наличии нескольких свидетельств в пользу решения ω_ℓ или нескольких частных решений с частными коэффициентами уверенности решается задача синтеза более общего правила принятия решений или окончательного решения [1, 2].

В ряде медицинских приложений, особенно в задачах прогнозирования риска появления и развития заболеваний ω_ℓ , принимаемое решение зависит от времени воздействия на организм некоторого вредного фактора Y_j .

Интенсивность действия фактора Y_j совместно со временем воздействия в частном решающем правиле по фактору с номером j могут быть учтены при использовании правила вида:

$$KU_{\omega_\ell, j} = \begin{cases} 0, & \text{если } \mu_{\omega_\ell}(Y_j) \cdot \mu_{\omega_\ell, Y_j}(t) = 0 \\ \mu_{\omega_\ell}(Y_j) + \mu_{\omega_\ell, Y_j}(t) \cdot [1 - \mu_{\omega_\ell}(Y_j)], & \text{если } \mu_{\omega_\ell}(Y_j) \cdot \mu_{\omega_\ell, Y_j}(t) > 0, \end{cases} \quad (10)$$

где $\mu_{\omega_\ell}(Y_j)$ - функция принадлежности к классу ω_ℓ по шкале интенсивности действия Y_j ; $\mu_{\omega_\ell, Y_j}(t)$ - функция принадлежности к классу ω_ℓ от действия фактора Y_j с носителем по шкале времени воздействия.

Если выбрать форму и параметры функций принадлежности таким образом, что функции принадлежности растут с ростом величин их носителей, то при выполнении условия $\mu_{\omega_\ell}(Y_j) \cdot \mu_{\omega_\ell, Y_j}(t) > 0$, $KU_{\omega_\ell, j}$ растет по мере роста интенсивности и времени воздействия по закону, определяемому формой и параметрами соответствующих функций принадлежности.

Характерной особенностью ряда информативных признаков (артериальное давление, энергетические характеристики биологически активных точек и т.д.) является то, что они могут изменяться под воздействием однократных относительно коротких возмущающих внутренних и внешних воздействий, после чего, если организм обладает достаточным адаптационным потенциалом, возвращаться в рамки допустимых значений или находиться за рамками значений, считающихся нормой, достаточно длительное время (неделя, месяц, год и т.д.)

В первом случае не всегда речь идет о патологическом отклонении в организме человека, а второй случай, как правило, свидетельствует о высоком риске появления и развития заболеваний или об имеющейся патологии. Причем, чем большее время наблюдается отклонение контролируемого параметра за рамки своих номинальных значений, тем увереннее можно говорить о наличии гипотезы ω_ℓ .

Учесть патологическое отклонение измеряемых показателей от номинальных значений в формулах расчета коэффициентов уверенностей в гипотезах ω_ℓ можно, введя понятие различных уровней доверия. Например, к первому уровню доверия со своими частными решающими правилами следует отнести выводы, делаемые по результатам однократных измерений, ко второму уровню доверия – результаты, получаемые при устойчивой тенденции выхода измеряемых параметров за рамки номинальных значений в течение недели и т.д. С учетом сказанного целесообразно соответствующие функции принадлежности строить для каждого уровня доверия q функции принадлежности $\mu_{\omega_\ell}^q(x_i)$.

Расчет частных коэффициентов уверенности для каждого из уровней доверия, в зависимости от существования решаемой задачи и предпочтений экспертов, можно производить различными способами.

В первом варианте для каждого из выбранных интервалов времени записывается продукционное правило вида:

$$\text{ЕСЛИ } T_q \text{ ТО } [КУ_{\omega_\ell} = ?^q_{\omega_\ell}(x_i)] , \quad (11)$$

где T_q – интервал времени, в течение которого наблюдается выход x_i за рамки номинальных значений.

Во втором варианте коэффициент уверенности для различных уровней доверия может быть пересчитан из коэффициента уверенности, рассчитанного для первого уровня доверия $KU_{\omega_\ell}^1$ через весовой коэффициент, отражающий время удержания x_i в определенных рамках $\gamma_{\omega_\ell}^r(x_i, t)$, то есть:

$$КУ_{\omega_\ell}^r = \gamma_{\omega_\ell}^r(x_i, t) \cdot КУ_{\omega_\ell}^1 . \quad (12)$$

При выборе значений $\gamma_{\omega_\ell}^r(x_i, t)$ следует учитывать, что $KU_{\omega_\ell}^{\text{rmax}} < 1$. В третьем варианте дополнительно к $?_{\omega_\ell}(x_i)$ вводится функция принадлежности к классу ω_ℓ по времени отклонения x_i от номинального значения, а соответствующий частный коэффициент уверенности рассчитывается аналогично (9). В этом варианте уровень доверия к гипотезе ω_ℓ в явном виде не присутствует. Здесь наблюдается рост уверенности в ω_ℓ по мере роста отклонения x_i от своего номинального значения со стремлением $KU_{\omega_\ell, j}^*$ к своему верхнему значению, не превышающему единицы. $KU_{\omega_\ell, j}^*$ изменяется по закону, определяемому параметрами соответствующих функций принадлежности.

Используя описанный механизм синтеза нечетких решающих правил, решались задачи прогнозирования, ранней и дифференциальной диагностики заболеваний: желудочно-кишечного тракта, нервной и костно-мышечной систем, возникающих и развивающихся в условиях действия постоянного магнитного поля Курской магнитной

СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЯ В МЕДИЦИНЕ НА ОСНОВЕ СИНТЕЗА СТРУКТУРИРОВАННЫХ МОДЕЛЕЙ ОБЪЕКТОВ ДИАГНОСТИКИ

А. И. ПОВОРОЗНИУК

*Национальный
технический
университет
«Харьковский
политехнический
институт»*

e-mail: perederiy@mail.ru

Разработан метод синтеза структурированных моделей объектов диагностики при построении компьютерных систем поддержки принятия решений в медицине на всех этапах преобразования информации. При этом учитывается не только структурный и функциональный базисы, но и неопределенность параметров модели и экспертные оценки.

Ключевые слова: медицинская диагностика, компьютерная система, структурированная модель, неопределенность параметров, экспертные оценки.

Постановка проблемы и анализ литературы

Внедрение новых информационных технологий в медицину обуславливает переход от традиционных медицинских информационно-поисковых систем к интеллектуальным компьютерным системам поддержки принятия решений в медицине (ИКСППМ) с развитым математическим аппаратом и элементами экспертных систем. Разработчики ИКСППМ отмечают следующие актуальные оптимизационные задачи, которые требуют первоочередного решения [1, 2]:

- выделения диагностически-значимых структурных элементов на фоне артефактов и помех (структурная идентификация) при анализе биологических сигналов и медицинских изображений;
- формализация, оценка информативности и формирование минимально необходимого множества информативных признаков;
- синтез решающего правила (РП), согласно которому выполняется процесс диагностики пациентов.

Как отмечается в [1], применение традиционных математических методов (детерминистических, вероятностных, и др.) для решения отмеченных оптимизационных задач в медицине является не совсем корректным, так как применяемые методы рассматривают модель объектов диагностики (ОД) в виде "черного ящика", не учитывают сложности организации диагностируемых подсистем организма, особенностей формирования системы разнотипных диагностических признаков (числовые, ранговые, дихотомические), разнородность методик их получения (анкетирования, осмотр, клинические исследования, анализ биологических сигналов и медицинских изображений), а также сложной итерационной процедуры постановки диагноза (от предварительного до уточненного), поэтому актуальной является задача разработки адекватных моделей ОД и РП на этих моделях.

Цель работы

Целью работы является разработка методов построения интеллектуальных компьютерных системам поддержки принятия решений в медицине на основе единого формализованного подхода синтеза структурированных моделей объектов диагностики на всех этапах преобразования информации.

Синтез структурированных моделей объектов диагностики.

Синтез структурированных моделей ОД основан на концепции эволюционной идентификации структурированных моделей (*C*-моделей), разработанной в работах Букатовой И. Л. [3], при этом базовым понятием является *C*-модель, реализующая преобразование $F: X \rightarrow Y$, то есть отображение входного вектора X в выходной Y через набор операторов f из множества F ($f \in F$) в соответствии с некоторой структурой S , которая определяет порядок следования f . Таким образом, *C*-модель задается графом, вершинами которого являются функциональные элементы f заданного класса моделей, а информационная связь вершин задается структурой S , то есть $C = \{f, S\}$.



Учитывая отмеченные выше оптимизационные задачи при построении ИКСППМ, разнородность диагностических признаков, итерационную процедуру постановки диагноза и необходимость учета экспертных оценок при формировании диагностического вывода, автором в [2] формализованы следующие этапы преобразования информации в ИКСППМ:

- структурная идентификация биологических квазипериодических сигналов (БКС), которые отражают циклические процессы в организме – электрокардиограмма (ЭКГ), реограмма (РГ), и др. и медицинских изображений внутренних органов (МИ) – томограмма, рентгенограмма, УЗИ;
- формализация описания разнородных диагностических признаков;
- синтез иерархических структур диагностируемых состояний S_D и диагностических признаков S_Z ;
- синтез диагностических РП при взаимодействии S_D и S_Z ;
- синтез индивидуальных диагностических прогнозирующих моделей;
- рекомендации по выбору оптимальной методики лечения.

На каждом из отмеченных этапов решается локальная оптимизационная задача, для которой используется свой набор исходных данных, свой критерий оптимальности и алгоритм оптимизационной процедуры, поэтому в [4] автором вводится понятие понятие расширенной C - модели (C' -модель) и процедура реконфигурации модели. C' -модель задается следующим образом

$$C' = \{S, F, E, \varepsilon_\sigma, \varepsilon_\Delta\}, \quad (1)$$

где: $S = \{P, V\}$ – структура модели, которая задана множествами вершин P и дуг V ;

F – функциональный базис модели, причем функции могут приписываться как вершинам f_p , так и дугам f_v ;

E – экспертные оценки, которые при необходимости дополняют F компоненты;

ε_σ – неопределенность параметров F -базиса, которая зависит от статистических свойств обучающей выборки;

ε_Δ – неопределенность параметров F -базиса, которая зависит от точности их задания (шага квантования).

Значения ε_σ и ε_Δ используется не только в качестве оценки параметров модели, но и являются параметрами оптимизационных процедур при синтезе модели. При реализации последовательности этапов преобразования информации, на каждом из которых используется критерий Q_i , выполняется реконфигурация модели i -1 этапа путем структурных и функциональных преобразований. Рассмотрим применение C' -моделей при реализации отмеченных ранее этапов преобразования информации.

Этап выделение информативных структурных элементов БКС, параметры которых являются диагностическими признаками, является наиболее ответственным и трудоемким этапом их обработки, так как ошибки структурной идентификации (пропуск структурного элемента или ложная идентификация) приводят к грубым ошибкам вычисления диагностических показателей и ошибкам компьютерного диагноза (если не предусмотрены специальные средства выявления указанного типа ошибок, в частности отказ от обработки "подозрительного" периода или подтверждение человека-оператора правильности структурной идентификации).

Сложность идентификации состоит в том, что амплитудные и временные характеристики структурных элементов содержат информацию о работе диагностируемых подсистем организма и подвержены вариабельности в широком диапазоне, кроме того, эти сигналы могут содержать артефакты (резкое изменение некоторых характеристик в отдельных периодах) и снимаются на фоне помех. Традиционные методы обработки БКС во временной области (контурный анализ) являются по сути эвристическими, с трудом адаптируются к другим типам сигналов и не удовлетворяют возрастающим требованиям к качеству структурной идентификации.

Автором разработана формализованная процедура идентификации структурных элементов БКС на основе построения индивидуального пространства параметров для каждого типа структурного элемента с использованием методологии и вычислительных процедур преобразования Хока [4, 5]. При этом строится С'-модель (1) эталона каждого типа структурного элемента, параметрами которой являются:

- S – граф неточечного преобразования Хока фрагмента БКС из временной области $x(t_i)$ в пространство параметров $Y(P)$;
- F – вычислительные процедуры преобразования Хока;
- E – априорная информация о структуре БКС (адаптивная временная маска);
- ε_Δ – неопределенность квантования пространства параметров.

Синтез S_D выполняется процедурой иерархической кластеризации F1: $\{D\} \rightarrow S_D$ множества диагностируемых состояний $\{D\}$ в признаковом пространстве $\{X\}$, при этом параметрами С'-модели (1) являются:

- S – граф диагностируемых состояний (бинарное дерево решений);
- F – вычислительные процедуры кластеризации;
- E – экспертная информация о структуре симптомокомплексов.

Синтез S_z выполняется в результате преобразований F2: $\{X\} \rightarrow S_x$, F3: $S_x \rightarrow S_y$ и F4: $S_y \rightarrow S_z$ (F3 и F4 – реконфигурация структур). Параметры С'-модели (1):

- $S = \{S_x, S_y, S_z\}$ – графы диагностических признаков;
- F – вычислительные процедуры синтеза S;
- E – экспертная информация о структуре симптомокомплексов;
- ε_Δ – неопределенность квантования числовых признаков;
- ε_σ – неопределенность, обусловленная ограниченностью обучающей выборки.

Преобразование F2 удовлетворяет критериям минимума корреляционной связи и для его реализации в [2] предложена процедура иерархической кластеризации, основанная на представлении задачи кластеризации в виде потоковой модели. При этом исходные признаки представляются вершинами полносвязного графа, а дугам такого графа приписываются некоторые численные данные (k_{ij} – коэффициент парной корреляции, хотя возможно применение других статистических мер связи). Тогда задача иерархической кластеризации вершин сводится к последовательной процедуре поиска минимального разреза графа в потоковой задаче с ограничениями, для решения которой предлагается адаптация алгоритма “дефекта” [6]. В потоковых задачах с ограничениями каждая дуга V_{ij} характеризуется следующими параметрами: f_{ij} – поток по дуге; L_{ij} – нижняя пропускная способность дуги; U_{ij} – верхняя пропускная способность дуги; C_{ij} – стоимость прохождения единицы потока по дуге.

Для рассматриваемой задачи кластеризации диагностических признаков $L_{ij} = 0$, $U_{ij} = k_{ij}$. Значение C_{ij} при отсутствии априорной информации принимается равным 1 для всех дуг, а при ее наличии устанавливаются с помощью экспертных оценок, т.е. $C_{ij} = E_{ij}$, $E_{ij} < 1$ для тех вершин, которые соответствуют признакам, входящим в симптомокомплекс одного заболевания. Значения потоков f_{ij} определяются после завершения работы алгоритма “дефекта”, для которого необходимо задать начальную конечную вершины (исток и сток t).

Задание s и t выполняется на основе оценочной кластеризации вершин по методу корреляционных плеяд, при котором исходное множество вершин N распадается на два непересекающихся подмножества N_s и N_t , после чего в каждом из них определяется наиболее связанная с остальными вершинами ($s \in N_s$ и $t \in N_t$), удовлетворяющая условию $N_i = \max_j \sum_j C_{ij}$, причем $s = N_i \forall V_{ij} \in N_s$ и $t = N_i \forall V_{ij} \in N_t$. После определения потоков на всех дугах определяется минимальный разрез графа R, для которого справедливы соотношения:



$$\begin{aligned} f_{ij} &= U_{ij} \quad \forall V_{ij} \in (N_s, N_t), \\ f_{ij} &= 0 \quad \forall V_{ij} \in (N_t, N_s). \end{aligned} \quad (2)$$

Минимальный разрез разделяет все вершины графа на два непересекающихся множества (N_s , включающее начальную вершину s и N_t , включающее конечную вершину t) таким образом, что потоки насыщают все прямые дуги разреза и нулевые на всех обратных дугах разреза (2). В результате иерархической кластеризации синтезируется иерархическая структура S_x , на нижнем уровне иерархии которой могут быть как отдельные диагностические признаки (если они информативны и независимы), так и кластеры коррелированных признаков, а также схема объединения кластеров на остальных уровнях.

Преобразование F_3 выполняется по критерию минимума интегральной неопределенности ε_k , которая включает ε_Δ и ε_σ и возникает при замене динамического диапазона изменения значений числового признака при разбижке Δ на m неравномерных диагностически-значимых интервалов Δ_k и построения соответствующих гистограмм. В [7] получено выражение оценки ε_k и итерационная процедура формирования Δ_k , при которой минимизируется ε_k .

Целью преобразования F_4 является согласование топологий S_z и S_D для реализации диагностического РП и обеспечения оптимального плана обследования конкретного пациента. При этом, каждой паре элементов S_D , имеющей общего родителя, ставится в соответствие один элемент структуры S_z . Элементы структуры S_z формируются из элементов структуры S_y с учетом их информационной полноты

$$k_{\text{ин}}(x_j, D) = I_D(x_j) / H(D),$$

где $I_D(x_j)$ – информативность признака x_j относительно состояний D ; $H(D)$ – энтропия D , и диагностической ценности

$$k_{\text{дц}}(x_j, D) = k_{\text{ин}}(x_j, D) / r_i(x_j),$$

где $r_i(x_j)$ – суммарная оценка сложности измерения признака x_j .

В работе реализуется метод синтеза уточняющего диагноза [8], который является модификацией метода последовательного анализа (метода Вальда) и основан на анализе взаимодействия иерархических структур диагностических признаков S_z и диагнозов S_D . На каждом i -м этапе последовательного анализа, при дифференциальной диагностике между двумя диагнозами D_q и D_l , вычисляются отношения правдоподобия

$$\Theta = \prod_i \frac{E_{ikq} P(x_{ik} / D_q)}{E_{ikl} P(x_{ik} / D_l)}, \quad (3)$$

которое сравнивается с порогами

$$\Theta > A, \quad \Theta < B, \quad (4)$$

где A и B – верхняя и нижняя границы неопределенности, необходимые для принятия решения.

При выполнении одного из условий (4) принимается решение о диагнозе D_q или D_l соответственно и выполняется переход на более низкий уровень иерархии диагнозов с целью уточнения диагноза. При невыполнении обеих неравенств (4) добавляется следующий $i + 1$ признак и процедура повторяется. Для реализации комбинированного решающего правила, в выражении (3) учитывается неопределенность оценок условных вероятностей, кроме того, условные вероятности взвешиваются системой экспертных оценок о структуре симптомокомплексов E_{ikq} и E_{ikl} , методика расчетов которых приводится в [9].

Разработка структурной схемы СППРМ.

Для программной реализации рассмотренных ранее математических методов структурной идентификации ОД, СППРМ должна работать в двух режимах: режим обу-

чения и рабочий режим постановки диагноза. В режиме обучения выполняется формирование БЗ, в процессе которого выполняется синтез и настройка параметров структурированных моделей структурных элементов БКС и медицинских изображений, диагностических признаков и диагностируемых состояний, комбинированного РП. В режиме постановки диагноза выполняется анализ биологической информации, полученной от пациента и формирование диагностического заключения. На рис. 1 показана структурная схема СППРМ, в которой сплошными стрелками показаны информационные потоки в режиме обучения, а пунктирными – в режиме постановки диагноза.

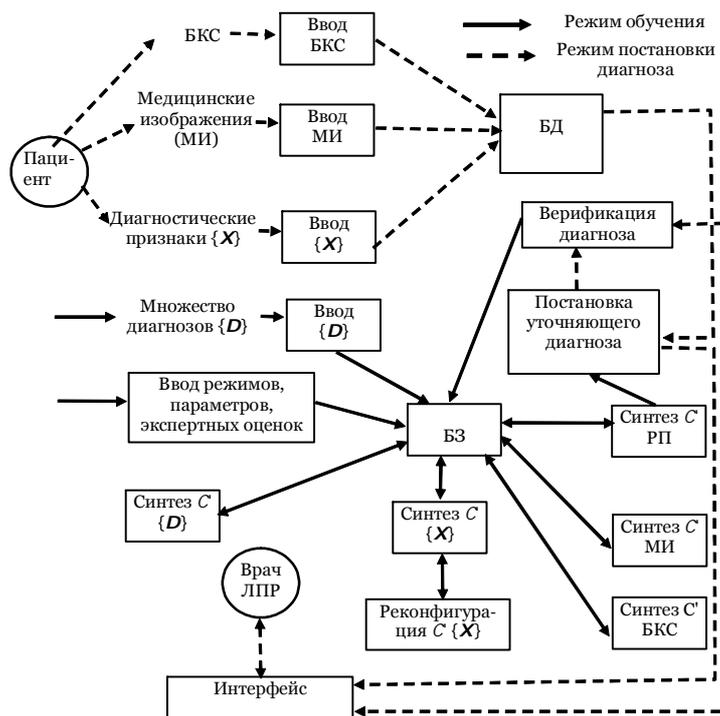


Рис. 1. Структурная схема СППРМ

Назначение модулей следующее:

- БД – база данных, которая хранит результаты обследований пациентов и верификацию диагностического заключения. Следует отметить, что при анализе БКС и медицинских изображений, в БД хранятся как исходные сигналы, так и результаты их структурной идентификации и сформированное множество вторичных диагностических признаков;
- БЗ – база знаний, которая хранит параметры отмеченных ранее структурированных моделей ОД.
- модуль ввода перечня диагностических признаков, диагностируемых состояний и экспертных оценок структур симптомокомплексов;
- модуль ввода значений первичных диагностических признаков;
- модули ввода БКС и медицинских изображений;
- модуль синтеза структурированных моделей структурных элементов БКС и медицинских изображений;
- модуль синтеза структурированной модели (иерархическая кластеризация) диагностируемых состояний;
- модуль синтеза и реконфигурации структурированной модели диагностических признаков;
- модуль реализации комбинированного РП;
- модуль формирования выходных экранных форм и документов, а также реализации интерактивного интерфейса с ЛПР.



Выполнен анализ средств для создания программного обеспечения, в результате которого обоснован выбор интерактивной среды программирования Inprise (Borland) Delphi на базе языка программирования высокого уровня Pascal, который характеризуется высокой точностью вычислений при работе с вещественными типами данных. Отмечена особенность языка явилась определяющей при значительном объеме вычислений, который выполняет СППРМ (особенно в режиме обучения). На этапе исследований и тестовых проверок использовался пакет "СТАТИСТИКА".

Апробация метода и обсуждение результатов.

Для апробация метода синтеза S_2 и реализации РП, взята выборка из 434 пациентов, у которых диагностировалось 10 диагнозов (в скобках указано количество пациентов): D_1 – здоров (243), D_2 – анемия неуточненная (29), D_3 – кардиомиопатия (10), D_4 – хроническая ишемическая болезнь сердца неуточненная (17), D_5 – аллергический геморрагический васкулит (24), D_6 – хронический холецистит (21), D_7 – пневмония неуточненная (16), D_8 – инфекционный мононуклеоз неуточненный (28), D_9 – хронический лимфоидный лейкоз (26), D_{10} – почечная недостаточность неуточненная (20). Анализировались 9 показателей клинического анализа крови: Eritr – эритроциты, Hb – гемоглобин, Leic – лейкоциты, Limf – лимфоциты, Eozin – эозинофилы, Pal – нейтрофилы палочкоядерные, Seg – нейтрофилы сегментоядерные, Mon – моноциты, Сое – скорость оседания эритроцитов. Полученные показатели качества классификации приведены в табл. 1.

Таблица 1

Показатели качества классификации

	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}	Среднее
Ψ %	97,53	75,86	40,00	29,41	91,67	85,71	62,50	85,71	88,46	95,0	88,48
Ψ' %	98,5	67	62	69	98,4	97,3	79	98,4	93	95	92,8
M_x	2,2	6,68	7,82	6,42	5,39	3,67	4,58	5,39	5,75	3,73	3,61

В табл. 1 в первой строке указана достоверность классификации Ψ стандартным методом дискриминантного анализа в пространстве 9 исходных признаков, а во второй Ψ' – предлагаемым методом уточняющего диагноза с использованием оптимальной схемы обследования пациентов. В третьей строке M_x – среднее число анализируемых признаков у одного пациента. Последний столбец таблицы рассчитан по формуле среднего для неравномерных групп

$$\bar{x} = \sum_i x_i \frac{N_i}{N},$$

где x_i – показатель качества; N_i – количество объектов в i -й группе; N – суммарный объем выборки.

Из табл. 1 следует, что в предлагаемом методе среднее число анализируемых признаков равно 3,61, что в 2,5 раза меньше исходных 9 признаков. При применении предлагаемого метода средняя достоверность диагноза повысилась с 88,48% до 92,8%, при этом из общего числа отрицательных ответов (31 ответ из 434) получена дифференциация 25 отрицательных ответов о невозможности постановки диагноза на предварительные диагнозы разного уровня. Кроме того, существенно повысилась достоверность тех диагнозов, у которых в обучающей выборке имеется недостаточно объектов (D_3 , D_4 , D_7 , ...) за счет незначительного снижения достоверности диагноза с наибольшей выборкой – D_1 . Это свидетельствует об эффективности применения предлагаемого метода на обучающих выборках с существенной неравномерностью количества объектов в анализируемых классах.

Выводы

Разработаны методы построения ИКСППМ на основе синтеза структурированных моделей ОД, которые учитывают не только структурный и функциональные

базисы, но и неопределенность параметров, а также позволяют учитывать экспертные оценки. Показано применение метода структурной идентификации на всех этапах преобразования информации. Разработанные методы позволяют снять ограничения на размерность пространства диагностических признаков, повысить достоверность компьютерного диагноза и адаптироваться к конкретным объектам медицинской диагностики.

Литература

1. Ахутин В.М. Оценка качества формализованных медицинских документов [Текст] / В.М. Ахутин, В.В. Шаповалов, М.О. Иоффе // Медицинская техника М.: 2002. № 2. С. 27 – 31.
2. Поворознюк А.И. Формализация этапов проектирования интеллектуальных компьютерных систем медицинской диагностики [Текст] / А.И. Поворознюк // Электронное моделирование К.: ИПМЕ, 2006. – Т. 28. №1 С. 85 – 97.
3. Букатова И.Л. Эвоинформатика: теория и практика эволюционного моделирования. [Текст] / И.Л. Букатова, Ю.И. Михасев, А.М. Шаров – М.: Наука, 1990.– 212с.
4. Поворознюк А.И. Система поддержки принятия решений в медицине на основе структурной идентификации объектов диагностики [Текст] / А.И. Поворознюк //Сборник научных трудов СНУЯЕиП. Севастополь: СНУЯЕиП, 2008. № 1 (25) С. 234 – 245.
5. Поворознюк А.И. Применение преобразования Хока для структурной идентификации физиологических сигналов // Моделювання та інформаційні технології. Збірник наукових праць інституту проблем моделювання в енергетиці. . К.: ИПМЕ, 2003. – Вип. 22, с.143 – 149.
6. Филлипс Д. Методы анализа сетей / Пер. с англ. [Текст] / Д. Филлипс, А. Гарсиа-Диас – М: Мир, 1984. – 648 с.
7. Поворознюк А.И. Формирование диагностических интервалов численных признаков при дифференциальной диагностике// Вісник Хмельницького національного університету. – Хмельницький: ХНУ, 2007.-№3, Т.1. С.106 – 109.
8. Поворознюк А.И. Метод постановки уточняющего диагноза в компьютерных системах медицинской диагностики при иерархической структуре диагностических признаков [Текст] / А.И. Поворознюк // Збірник наукових праць Харківського університету Повітряних Сил. Х: ХУПС, 2006. – Вип. 3(9). С. 125 – 130.
9. Поворознюк А.И. Интеллектуальная система поддержки принятия решений в медицине [Текст] / А.И. Поворознюк // Материалы XIII международной конференции с автоматического управления (Автоматика 2006), г.Винница 25-28 сентября 2006г. – Винница: УНИ-ВЕРСУМ-Винница, 2007. С. 364 – 368.

DECISION SUPPORT SYSTEM IN MEDICINE ON BASE OF THE SYNTHESSES OF THE OUTLINE MODELS OBJECT DIAGNOSTICES

A. I. POVOROZNYUK

*National technical university
"Kharkov polytechnic institute",
Kharkov, Ukraine*

e-mail: perederiy@mail.ru

The Designed method of the syntheses of the outline models object diagnostics at building computer decision support system in medicine on all stage of the transformation to information. Is it Here-with taken into account not only structured and functional bases, but also uncertainty parameter to models and expert estimations.

Key words: medical diagnostics, computer system, outline model, uncertainty parameter, expert estimations.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ФИНАНСОВЫХ ПРОЦЕССОВ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

М. Ф. ТУБОЛЬЦЕВ
В. М. МИХЕЛЕВ

*Белгородский
государственный
университет*

e-mail:
tuboltsev@bsu.edu.ru

e-mail:
mikhelev@bsu.edu.ru

Рассматриваются вопросы математического моделирования финансовых процессов накопительного типа в условиях неопределенности. Отличительной особенностью постановки рассматриваемой здесь задачи прогнозирования является то, что показана роль макро – и микроэкономических факторов.

В этой постановке модель создания накопительных фондов адекватно отражает реальную ситуацию, когда источник финансирования накопительного процесса не может быть точно задан, а подвержен случайным изменениям. Предложенный алгоритм решения задачи прогнозирования допускает эффективную реализацию с помощью современных вычислительных средств.

Ключевые слова: финансовые потоки, прогнозирование, случайные процессы, накопительные фонды, моделирование, компьютерное моделирование.

Введение

Во время экономического кризиса риски заимствования возрастают, поэтому большее внимание уделяется финансовым инструментам, которые являются альтернативой заимствованиям в качестве инструмента финансирования инвестиционных проектов [1]. Методика оптимального накопления фондов является достаточно проработанной для случая постоянного источника финансирования [2, 3, 4, 5]. В этих работах имеется ряд ограничений, которые в условиях экономического кризиса являются мало вероятными и затрудняют практическое применение.

Во-первых, ограничение на постоянство во времени источника финансирования накопительных фондов в условиях экономического и финансового кризиса необоснованно. В настоящее время мощность источников финансирования уменьшилась по сравнению с докризисным периодом, а по мере преодоления кризиса, она может снова увеличиться. В условиях относительной стабильности можно считать, что мощность источника финансирования накопительных фондов $u(t)$ является кусочно-постоянной функцией времени. В условиях нестабильности и неопределенности вложения в накопительный фонд $u(t)$ являются случайной величиной, и только от математического ожидания $u(t)$ можно требовать постоянства.

Во-вторых, в условиях кризиса редко выполняется требование постоянства процентных ставок, по которым начисляются проценты на средства накопительных фондов. Накопительные фонды формируются на счетах коммерческих банков, которые устанавливают размеры ставок самостоятельно и могут их изменять в зависимости от экономической ситуации. Поэтому ставка процентов $r(t)$ также является случайной функцией.

Процесс формирования (накопления) фонда в рамках динамических систем моделируется дифференциальным уравнением, содержащим в правой части случайные функции:

$$\dot{x}(t) = p(t)x(t) + u(t), \quad (1)$$

где $u(t)$ – интенсивность вложений в накопительный фонд, а $p(t) = \ln(1+r(t))$.

Если предполагать, что $u(t)$ и $p(t)$ – случайные функции, то нужно учитывать их различную относительно временных интервалов изменчивость. Случайная функция $p(t)$ является борелевской функцией ставки процентов, которая определяется макроэкономической ситуацией и изменяется, главным образом, из-за инфляции, как и ставка рефинансирования Центробанка. Ее изменчивость даже в условиях кри-

зиса не велика, и при рассмотрении краткосрочных финансовых операций может, в первом приближении, считаться постоянной. Только на интервалах времени больше базового периода ее изменчивость должна обязательно учитываться. Напротив, $u(t)$ – интенсивность вложений в накопительный фонд является быстро осциллирующей функцией, которая может изменяться подобно биржевым индексам. Случайным характером этой функции нельзя пренебрегать, если только она гарантированно не является детерминированной.

Таким образом, $x(t)$ является случайной функцией во всех случаях, когда $u(t)$ гарантированно не может считаться детерминированной и, следовательно, для исследования процесса накопления финансового актива (фонда) необходимо использовать теорию случайных процессов. Имеется большое число результатов по применению теории случайных процессов в изучении процессов на фондовых рынках, с чем можно ознакомиться в [6]. Однако накопительные финансовые случайные процессы исследованы мало, что является следствием недостаточного внимания теоретиков и практиков к накопительным финансовым инструментам. Возможно, финансовый кризис изменит ситуацию, и вопросы прогнозирования финансовых накопительных процессов будут должным образом проработаны, тем более что требуемый математический аппарат уже существует [6, 7].

В данной работе рассматриваются вопросы прогнозирования параметров краткосрочного накопительного процесса, что позволяет считать случайной только функцию вложений $u(t)$.

Теоретический анализ

Уточним постановку задачи формирования накопительного фонда. Пусть $x(t)$ – величина накопительного фонда в момент времени t и в начальный момент времени t_n все траектории случайного процесса выходят из нуля $x(t_n)=0$, производная $x(t)$ в левой части (1) понимается в смысле средней квадратичной сходимости. Поскольку рассматриваются вопросы прогнозирования параметров краткосрочного накопительного процесса, то изменчивостью $p(t)$ можно пренебречь и считать силу роста постоянной ($p(t)=p$). В первом приближении функцию вложений $u(t)$ можно считать при каждом t равномерно распределенной случайной величиной ($u(t)=U$) на отрезке $[U_{\min}, U_{\max}]$ и $U_{\max} > U_{\min} > 0$.

Найдем, при этих предположениях, математическое ожидание и дисперсию случайной величины $x(t_k)$, где $t_k > t_n$ некоторый произвольный момент времени. Поскольку уравнение (1) имеет в данном конкретном случае вид:

$$\dot{x}(t) = px(t) + U, \quad (2)$$

то его решение можно представить следующим образом:

$$x(t) = \int_{t_n}^t \exp(p(t-s))U ds, \quad (3)$$

поскольку интеграл в (3) существует [8, с. 158], то дифференцируя его по верхнему пределу, убеждаемся, что он является решением уравнения (2).

Используя свойства интеграла от случайной функции [8, с.159], можно легко получить выражение для математического ожидания $x(t)$:

$$M(x(t)) = \int_{t_n}^t \exp(p(t-s))M(U) ds. \quad (4)$$

Зная, что для равномерно распределенной на отрезке $[U_{\min}, U_{\max}]$ случайной величины U выполняется соотношение $M(U) = (U_{\max} + U_{\min})/2$, легко получаем из (4) значение математического ожидания для $x(t)$:

$$M(x(t)) = \frac{(U_{\max} + U_{\min})}{2p} (e^{p(t-t_n)} - 1). \quad (5)$$

Соотношение (5) позволяет найти момент времени t_k , когда математическое ожидание величины накопленного фонда достигнет некоторого заданного значения S . Подставив в (5) $t=t_k$ $M(x(t_k))=S$, получаем:

$$t_k = t_n + \frac{1}{p} \ln \left(1 + \frac{2pS}{U_{\max} + U_{\min}} \right). \quad (6)$$

Соотношение (6) в отличие от детерминированного случая, позволяет только оценить среднее значение накопительного фонда к некоторому заданному моменту времени. Исчерпывающую характеристику случайной величине $x(t)$ дает ее функция распределения. С ее помощью можно найти вероятность попадания случайной величины в любой заданный интервал. Поскольку интеграл в (3) можно понимать как траекторный, то имеет место равенство:

$$x(t) = \frac{U}{p} \left(e^{p(t-t_n)} - 1 \right), \quad (7)$$

из которого легко определить функцию распределения $x(t)$:

$$F_{x(t)}(y) = F_U \left(\frac{py}{e^{p(t-t_n)} - 1} \right). \quad (8)$$

Таким образом, можно не только оценить с помощью (6) среднее время формирования накопительного фонда, но и найти вероятность попадания величины накопленного фонда в заданный интервал, поскольку функция распределения U $F_U(y)$ известна.

Поэтому, задачу прогнозирования можно считать при сделанных предположениях полностью решенной. Можно отметить, что предположение о равномерном распределении $u(t)$ не является существенным. Тем же способом задача прогнозирования решается при любом законе распределения, если он не зависит от времени. Например, можно считать, что $u(t)$ нормальный закон распределения, а математическое ожидание и дисперсия не зависят от времени. Однако использование равномерного распределения выглядит в данной постановке наиболее естественно. Явная зависимость источника финансирования от времени может быть задана директивно либо спрогнозирована экономико-математическими методами. В этом случае модель требуется расширить с помощью стохастических дифференциальных уравнений.

Литература

1. Тубольцев М.Ф. Методы оптимального накопления фондов в бюджете развития муниципального образования. // «Научная мысль Кавказа», Ростов-на-Дону, Изд-во Северо-Кавказского научного центра высшей школы, 2005, №8. – с.82-91.
2. Тубольцев М.Ф. Оптимальные по быстродействию стратегии создания накопительных фондов. // «Научные ведомости», серия «Информатика, Прикладная математика, Управление», том 1 выпуск 1(19).- Белгород: Изд-во БелГУ, 2004.- стр.65-70.
3. Тубольцев М.Ф. Оптимальные по критерию минимума вложения средств стратегии создания накопительных фондов. // «Научные ведомости», серия «Информатика, Прикладная математика, Управление», № 1 (21) выпуск 2.- Белгород: Изд-во БелГУ, 2006.- стр.50-55.
4. Тубольцев М.Ф. Математическое моделирование систем накопительных фондов. // «Информационные технологии моделирования и управления», выпуск 8(33). – Воронеж: Изд-во «Научная книга», 2006. – стр. 990-995.
5. Тубольцев М.Ф. Математическое моделирование систем накопительных фондов // «Научные ведомости», серия «История, Политология, Экономик, Информатика», №1 (56) выпуск 9/1.- Белгород: Изд-во БелГУ, 2009.- стр.45-51.
6. Ширяев А.Н. Основы стохастической финансовой математики. Т. 1,2. –М.: Фазис, 1998.
7. Пугачев В.С., Сеницын И.Н. Теория стохастических систем. Учеб. Пособие. – М.: Логос, 2000. – 1000 с.: ил.
8. Миллер В.М., Панков А.Р. Случайные процессы в примерах и задачах. – М.: Изд-во МАИ, 2001. – 316 с.: ил.



MATHEMATICAL MODELLING OF FINANCIAL PROCESSES IN THE CONDITIONS OF UNCERTAINTY

M. F. TUBOLTSEV
V. M. MIKHELEV

*Belgorod State
University*

*e-mail:
tuboltsev@bsu.edu.ru*

*e-mail:
mikhelev@bsu.edu.ru*

Questions of mathematical modeling of financial processes of memory type in the conditions of uncertainty are considered. Distinctive feature of statement of a problem of forecasting considered here is that the role of macro – and micro economic factors is shown.

In this statement the model of creation of memory funds adequately reflects a real situation when the source of financing of memory process cannot be precisely set, and is subject to casual changes. The offered algorithm of the decision of a problem of forecasting supposes effective realization by means of modern computing means.

Key words: financial streams, forecasting, casual processes, memory funds, modeling, computer modeling.

ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 519.6

НОВЫЕ ТЕХНОЛОГИИ ПЕРЕДАЧИ И ОБРАБОТКИ ИНФОРМАЦИИ НА ОСНОВЕ ХОРОШО ЛОКАЛИЗОВАННЫХ СИГНАЛЬНЫХ БАЗИСОВ

В. П. ВОЛЧКОВ

*Московский
технический
университет
связи и информатики*

e-mail: volchkovvalery@mail.ru

Анализируются основные проблемы, возникающие при передаче и обработке сигналов в беспроводных высокоскоростных сетях и существующие методы их решения. Предлагается новый более эффективный подход к формированию и обработке сигналов на основе OFTDM технологии, использующей хорошо локализованные сигнальные базисы Вейля-Гейзенберга. Рассматриваются основные преимущества этой технологии и возможные применения для обработки информации в различных технических приложениях.

Ключевые слова: цифровые системы связи, OFDM система, OQAM, базис Вейля-Гейзенберга, ортогональный базис, хорошая локализация.

1. Постановка проблемы и существующие методы решения

При построении беспроводных высокоскоростных цифровых систем связи часто возникает ситуация, когда реальный радиоканал (среда распространения) обладает частотно-временным рассеянием. Последнее обычно вызвано тем, что сигнал в точку приема приходит многочисленными путями после многократных отражений от нестационарных неоднородностей среды (городские постройки, движущиеся предметы, гидрометеоры, слои ионосферы и др.).

Примерами таких дисперсионных каналов могут служить, например, КВ и УКВ радиолнии спецназначения, радиолнии мобильных сетей широкополосного доступа (мобильный Wi-Max) и цифрового телевидения (DVB-T /H).

Наиболее эффективной технологией передачи данных в таких каналах является мультиплексирование (уплотнение) с ортогональным частотным разделением (OFDM – Orthogonal Frequency Devision Multiplexing). Поэтому в дальнейшем в качестве сигнала- переносчика информации в радиолнии будет пониматься OFDM сигнал.

В результате частотно-временного рассеяния OFDM сигнала, на приемной стороне наблюдаются эффекты многолучевости, амплитудно-фазовых замираний, доплеровского расширения и частотного сдвига. Их действие на приемник ведет к меж-

символьной (МСИ) и межканальной (МКИ) интерференциям, которые значительно ухудшают характеристики приема [1].

Отметим, что структура и свойства любого цифрового сигнала однозначно определяются собственным конечномерным базисом, из которого он формируется, как некоторая линейная комбинация базисных функций. Коэффициенты линейной комбинации (модулирующие символы) могут быть вещественными или комплексными и определяются выбранным сигнальным созвездием (например QAM, PSK, и др.). Каждая базисная функция является независимым переносчиком модулирующих символов и в дальнейшем ассоциируется с некоторым поднесущим каналом. Обычно, сигнальный базис выбирается ортогональным, чтобы в случае гауссовского канала обеспечить отсутствие межсимвольной интерференции (выполнение теоремы Найквиста) и упростить реализацию оптимального приемника.

В каналах с частотно-временным рассеянием (ЧВР) на сигнал кроме аддитивного шума действует сложная мультипликативная помеха. У такой помехи фактор временного рассеяния характеризуется мультипликативным воздействием в частотной области (сверткой во временной области), а фактор частотного рассеяния – мультипликативным воздействием во временной области (сверткой в частотной области). В результате, базис сигнала деформируется, становясь неортогональным, а теорема Найквиста не выполняется.

Фактически возникновение МСИ и МКИ объясняется потерей ортогональности между «возмущенными» базисными функциями сигнала на выходе канала. В результате чего, процедура демодуляции сигнала на приемной стороне оказывается неоптимальной. Возникает просачивание информации из каждого поднесущего канала, в соседние. Причем, величина этих взаимных помех зависит от частотно-временной локализации (ЧВЛ) сигнальных базисных функций и определяется эффективным носителем их функции неопределенности. Чем быстрее спадают «хвосты» функции неопределенности, тем лучше ЧВЛ сигнального базиса, а значит меньше уровень МСИ и МКИ.

Известно, что в каналах с частотно-временным рассеянием технология OFDM позволяет очень эффективно бороться с межсимвольной интерференцией (МСИ), благодаря циклическому префиксу (защитному интервалу), передаваемому в начале каждого OFDM символа. В результате, влияние временного рассеяния на сигнал можно практически полностью скомпенсировать. Учитывая это, в дальнейшем основное внимание будет уделено борьбе с МКИ, возникающей из-за частотного рассеяния сигнала.

Рассмотрим некоторые известные подходы к подавлению МКИ, которые предлагались в разное время разными авторами в журнальных публикациях и докладах научно-технических конференций и семинаров.

В системах связи, использующих принцип OFDM передачи (Wi-Fi, Wi-Max, DVB-T/H и др.), сигнальные базисные функции представляют собой отрезки гармоник, а их амплитудный спектр имеет узкий основной лепесток и медленно спадающие хвосты, как у функции $|\sin(kx)/kx|$. В этих условиях уменьшить чувствительность к доплеровским эффектам канала связи можно, расширяя спектр базисных функций, т.е. уменьшая длительность передаваемых гармоник. Это ведет к расширению спектра OFDM сигнала и увеличению расстояния между поднесущими каналами, что не всегда допустимо в рамках используемого стандарта связи. Кроме того, это не спасает от межканальных помех, вызванных просачиванием информации через боковые лепестки спектра $|\sin(kx)/kx|$ соседних гармоник. Однако, поскольку МКИ действует во всей полосе частот OFDM сигнала, ее дополнительное подавление может быть получено созданием защитных интервалов (в виде «нулевых» поднесущих) на границах частотного диапазона и между информационными поднесущими.

Отметим, что оба указанных инструмента борьбы с МКИ (расширение спектра и «нулевые» поднесущие) уменьшают спектральную эффективность системы связи.



Однако на это идут и в разумных пределах используют при разработке соответствующих стандартов.

Другой подход к уменьшению МКИ связан с применением оконного преобразования Фурье [2]. В этом случае, мы ничего не меняем на передающей стороне, но в алгоритме демодуляции OFDM сигнала исходный ортогональный базис БПФ заменяем на неортогональный базис взвешенного БПФ, обладающий хорошей локализацией в частотной области. Это позволяет уменьшить составляющую межканальной помехи, вызванную просачиванием информации через боковые лепестки спектров соседних гармоник. Однако, весовая функция окна расширяет полосу каждого поднесущего канала приемника, увеличивая просачивание информации через основной лепесток спектра. Кроме того, нарушение свойства ортогональности в случае оконного БПФ увеличивает шумовую составляющую.

Третий подход к решению описанной проблемы связан с различными обобщениями теоремы Котельникова [3]. Суть их сводится тому или иному варианту «пересыщения» (oversampling) отсчетами и использовании для интерполяции сигналов обобщенных рядов Котельникова с хорошо локализованным по частоте ядром. Последнее означает, что принимаемый сигнал дискретизируется с частотой гораздо большей, чем критическая частота Найквиста, а для восстановления или цифровой обработки сигнала используются, так называемые, атомарные функции. Набор таких равномерно сдвинутых по частоте атомарных функций образует сигнальный базис, который хорошо локализован в частотной области, но не является обязательно ортогональным.

За счет «пересыщения» отсчетами, основной лепесток спектра атомарной базисной функции может быть выбран по ширине таким же, как у функции $|\sin(kx)/kx|$, описывающей поднесущий канал OFDM сигнала. Причем сохраняется быстрый спад боковых лепестков спектра. Это является достоинством третьего подхода. Однако, к сожалению, подходящий (по локализации в частотной области) сигнальный базис из атомарных функций обычно оказывается неортогональным. Это усложняет вычислительную реализацию и уменьшает помехоустойчивость по отношению к шумовым помехам.

Напомним, что все перечисленные подходы использовались для каналов связи с частотным рассеянием, которые характерны для мобильных систем передачи и сетей широкополосного доступа (например, DVB-H, мобильный Wi-Max и др.). При разработке стандартов для этих мобильных приложений складывалась не самая лучшая ситуация, когда ряд положений стандартов были сформулированы раньше, чем производители приступили к выпуску соответствующей аппаратуры. Кроме того, фактор времени и конкуренция не позволяли тщательно проверить и обосновать выбор всех параметров сигнально-кодовой конструкции. В результате, первые попытки производителей создать работающую аппаратуру мобильной связи по этим стандартам, используя свои старые наработки, натолкнулись на серьезные проблемы.

Часто возникала следующая ситуация – аппаратура прекрасно работает, если мобильный абонент передвигается с небольшой скоростью, и связь резко ухудшается или обрывается, если абонент увеличивает скорость. Причина – срыв синхронизации и недостаточно точная оценка параметров канала в условиях частотного рассеяния сигнала, т.е. повышенного уровня МКИ.

Поскольку принятый стандарт изменить нельзя, у разработчиков оставался только один выход – соревноваться между собой в совершенствовании приемной аппаратуры, применяя более сложные методы обработки сигналов и оптимизируя вычислительные и аппаратные затраты на их реализацию.

Вывод: Задача борьбы с межканальной интерференцией в мобильных OFDM системах является очень актуальной, и во многих случаях пока не находит удовлетворительного решения.

2. Разработка новых методов формирования и обработки сигналов для мобильных беспроводных широкополосных систем связи

2.1. OFTDM технология

Как уже отмечалось выше, борьба с частотным рассеянием представляет серьезную проблему в мобильных сетях. Возникающие при этом взаимные помехи между поднесущими (МКИ) действуют во всем диапазоне частот OFDM сигнала, а их интервал корреляции в частотной области зависит от формы основного лепестка и скорости убывания боковых лепестков спектра базисной функции. Причем, природа межканальных помех такова, что их нельзя спекенсировать или отфильтровать обычными методами цифровой обработки, а уровень МКИ зависит от структуры сигнального базиса.

С другой стороны, если одновременно эффективно не бороться с временным рассеянием сигнала, многолучевое сложение в точке приема принесет дополнительную составляющую в МКИ.

Поэтому в качестве основного метода борьбы с частотно-временным рассеянием (ЧВР) сигнала *предлагается разработка и применение специальных ортогональных хорошо локализованных сигнальных базисов*, минимизирующих уровень взаимного влияния поднесущих каналов, как в частотной, так и во временной областях, но при этом не ухудшающих спектральную эффективность системы [4–8].

Оптимально локализованным назовем такой ортогональный сигнальный базис $B = \{\psi_{k,l}(t)\}_{k=0, l=0}^{M-1, L-1}$, который в условиях ЧВР приводит к минимальным значениям МСИ и МКИ и обладает максимальной частотно-временной плотностью упаковки.

Последнее требование означает, что упаковка, составленная из эффективных носителей функций неопределенности всех базисных функций $\psi_{k,l}(t)$, образует в частотно-временной плоскости прямоугольную однородную решетку с параметрами $\tau_o, \nu_o : \tau_o \nu_o = 1/2$, т.е. с минимальным фундаментальным объемом (см. рис. 1).

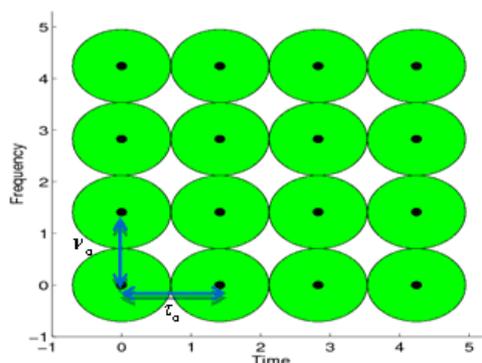


Рис. 1. Упаковка носителей WH-базисных функций в частотно-временной области

Можно показать, что в наиболее плотный хорошо локализованный ортогональный базис с указанной структурой может быть получен из равномерно сдвинутых по времени и частоте версий нескольких комплексных инициализирующей функций $\psi_{k,l}(t)$ с хорошей частотно-временной локализацией (обобщенный базис Вейля-Гейзенберга) и представляется в виде

$$B = \{\psi_{k,l}(t) = \Phi_{k,l}(t - l\tau_o) \exp(2\pi j k \nu_o t)\}, \quad k = 0, \dots, M-1, \quad l = 0, \dots, 2L-1 \quad (1)$$

$$\Phi_{k,l}(t) = (-j)^{\text{rem}(l,2)} g(t) \exp(j\alpha k), \quad (2)$$

где $g(t)$ – вещественный формирующий импульс; $\text{rem}(l,2)$ – остаток от деления $l/2$; α – фазовый параметр. Количество различных инициализирующих функций зависит от значения фазового параметра α и равно $p = 2 \text{fix}(2\pi/\alpha)$, где $\text{fix}(\cdot)$ – оператор взятия целой части числа. Размерность соответствующего сигнального пространства, натянутого на базис (1), равна $N_0 = 2ML$.



Параметры локализации базисных функций в частотной и временной областях (размеры осей эллипсов на рис. 1) можно настраивать и размещать между собой в любом требуемом соотношении, адаптируя их к конкретным условиям решаемой задачи.

Для сравнения отметим, что хорошо локализованных ортонормированных базисов той же размерности N_0 и состоящих из равномерно сдвинутых по времени и частоте версий одной инициализирующей функции $\Phi(t) = g(t)$ (классический базис Вейля-Гейзенберга):

$$B = \{ \psi_{k,l}(t) = \Phi(t - l\tau_o) \exp(2\pi j k v_o t) \}, \quad k = 0, \dots, M-1, \quad l = 0, \dots, 2L-1, \quad (3)$$

не существует (теорема Balian-Low, [2]). Более того, при $M > 2$ базисы (3) оказываются плохо локализованными, как в частотной, так и во временной области.

Таким образом, наличие нескольких инициализирующих функций является важнейшим требованием существования плотно упакованных хорошо локализованных ортогональных базисов.

В дальнейшем базисы (1) будем выделять аббревиатурой WH (Weyl-Heisenberg), а методы формирования и обработки сигналов с использованием WH-базисов будем называть OFTDM технологией (Orthogonal Frequency-Time Devision Multiplexing – мультиплексирование с ортогональным частотно-временным разделением).

Систему цифровой связи, построенную на основе данной технологии, будем называть WH-OFTDM или, сокращенно, OFTDM системой (т.е. системой с ортогональным частотно-временным мультиплексированием).

Пакетная передача информации в такой системе осуществляется с помощью OFTDM символов, представляющих собой линейную комбинацию базисных функций (2):

$$s(t) = \sum_{l=0}^{2L-1} \sum_{k=0}^{M-1} a_{k,l} \psi_{k,l}(t), \quad (4)$$

где $a_{k,l}$ – реальные (при l четном) или мнимые (при l нечетном) части комплексных QAM символов $c_{k,l} = a_{k,2l} + j a_{k,2l+1}$. В результате, пакет информации из $N = ML$ комплексных модулирующих QAM символов с помощью WH-базисных функций равномерно распределяется по частотно-временной плоскости в прямоугольнике размером $(T, W) = (2L\tau_o, Mv_o)$.

Так как WH-базис (1) обладает наибольшей (критической) частотно-временной плотностью, то спектральная эффективность OFTDM системы максимальна и, без учета защитных интервалов, равна $\eta = \beta / (2\tau_o v_o)$, где β – количество бит, передаваемых в одном QAM символе.

Заметим, что если $L = 1$, а формирующий импульс $g(t)$ имеет прямоугольную форму, то формулы (1)-(2) описывают гармонический базис Фурье классического OFDM сигнала, который плохо локализован в частотной области. Соответствующее выражение для OFDM символа преобразуется к виду

$$s(t) = \sum_{k=0}^{M-1} c_{k,l} \psi_{k,l}(t). \quad (5)$$

Из (4), (5) следует, что каждая базисная функция в OFTDM символе переносит только одну (вещественную) размерность, а в OFDM символе – две размерности. Поэтому, без учета защитных интервалов, спектральные эффективности OFDM и OFTDM систем одинаковы.

С другой стороны, если для обеих систем определены одна и та же полоса частот W и количество поднесущих M , то за счет выбора $L > 1$ OFTDM система в рамках одного символа (4) позволяет получить дополнительное внутрисимвольное уплотнение каналов во временной области. Причем, это уплотнение практически идеальное, поскольку на каждой поднесущей частоте k соседние сдвинутые по времени базисные функции $\psi_{k,l}(t)$ и $\psi_{k,l+1}(t)$ хорошо локализованы и ортогональны, а защитные интервалы между ними отсутствуют (см. рис. 1)

Отметим, что в OFDM системах внутрисимвольное временное уплотнение отсутствует ($L=1$), и для временного уплотнения символов (5) в кадр приходится использовать специальные защитные интервалы (циклические префиксы). Величина циклического префикса для стандартов Wi-Max и DVB-T может достигать 1/4 длительности OFDM символа, что ведет к значительным энергетическим потерям и снижению реальной спектральной эффективности системы.

Вывод 1. Применение OFTDM технологии (вместо технологии OFDM) в каналах с частотно-временным рассеянием позволяет за счет организации внутрисимвольного временного уплотнения существенно повысить, как спектральную, так и энергетическую эффективность системы.

Другое преимущество связано с хорошей локализацией базисных функций (1) в частотной области. В результате, спектр OFTDM символа (4) имеет быстро спадающие хвосты, а значит низкий уровень внеполосного излучения. Это существенно снижает требования к фильтрации сигнала на выходе передатчика, и, как следствие, уменьшению необходимого числа защитных «нулевых» поднесущих на границах частотного диапазона. Освободившиеся поднесущие могут быть использованы для передачи полезной информации.

Вывод 2. Применение OFTDM технологии позволяет значительно понизить уровень внеполосного излучения, и тем самым ослабить требования к выходному фильтру передатчика и защитному интервалу на границах частотного диапазона.

Третье преимущество связано с устойчивостью (робастностью) системы к МСИ и МКИ внутри каждого OFTDM символа. Оно обусловлено тем, что квадрат модуля WH-базисной функции и ее спектра (т.е. соответствующие сечения функции неопределенности) имеют достаточно плоские вершины и быстро спадающие хвосты. В результате, частотно-временные возмущения деформируют каждую базисную функцию (2) не столь значительно (как в случае OFDM символа), а возникающее при этом просачивание информации в соседние поднесущие каналы распространяется только на ближайших соседей.

Чтобы минимизировать уровень взаимных помех, параметры функции неопределенности WH-базиса должны быть согласованы с параметрами функции рассеяния канала. Базис с такими свойствами мы выше определили как оптимально локализованный.

Вывод 3. Применение OFTDM технологии позволяет повысить устойчивость телекоммуникационной системы к межканальной и межсимвольной интерференции и адаптировать ее к параметрам частотно-временного рассеяния среды.

Из проведенного выше сравнительного анализа следует, что *технология OFTDM является очень перспективным направлением для разработки новых мобильных широкополосных систем, как специального, так и гражданского назначения.*

2.2. Применение OFTDM технологии для борьбы с межканальной интерференцией в рамках существующих стандартов Wi-Max и DVB

Отметим, что в наиболее полном объеме преимущества OFTDM технологии реализуются, если WH-базис применяется и для формирования сигнала, и для его обработки на приемной стороне (см. выше выводы 1-3).

Однако, если стандарт мобильной связи или вещания уже принят и основан на использовании классической OFDM технологии (например, мобильная сеть Wi-Max или DVB-T/H), то WH-базис может быть успешно использован для проектирования приемной части, устойчивой к действию МСИ и МКИ. (Поскольку практически любой стандарт оставляет большую степень свободы для разработчиков приемной аппаратуры).

В этом случае для расфилтровки и демодуляции OFDM сигнала предлагается использовать модифицированный вариант OFTDM технологии, при которой сигнал обрабатывается в сдвинутых по времени ортогональных окнах, формируемых с помощью WH-базиса (1). В каждом окне выполняется расфилтровка сигнала на M ор-

тогональных поднесущих каналов. Параметры локализации этих каналов в частотной области, их количество M и расстояние между ними ν_0 согласуются со структурой передаваемого OFDM символа и выбираются с учетом максимально возможного подавления МКИ. Далее выполняются оптимальное комплексирование результатов расфилтровки и демодуляция OFDM символа.

Отметим, что такая обработка сигналов в неявном виде содержит L -кратный oversampling (пересыщение отсчетами), т.е. может рассматриваться как дальнейшее развитие метода атомарных функций на случай, когда равномерным сдвигам по времени и частоте подвергается не одна, а несколько таких функций при обязательном сохранении их взаимной ортогональности.

2.3. Экспериментальное сравнение базисных функций OFTDM и OFDM систем

В качестве иллюстрации, на рис. 2 приведены графики модулей базисных функций классической OFDM системы и оптимальной WH-OFTDM системы, а на рис. 3 и рис. 4 – графики модулей спектров этих же функций в линейном и логарифмическом масштабах, соответственно. Для наглядности, на рисунках изображены только по одному представителю из соответствующих сигнальных базисов, отвечающих одному и тому же сдвигу по времени и частоте.

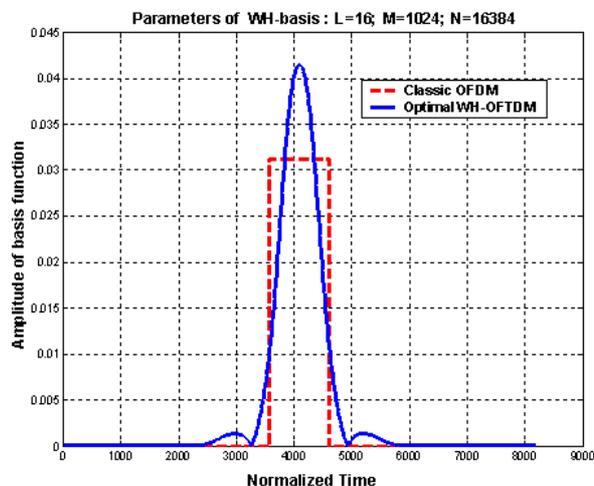


Рис. 2. Модули базисных функций OFDM и OFTDM систем

Число временных и частотных сдвигов в сигнальном базисе WH-OFTDM системы обозначается, соответственно, через L и M . Их произведение $N = LM$ определяет размерность WH-базиса.

Чтобы корректно сопоставить базисные функции OFDM и WH-OFTDM систем, их эффективная ширина спектра в частотной области (равная расстоянию между поднесущими частотами) и общее число поднесущих частот выбраны одинаковыми.

Из сравнения кривых на рис. 3 видно, что спектр базисных функций классической OFDM системы имеет достаточно острый пик и медленно спадающие «хвосты» типа $|\sin(kx)/kx|$. В то время, как аналогичный спектр у WH-OFTDM системы имеет пологую вершину, быстро спадает и практически не имеет боковых лепестков.

Согласно рис. 4, первый боковой лепесток подавляется на 60 дБ, а на расстоянии двух поднесущих от главного лепестка подавление составляет более 80 дБ. Последнее означает, что резервируя на границах частотного диапазона OFTDM сигнала как минимум две нулевых поднесущих, можно уже гарантировать уровень подавления внеполосного излучения 80 дБ без применения специальных фильтров.

Во временной области (рис. 2) хорошая локализация у WH-базисных функций сохраняется, но по форме они существенно отличаются от прямоугольных базисных функций классической OFDM системы.

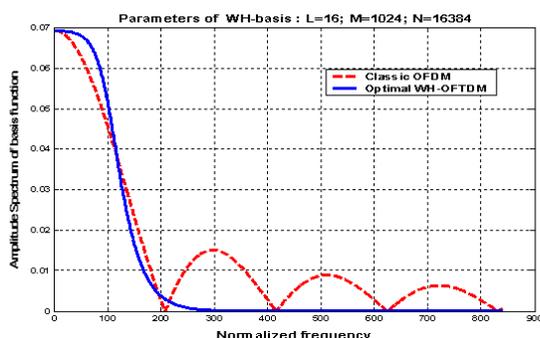


Рис. 3. Модули спектров базисных функций OFDM и OFTDM систем в линейном масштабе

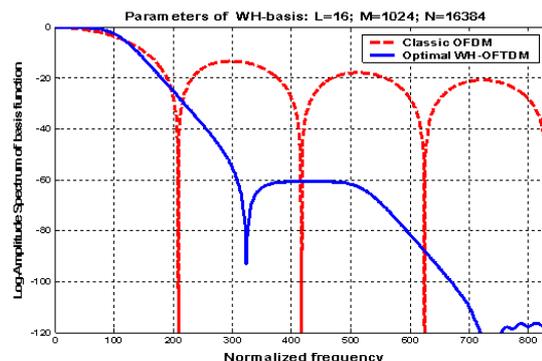


Рис. 4. Модули спектров базисных функций OFDM и OFTDM систем в логарифмическом масштабе

3. Другие области применения OFTDM технологии

Кроме перечисленных связанных приложений, области практического применения WH-базисов гораздо шире. Наиболее важные из них перечислены ниже.

3.2. Применение хорошо локализованных базисов для эффективного спектрально-временного анализа различных процессов

В этом случае удается получить гибкие многоуровневые алгоритмы анализа процессов, наблюдаемых на выходе различных устройств регистрации. В качестве таких устройств, например, могут выступать

- датчики биомедицинских приборов;
- приемники эхо сигналов в радиолокаторах (или гидролокаторах) различного назначения;
- сейсмографические датчики, и т.д.

Особенность спектрально-временного WH-анализа состоит в том, что базисные функции, по которым разлагается наблюдаемый процесс, ортогональны и обладают хорошим разрешением, как в частотной, так и во временной области (см. рис. 1 – 4). Более того, уровень разрешения в этих двух областях может гибко изменяться, позволяя исследователю более детально изучать свойства наблюдаемого процесса. Важным достоинством является также существование быстрых прямых и обратных алгоритмов спектрального WH-анализа.

3.3. Применение хорошо локализованных базисов для идентификации и классификации объектов (процессов) по частотно-временным признакам

Преимущество использования WH-базисов в этом случае состоит в том, что процесс (объект), подлежащий обнаружению или различению (на фоне других аналогичных), аппроксимируется многофакторной параметрической моделью в частотно-временной области. В качестве факторов выступают задействованные временные и частотные размерности базиса (1), а в качестве параметров – коэффициенты разложения процесса по соответствующему этим факторам ортогональному WH-базису. При этом, характеристики частотно-временной локализации базисных функций могут гибко меняться и служат дополнительным параметром для адекватной настройки модели. Достоинства такой WH-модели :

- хорошие аппроксимирующие свойства (широкий диапазон аппроксимации, обусловленный большим набором факторов и параметров);
- существование быстрого оптимального алгоритма идентификации, который фактически строится на быстром разложении процесса в ортогональном WH-базисе.



Заключение

Благодаря высокой плотности упаковки WH -базисных функций, их взаимной ортогональности, возможности гибкой перестройки параметров локализации и существованию быстрых вычислительных алгоритмов, OFTDM технология может быть использована для эффективной частотно-временной обработки практически любой информации природного характера. Более того, во многих приложениях эта технология составляет мощную конкуренцию вейвлет-технологии [2], где совместить хорошую локализацию, симметрию и ортогональность базисных функций не всегда удается.

Литература

1. Прокис Дж. Цифровая связь [Текст]: пер. с англ. / Под ред. Д. Д. Кловского. – М.: Радио и связь, 2000.
2. Добеши И. Десять лекций по вейвлетам [Текст]: пер. с англ. / Под ред. А.П. Петухова. – Ижевск: НИЦ «Регулярная и хаотическая динамика», 2001.
3. Кравченко В. Ф. Цифровая обработка сигналов и изображений [Текст]/Под ред. В. Ф. Кравченко. – М.: ФИЗМАТЛИТ, 2007.
4. Волчков В.П., Казаков Д.Ю. Синтез оптимальных сигнальных базисов Вейля-Гейзенберга для OFDM систем [Текст]// Научные ведомости БелГУ. Серия «Информатика и прикладная математика», №1(21), Вып. 2. – Белгород: Изд-во БелГУ, 2006. – С.107-118.
5. Волчков В.П. Сигнальные базисы с хорошей частотно-временной локализацией [Текст] // Электросвязь, №2. – Москва, 2007. – С. 21-25.
6. Волчков В.П., Петров Д.А. Оптимизация ортогонального базиса Вейля-Гейзенберга для цифровых систем связи, использующих принцип OFDM/OQAM передачи [Текст]// Научные ведомости БелГУ. Серия «История. Политология. Экономика. Информатика», №1(56), Вып. 9/1. – Белгород: Изд-во БелГУ, 2009. – С.102-110.
7. Волчков В.П., Петров Д.А. Обобщенная теорема Найквиста для OFTDM сигналов [Текст] // Сборник докладов. Всероссийский научно-технический семинар «Системы синхронизации формирования и обработки сигналов для связи и вещания» – Воронеж, 2009. – С. 28-32.
8. Volchkov V.P., Petrov D.A. Orthogonal Well-Localized Weyl-Heisenberg Basis Construction and Optimization for Multicarrier Digital Communication Systems [Электронный ресурс] // International Conference on Ultra Modern Telecommunications (ICUMT 2009), St. Petersburg, Russia, Oct 12-14, 2009, ISBN: 978-1-4244-3941-6, IEEE Catalog Number: CFP0963G-CDR (<http://ieeexplore.ieee.org>)

A NEW TECHNOLOGY OF TRANSMITTING AND PROCESSING OF INFORMATION BASED ON WELL-LOCALIZED SIGNAL BASIS

V. P. VOLCHKOV

*Moscow Technical
University
of Communications
and Informatics*

*e-mail:
volchkovvalery@mail.ru*

There are analysed main problems of transmitting and processing of signals in wireless high rate networks and known methods of their decision. There are proposed a new more effective approach for transmitting and processing of signals based on OFTDM technology with using well-localized Weyl-Heisenberg basis. In the article are discussed main advantages of this technology and possibility applications for processing of information in various kind of natural science.

Key words: digital mobile wideband communication systems, digital processing, OFDM system, OFTDM system, Weyl-Heisenberg basis, well localization, orthogonalization, OQAM, well-localized basis.

УСЛОВИЯ ОРТОГОНАЛЬНОСТИ ОБОБЩЕННЫХ БАЗИСОВ ВЕЙЛЯ-ГЕЙЗЕНБЕРГА ДЛЯ OFTDM СИГНАЛОВ

В. П. ВОЛЧКОВ¹⁾

Д. А. ПЕТРОВ²⁾

*¹⁾ Московский
технический
университет связи
и информатики*

*e-mail:
volchkovvalery@mail.ru*

*²⁾ Московский
государственный
университет
им. М.В. Ломоносова*

e-mail: dapetroff@gmail.com

В статье рассматриваются технология передачи информации с ортогональным частотно-временным мультиплексированием (OFTDM). OFTDM сигнал формируется при помощи ортогонального обобщенного базиса Вейля-Гейзенберга, обладающего хорошей частотно-временной локализацией. Для этого базиса доказываются условия ортогональности, критерий и теорема Найквиста, которые гарантируют отсутствие межканальной и межсимвольной интерференции сигнала в гауссовских каналах. На основе этих критериев предложен вычислительно-эффективный метод ортогонализации на основе быстрого преобразования Фурье.

Ключевые слова: OFTDM, критерий Найквиста, ортогонализация.

В настоящее время одной из наиболее распространенных технологий, лежащих в основе передачи информации, как по проводным, так и по беспроводным каналам, является ортогональное частотное мультиплексирование (Orthogonal Frequency Division Multiplexing или OFDM). Структура и свойства OFDM сигнала определяются собственным конечномерным базисом, из которого он формируется, как некоторая линейная комбинация базисных функций. Коэффициенты линейной комбинации (модулирующие символы) могут быть вещественными или комплексными и определяются выбранным сигнальным созвездием (например, QAM, PSK, и др.).

Однако, прямоугольная форма формирующего импульса, характерная для классических OFDM систем [1, 2], не является оптимальной с точки зрения устойчивости к межканальной интерференции (МКИ). Поскольку в этом случае локализация формируемых на основе неё базисных функций в частотной области будет наихудшей. По этой же причине в таких OFDM системах уровень внеполосного излучения оказывается завышенным. Причем, природа межканальных помех такова, что их нельзя скомпенсировать или отфильтровать обычными методами цифровой обработки.

В литературе [3, 4] был предложен ряд подходов к подавлению МКИ. Однако, эти методы приводят либо к весьма значительной потере спектральной и энергетической эффективности, либо к нарушению ортогональности. Поэтому задача борьбы с межканальной интерференцией в мобильных OFDM системах является очень актуальной, и во многих случаях пока не находит удовлетворительного решения.

Одним из основных методов борьбы с частотно-временным рассеянием сигнала является использование специальных ортогональных хорошо локализованных базисов, минимизирующих уровень взаимного влияния поднесущих каналов, как в частотной, так и во временной областях, но при этом не ухудшающих спектральную и энергетическую эффективность системы. Однако, теорема Балиан-Лоу (Balin-Low) [5] ограничивает возможность построения ортогональных плотно упакованных и хорошо локализованных базисов на основе сдвигов по времени и частоте единственного формирующего импульса. Другими словами, в этом случае невозможно построить ортогональный сигнальный базис с хорошей частотно-временной локализацией без потери спектральной эффективности. С другой стороны, ортогональность остается обязательным требованием, приводящим, в частности, к минимизации ошибок, обусловленных действием в канале аддитивного белого гауссовского шума. Альтерна-

тивной технологией, которая позволяет в значительной степени преодолеть указанные недостатки и получить наилучшее частотно-временное уплотнение базисных функций, является ортогональное частотно-временное мультиплексирование – OFTDM (Orthogonal Frequency Time Division Multiplexing) [2].

В OFTDM системах при построении ортогонального сигнального базиса вместо одного формирующего импульса используется несколько инициализирующих функций с хорошей частотно-временной локализацией. Эти функции отличаются друг от друга только фазами, а сам базис получается в результате их равномерных сдвигов по времени и частоте (обобщенный базис Вейля-Гейзенберга).

До последнего времени одной из основных проблем при применении на практике (в частности, в абонентских телекоммуникационных устройствах) таких базисов являлась сложность алгоритма их формирования. В статье выводятся упрощенные условия ортогональности обобщенных базисов Вейля-Гейзенберга для OFTDM сигналов, учитывающие свойства формирующего импульса. В теории связи эти условия часто называются критерием или теоремой Найквиста [1]. Данные критерии имеют важное теоретическое и прикладное значение, так как позволяют получить вычислительно эффективный алгоритм построения обобщенных ортогональных базисов Вейля-Гейзенберга с наилучшей частотно-временной локализацией. Таким образом, упрощается процедура синтеза базиса, что позволяет приблизиться к практической реализации OFTDM системы.

Передаваемый OFTDM сигнал $s(t)$ в дискретном времени можно эквивалентно представить в виде:

$$s[n] = \sum_{k=0}^{M-1} \left(\sum_{l=0}^{L-1} c_{k,l}^R \psi_{k,l}^R[n] + \sum_{l=0}^{L-1} c_{k,l}^I \psi_{k,l}^I[n] \right), n \in J_N, \quad (1)$$

$$\psi_{k,l}^R[n] = g[(n - lM)_N] \exp\left(j \frac{2\pi}{M} k(n - \alpha/2)\right), \quad (2)$$

$$\psi_{k,l}^I[n] = -jg[(n + M/2 - lM)_N] \exp\left(j \frac{2\pi}{M} k(n - \alpha/2)\right), \quad (3)$$

где $c_{k,l}^R = \text{Re}(a_{k,l})$, $c_{k,l}^I = \text{Im}(a_{k,l})$ – действительные и мнимые части комплексных информационных QAM символов $a_{k,l}$; $s[n] = s(nT/M)$, $g[n] = g(nT/M)$, $J_N = \{0, 1, \dots, N-1\}$, $\psi_{k,l}^R[n]$ и $\psi_{k,l}^I[n]$ – комплексные функции, полученные в результате равномерных сдвигов по времени и частоте двух классов инициализирующих функций: $g[n] \exp(j\pi/M k\alpha)$ и $g[n + M/2] \exp(j\pi/M k\alpha)$; $M \geq 2$ – количество поднесущих, $N = ML \geq M$; $F = 1/T$ – расстояние между поднесущими, T – символный временной период; $\alpha \in \mathbb{R} = (-\infty, \infty)$ – фазовый параметр; $(a)_N = (a)_{\text{mod } N}$ – значение числа a по модулю N . Система базисных функций $\mathbf{B}[J_N] = \{\psi_{k,l}^R[n], \psi_{k,l}^I[n]\}$ ортонормирована на дискретном интервале J_N в смысле вещественного скалярного произведения

$$\langle x[n], y[n] \rangle_R = \text{Re}(\langle x[n], y[n] \rangle), \quad \langle x[n], y[n] \rangle = \sum_{n=0}^{N-1} x[n] y^*[n].$$

Важно упростить условия ортогональности базиса $\mathbf{B}[J_N]$ и привести их к виду, удобному для использования на практике.

Пусть OFTDM сигнал $s[n]$ передается по гауссовскому каналу с идеальной частотной характеристикой, а его выход описывается уравнением:

$$r[n] = s[n] + w[n], \quad w[n] \sim \mathbf{N}(0, \sigma_w^2), \quad (4)$$

где $w[n]$ – белый гауссовский шум с нулевым математическим ожиданием и дисперсией σ_w^2 . Оптимальный прием такого сигнала обеспечивается схемой, состоящей из согласованного фильтра с импульсной характеристикой $h_{сф}[n] = s^* [(-n)_N]$ и оператора взятия реальной части комплексного числа $\text{Re}(\cdot)$. Выход данной схемы описывается уравнением:

$$y[n] = \text{Re}(r[n] \otimes h_{сф}[n]) = \text{Re} \left(\sum_{m=0}^{N-1} r[m] s^* [(m-n)_N] \right),$$

где \otimes означает циклическую свертку двух функций.

Отсюда видно, что в момент времени $n = 0$ на выходе оптимального приемника формируется реальная часть корреляционной суммы:

$$y[0] = \text{Re}(\langle r, s \rangle) = \text{Re} \left(\sum_{n=0}^{N-1} r[n] s^* [n] \right) = \text{Re} \left(r[n] \otimes s^* [(-n)_N] \Big|_{n=0} \right).$$

Учитывая вид OFTDM сигнала (1), после преобразований получим:

$$y[0] = \sum_{k=0}^{M-1} \sum_{l=0}^{L-1} (c_{k,l}^R y_{k,l}^R + c_{k,l}^I y_{k,l}^I), \quad (5)$$

$$y_{k,l}^R = \text{Re} \left(\sum_{n=0}^{N-1} r[n] \psi_{k,l}^{*R} [n] \right) = \text{Re} \left(r[n] \otimes \psi_{k,0}^{*R} [(-n)_N] \Big|_{n=lM} \right),$$

$$y_{k,l}^I = \text{Re} \left(\sum_{n=0}^{N-1} r[n] \psi_{k,l}^{*I} [n] \right) = \text{Re} \left(r[n] \otimes \psi_{k,0}^{*I} [(-n)_N] \Big|_{n=lM} \right).$$

реальные части выходов фильтров, согласованных с базисными функциями (2) и (3) соответственно (СФБИ фильтров).

Таким образом, оптимальный прием OFTDM сигнала может быть осуществлен при помощи банка из $2M$ СФБИ фильтров, реальные выходы которых замеряются в моменты времени, кратные M .

Условия накладываемые на структуру сигнала $s[n]$ в гауссовском канале, при которых на выходе оптимального приемника (5) отсутствуют МСИ и МКИ, принято называть критерием Найквиста.

Очевидно, что сложность структуры OFTDM сигналов не позволяет воспользоваться напрямую критерием Найквиста, сформулированным для модели посимвольной передачи [6]. Однако критерий может быть расширен и на такие сигналы.

Нетрудно убедиться, что в случае модели канала (4) межсимвольная и межканальная интерференция будут отсутствовать, если выполнены условия ортогональности для обобщенного базиса Вейля-Гейзенберга, т.е.

$$\langle \psi_{k,l}^R [n], \psi_{k',l'}^R [n] \rangle_R = \delta_{k,k'} \delta_{l,l'}, \quad \langle \psi_{k,l}^R [n], \psi_{k',l'}^I [n] \rangle_R = 0, \quad \langle \psi_{k,l}^I [n], \psi_{k',l'}^I [n] \rangle_R = \delta_{k,k'} \delta_{l,l'}. \quad (6)$$

Действительно, если условия (6) выполнены, то при $r[n] = s[n]$ (отсутствии шума) сигнал на выходе оптимального приемника (5) не содержит интерференционных составляющих и равен:

$$y[0] = \sum_{k=0}^{M-1} \sum_{l=0}^{L-1} (c_{k,l}^R)^2 + (c_{k,l}^I)^2 = E,$$

где E – энергия OFTDM сигнала. Таким образом, условия (6) описывают критерий Найквиста для OFTDM сигнала.

Воспользовавшись свойствами скалярного произведения, нетрудно доказать следующую лемму, позволяющую записать критерий в более компактном виде:

Лемма 1. Необходимыми и достаточными условиями ортогональности базиса $\mathbf{B}[J_N]$ являются равенства:

$$\langle \psi_{k,l}^R[n], \psi_{0,0}^R[n] \rangle_R = \delta_{k,0} \delta_{l,0}, \quad (7)$$

$$\langle \psi_{k,l}^R[n], \psi_{0,0}^I[n] \rangle_R = 0, \quad \forall k \in J_M, \forall l \in J_L. \quad (8)$$

Если учесть конкретный вид базисных функций (2), (3), то лемма 1 позволяет сформулировать критерий Найквиста в следующей форме:

Критерий Найквиста (для OFTDM сигнала). Необходимым и достаточным условием отсутствия МСИ и МКИ в канале (4) с OFTDM сигналом является условие ортогональности базиса $\mathbf{B}[J_N]$, которое сводится к выполнению следующих равенств:

$$\operatorname{Re} \left(\sum_{n=0}^{N-1} g[(n-lM)_N] \exp(j2\pi/M k(n-\alpha/2)) g^*[n] \right) = \delta_{k,0} \delta_{l,0}, \quad \forall k \in J_M, \forall l \in J_L, \quad (9)$$

$$\operatorname{Re} \left(\sum_{n=0}^{N-1} g[(n-lM)_N] \exp(j2\pi/M k(n-\alpha/2)) g^* \left[\left(n + M/2 \right)_N \right] \right) = 0. \quad (10)$$

Рассмотрим частный случай, когда $g[n]$ является сопряженной N -симметричной функцией:

$$g[n] = g^* \left[(-n)_N \right]. \quad (11)$$

Как было показано в работе [7], этому типу симметрии соответствует оптимальное значение фазового параметра $\alpha = M/2$, позволяющее получить наилучшую частотно-временную локализацию ортогонального базиса (5), (6). Причем, отклонение от оптимально выбранного значения α приводит как к ухудшению локализации базисных функций, так и к потере симметрии [8].

При указанных выше условиях равенство (10) выполняется автоматически, а равенство (9) преобразуется к виду

$$p(l, k) = \sum_{n=0}^{N-1} g[n] g[(n-lM)_N] \exp\left(j \frac{2\pi k n}{M/2}\right) = \delta_{l,0} \delta_{k,0}, \quad \forall k \in J_{M/2}, \forall l \in J_L. \quad (12)$$

Определим импульс Найквиста для OFTDM сигнала выражением

$$B(\tau, \nu) = \sum_{n=0}^{N-1} g[n] g^* \left[(n+\tau)_N \right] \exp\left(-j \frac{2\pi}{N} \nu n\right), \quad \tau, \nu \in J_N. \quad (13)$$

В отличие от обычного импульса Найквиста [1], функция $B(\tau, \nu)$ зависит от двух переменных (времени и частоты) и совпадает по смыслу с дискретной функцией неопределенности

$$B(\tau, \nu) = \langle g_o[n, m], g_o[(n+\tau)_N, (m+\nu)_N] \rangle, \quad \tau, \nu \in J_N$$

базового импульса $g_o[n, m] = g[n] \exp\left(j \frac{2\pi}{N} mn\right)$, $n, m \in J_N$.

Из (12) и (13) следует, что

$$B(Ml, 2Lk) = p(l, k) = \delta_{l,0} \delta_{k,0}, \quad (14)$$

т.е. импульс Найквиста равен нулю во всех точках (τ, ν) частотно-временной плоскости, кратных M и $2L$, соответственно, исключая точку $(\tau, \nu) = (0, 0)$, где он отличен от нуля.

Таким образом, при $\alpha = M/2$ и выполнении условия сопряженной N -симметрии (11) критерий Найквиста для OFTDM сигнала упрощается и справедлива следующая теорема.

Теорема Найквиста (для OFTDM сигналов). Если формирующий импульс $g[n]$ базиса OFTDM сигнала обладает свойством сопряженной N -симметрии (11), а фазовый параметр выбран оптимально ($\alpha = M/2$), то необходимым и достаточным условием отсутствия МСИ и МКИ в канале (4) является равенство (14), которое во временной области эквивалентно условию

$$\sum_{r=0}^{2L-1} g \left[\left(n - r \frac{M}{2} \right)_N \right] g \left[\left(n - r \frac{M}{2} - lM \right)_N \right] = \frac{2}{M} \delta_{l,0}, \quad \forall n \in J_N, \quad (15)$$

а в частотной области – условию

$$\sum_{k=0}^{M-1} G \left[(p - kL)_N \right] G^* \left[(p - kL - 2lL)_N \right] = M \delta_{l,0}, \quad \forall p \in J_N, \quad (16)$$

где $G[p] = \sum_{n=0}^{N-1} g[n] \exp(-j 2\pi / N pn)$ – дискретное преобразование Фурье функции $g[n]$.

Доказательство. Введем функцию $f_1'[n] = g \left[(n - lM)_N \right] g \left[\left(n + \frac{M}{2} \right)_N \right]$. Она является симметричной на интервале $n \in [0; lM - M/2]$, т.е.

$$f_1' \left[lM - \frac{M}{2} - n \right] = g \left[-\frac{M}{2} - n \right]_N g \left[lM - n \right]_N = g \left[n + \frac{M}{2} \right]_N g \left[n - lM \right]_N = f_1' [n],$$

а также на интервале $n \in [lM - M/2 + 1; N - 1]$, т.к. для любых $p \in [1; N - 1 - lM + M/2]$

$$\begin{aligned} f_1' [N - p] &= g \left[(N - p - lM)_N \right] g \left[\left(N - p + \frac{M}{2} \right)_N \right] = \\ &= g \left[\left(p + lM - \frac{M}{2} + \frac{M}{2} \right)_N \right] g \left[\left(p + lM - \frac{M}{2} - lM \right)_N \right] = f_1' \left[p + lM - \frac{M}{2} \right] \end{aligned}$$

На этих же интервалах функция $\sin \left(\frac{2\pi}{M} k \left(n - \frac{\pi}{2} \right) \right)$ является антисимметричной:

$$\begin{aligned} \sin \left(\frac{2\pi}{M} k \left(lM - \frac{M}{2} - n - \frac{M}{4} \right) \right) &= \sin \left(\frac{2\pi}{M} k \left(-n - \frac{3M}{4} \right) \right) = -\sin \left(\frac{2\pi}{M} k \left(n - \frac{M}{4} \right) \right), \\ \sin \left(\frac{2\pi}{M} k \left(lM - \frac{M}{2} + p - \frac{M}{4} \right) \right) &= \sin \left(\frac{2\pi}{M} k \left(p + \frac{M}{4} \right) \right) = -\sin \left(\frac{2\pi}{M} k \left(N - p - \frac{M}{4} \right) \right). \end{aligned}$$

Поэтому условие (8) ортогональности базиса k , которое можно записать в виде

$$\begin{aligned} \operatorname{Re} \left\{ \sum_{n=0}^{N-1} j f_1' [n] \exp \left(j \frac{2\pi}{M} k \left(n - \frac{\alpha}{2} \right) \right) \right\} &= -\sum_{n=0}^{N-1} f_1' [n] \sin \left(\frac{2\pi}{M} k \left(n - \frac{M}{4} \right) \right) = \\ &= -\left\{ \sum_{n=0}^{lM - M/2} f_1' [n] \sin \left(\frac{2\pi}{M} k \left(n - \frac{M}{4} \right) \right) \right\} - \left\{ \sum_{n=lM - M/2 + 1}^{N-1} f_1' [n] \sin \left(\frac{2\pi}{M} k \left(n - \frac{M}{4} \right) \right) \right\} = 0, \end{aligned}$$

выполняется автоматически для любых $k \in J_M$ и любых $l \in J_M$, т.к. обе суммы произведений симметричной и антисимметричной функции обращаются в ноль.

Отсюда непосредственно следует, что необходимым и достаточным условием ортогональности базиса $\mathbf{B}[J_N]$ является только условие (7).

Определим следующую вспомогательную функцию

$$f_2'[n] = g \left[(n - lM)_N \right] g \left[(n)_N \right]. \quad (17)$$

Она является N -периодической и симметричной на интервале $n \in [0; lM]$, т.к.

$$f_2' [lM - n] = g \left[(-n + lM)_N \right] g \left[(-n)_N \right] = g \left[(n - lM)_N \right] g [n] = f_2' [n], \quad (18)$$

а также на интервале $n \in [lM + 1; N - 1]$, т.к. для любого $p \in [1; N - lM - 1]$

$$f_2' [p + lM] = g \left[(p + lM)_N \right] g \left[(p)_N \right] = g \left[(-p - lM)_N \right] g \left[(-p)_N \right] = f_2' [N - p]. \quad (19)$$

Причем, на этих же интервалах симметричность функции $\cos \left(\frac{2\pi}{M} k \left(n - \frac{M}{4} \right) \right)$ зависит от четности k :

$$\begin{aligned} \cos \left(\frac{2\pi}{M} k \left(lM - n - \frac{M}{4} \right) \right) &= \cos \left(\frac{2\pi}{M} k \left(n - \frac{M}{4} \right) + \pi k \right) = (-1)^k \cos \left(\frac{2\pi}{M} k \left(n - \frac{M}{4} \right) \right), \\ \cos \left(\frac{2\pi}{M} k \left(lM + p - \frac{M}{4} \right) \right) &= \cos \left(\frac{2\pi}{M} k \left(N - p - \frac{M}{4} \right) + \pi k \right) = (-1)^k \cos \left(\frac{2\pi}{M} k \left(N - p - \frac{M}{4} \right) \right), \end{aligned}$$

Условие (7), которое можно записать в виде

$$\begin{aligned} \operatorname{Re} \left\{ \sum_{n=0}^{N-1} f_2^l[n] \exp \left(j \frac{2\pi}{M} k (n - \alpha/2) \right) \right\} &= \sum_{n=0}^{N-1} f_2^l[n] \cos \left(\frac{2\pi}{M} k (n - M/4) \right) = \\ &= \sum_{n=0}^{Ml} f_2^l[n] \cos \left(\frac{2\pi}{M} k (n - M/4) \right) + \sum_{n=Ml+1}^{N-1} f_2^l[n] \cos \left(\frac{2\pi}{M} k (n - M/4) \right) = \delta_{l,0} \delta_{k,0}, \end{aligned}$$

выполняется автоматически для нечетных k , т.к. обе суммы в левой части равенства, очевидно, обращаются в 0.

Остается рассмотреть случай, когда k является четным, т.е. $k = 2m$, $m \in J_{M/2}$.

Дискретное преобразование Фурье функции $f_2^l[n]$ имеет вид

$$F_2^l[v] = \sum_{n=0}^{N-1} f_2^l[n] \exp \left(-j \frac{2\pi}{N} vn \right), v \in J_N, \quad (20)$$

следовательно

$$F_2^l[2Lm] = \sum_{n=0}^{N-1} f_2^l[n] \exp \left(-j \frac{2\pi}{N} 2Lmn \right). \quad (21)$$

Учитывая свойства симметрии (18) и (19) функции $f_2^l[n]$ можно показать, что:

$$F_2^l[-2Lm] = F_2^l[2Lm]. \quad (22)$$

Действительно,

$$\begin{aligned} F_2^l[-2Lm] &= \sum_{n=0}^{N-1} f_2^l[n] \exp \left(-j \frac{2\pi}{N} (-2Lm)n \right) = \\ &= \sum_{n=0}^{lM} f_2^l[lM - n] \exp \left(j \frac{2\pi}{N} 2Lmn \right) + \sum_{n=lM+1}^{N-1} f_2^l[N - n + lM] \exp \left(j \frac{2\pi}{N} 2Lmn \right) = \\ &= \sum_{n'=0}^{lM} f_2^l[n'] \exp \left(-j \frac{2\pi}{N} 2Lmn' \right) \exp \left(j \frac{2\pi}{N} 2LmlM \right) + \\ &+ \sum_{n''=lM+1}^{N-1} f_2^l[n''] \exp \left(-j \frac{2\pi}{N} 2Lmn'' \right) \exp \left(j \frac{2\pi}{N} 2Lm(lM + N) \right) = \\ &= \sum_{n=0}^{N-1} f_2^l[n] \exp \left(-j \frac{2\pi}{N} 2Lmn \right) = F_2^l[2Lm], \end{aligned}$$

где была использована замена переменных: $n' = lM - n$, $n'' = N - n + lM$.

Для четных $k = 2m$ левая часть условия (7) принимает вид

$$\begin{aligned} \operatorname{Re} \left\{ \sum_{n=0}^{N-1} f_2^l[n] \exp \left(j \frac{2\pi}{M} 2m (n - M/4) \right) \right\} &= (-1)^m \sum_{n=0}^{N-1} f_2^l[n] \cos \left(\frac{2\pi}{M} 2mn \right) = \\ &= \frac{(-1)^m}{2} \left(\sum_{n=0}^{N-1} f_2^l[n] \exp \left(j \frac{2\pi}{M} 2mn \right) + \sum_{n=0}^{N-1} f_2^l[n] \exp \left(-j \frac{2\pi}{M} 2mn \right) \right) = \\ &= \frac{(-1)^m}{2} (F_2^l[2Lm] + F_2^l[-2Lm]) \end{aligned}$$

Учёт равенства (22) приводит к следующему условию ортогональности базиса $\mathbf{B}[J_N]$, являющегося аналогом соотношения (7):

$$F_2^l[2Lm] = F_2^l[-2Lm] = \delta_{l,0} \delta_{m,0}, \forall m \in J_{M/2}; \forall l \in J_L. \quad (23)$$

Принимая во внимание вид функции $f_2^l[n]$ (17) и соотношение (21) для ее преобразования Фурье, получим из равенства (23) следующую форму необходимого и достаточно условия ортогональности базиса Вейля-Гейзенберга:

$$\sum_{n=0}^{N-1} g[n] g[(n - lM)_N] \exp \left(-j \frac{2\pi mn}{M/2} \right) = \delta_{l,0} \delta_{m,0}, \forall m \in J_{M/2}, \forall l \in J_L. \quad (24)$$

Докажем сначала необходимость условия (15) во временной области.

Учитывая выражение (20) для дискретного преобразования Фурье функции $f_2^l[n]$, она может быть представлена в виде

$$f_2^l[n] = \frac{1}{N} \sum_{v=0}^{N-1} F_2^l[v] \exp\left(j \frac{2\pi}{N} nv\right).$$

Поэтому, для любого $r \in J_{2L}$

$$f_2^l\left[n - r \frac{M}{2}\right] = \frac{1}{N} \sum_{v=0}^{N-1} F_2^l[v] \exp\left(j \frac{2\pi}{N} \left(n - r \frac{M}{2}\right)v\right).$$

Произведем в последнем выражении суммирование по r от 0 до $2L-1$ и учтем явный вид функции $f_2^l[n]$ (17)

$$\begin{aligned} \sum_{r=0}^{2L-1} f_2^l\left[n - r \frac{M}{2}\right] &= \sum_{r=0}^{2L-1} g\left[\left(n - r \frac{M}{2}\right)_N\right] g\left[\left(n - r \frac{M}{2} - lM\right)_N\right] = \\ &= \frac{1}{N} \sum_{v=0}^{N-1} \exp\left(j \frac{2\pi}{N} vn\right) F_2^l[v] \sum_{r=0}^{2L-1} \exp\left(j \frac{2\pi}{2L} r(-v)\right). \end{aligned}$$

Заметим, что

$$\frac{1}{2L} \sum_{r=0}^{2L-1} \exp\left(j \frac{2\pi}{2L} r(-v)\right) = \begin{cases} 1, & \text{при } v \text{ кратных } 2L \\ 0, & \text{для остальных } v \end{cases},$$

а так как $v \in J_N$, то можно заключить, что $\sum_{r=0}^{2L-1} \exp\left(j \frac{2\pi}{2L} r(-v)\right)$ отлична от 0 и равняется $2L$ только при $v = 2Lm$, $m \in J_{M/2}$, поэтому

$$\sum_{r=0}^{2L-1} g\left[\left(n - r \frac{M}{2}\right)_N\right] g\left[\left(n - r \frac{M}{2} - lM\right)_N\right] = \frac{2}{M} \sum_{m=0}^{M/2-1} \exp\left(j \frac{4\pi}{M} mn\right) F_2^l[2Lm] \quad (25)$$

Очевидно, что в случае, когда выполняется условие (23), выполняется и равенство (15), которое является необходимым для ортогональности базиса Вейля-Гейзенберга.

Для доказательства достаточности этого условия, заметим, что если оно выполнено, то из соотношения (25) следует:

$$\begin{aligned} \frac{2}{M} \sum_{m=0}^{M/2-1} \exp\left(j \frac{4\pi}{M} mn\right) F_2^0[2Lm] &= 1, \forall m \in J_{M/2}. \\ F_2^l[2Lm] &= 0, \forall l \in \{1, 2, \dots, L-1\}, \forall m \in J_{M/2}, \end{aligned}$$

Причем, последнее равенство выполняется только в том случае, когда

$$F_2^0[2Lm] = \delta_{m,0}, \forall m \in J_{M/2},$$

следовательно, полностью выполняется необходимое и достаточное условие ортогональности базиса в виде (23).

Для доказательства необходимости и достаточности условия ортогональности (16) в частотной области, отметим, что условие (24) можно записать в виде

$$\sum_{n=0}^{N-1} g[n] \left(g\left[\left(n - lM\right)_N\right] \exp\left(j \frac{2\pi}{N} 2Lmn\right) \right)^* = \delta_{l,0} \delta_{m,0}, \forall l \in J_L, \forall m \in J_{M/2}.$$

Применив равенство Парсеваля, получим

$$\sum_{\kappa=0}^{N-1} G[\kappa] G^* \left[\left(\kappa - 2Lm\right)_N \right] \exp\left(-j \frac{2\pi}{L} l\kappa\right) = \delta_{l,0} \delta_{m,0}, \quad (26)$$

где $G[\kappa]$ – дискретное преобразование Фурье функции $g[n]$.

Левую часть условия (26) представим в виде разложение в ряд Фурье некоторой L -периодической функции $\Gamma_m[p] = \sum_{k=0}^{M-1} G[(p+kL)_N] G^*[(p+kL-2Lm)_N]$, $p \in J_L$, поэтому

$$\sum_{p=0}^{L-1} \exp\left(-j \frac{2\pi}{L} pl\right) \Gamma_m[p] = \delta_{l,0} \delta_{m,0}.$$

Очевидно, что при $m \neq 0$ это равенство выполняется только в том случае, если $\Gamma_m[p] = 0, \forall p \in J_L$, а при $m = 0$, если $\Gamma_m[p] = 1/L, \forall p \in J_L$. Таким образом, условие

$$\Gamma_m[p] = \frac{1}{L} \delta_{l,0}, \forall p \in J_L$$

представляет собой необходимое и достаточное условие ортогональности базиса $\mathbf{B}[J_N]$.

Если учесть вид функции $\Gamma_m[p]$ и ее L -периодичность, то мы получим искомый вид условия ортогональности (10) в частотной области для любых $p \in J_N$.

Теорема полностью доказана.

Форма базиса Вейля-Гейзенберга $\mathbf{B}[J_N]$ фактически определяется инициализирующей функцией $g[n]$, поэтому можно упростить алгоритм его синтеза, получив специальные критерии ортогональности, аналогичные (15), в виде условий на эту функцию. Для этого введем базис Винера:

Определение 1 [9]. Совокупность функций f_0, f_1, \dots, f_{N-1} линейного подпространства $V \subset \tilde{\mathcal{C}}^N$ ($\tilde{\mathcal{C}}^N$ – пространство N -периодических комплексных функций целочисленного аргумента) называется базисом Винера подпространства V , если для нее справедлив дискретно-периодический аналог теоремы Винера: функции $\{g[\cdot - kN/J]\}_{k=0}^{J-1}$ образуют базис пространства V в том и только том случае, когда $g \in \tilde{\mathcal{C}}^N$ и представима в виде $g = \sum_{k=0}^{J-1} a_k f_k$, где все a_k отличны от нуля.

Таким образом, существование базиса Винера необходимо и достаточно для существования функции g , чей вектор сдвигов

$$\vec{g}[n] = (g[n], g[(n-N/J)_N], \dots, g[(n-(J-1)N/J)_N])^T$$

является базисом подпространства V .

Определение 2. Преобразование, позволяющее перейти от вектора сдвигов к ортогональному базису Винера, называется преобразованием Винера и может быть записано в виде:

$$\eta_k[n] = \sum_{r=0}^{2L-1} g\left[\left(n - r \frac{N}{2L}\right)_N\right] \exp\left(\frac{2\pi j}{2L} rk\right). \quad (27)$$

Обратное преобразование Винера определим следующим образом:

$$g[n] = \frac{1}{2L} \sum_{k=0}^{J-1} \eta_k[n]. \quad (28)$$

Используя (15) и вид преобразования Винера (27), можно показать, что необходимым и достаточным условием ортонормированности обобщенного базиса Вейля-Гейзенберга $\mathbf{B}[J_N]$ (а следовательно, отсутствия МКИ и МСИ) во временной области является следующее равенство

$$|\eta_k^{M/2}[n]|^2 + |\eta_{k+L}^{M/2}[n]|^2 = 4/M. \quad (29)$$

Этот критерий позволяет сформировать эффективную процедуру ортогонализации.

Действительно, возьмем некоторую функцию $g_o[n]$, $n \in J_N$, являющуюся сопряженной N -симметричной. Для получения хорошей локализации ортогонального базиса в качестве $g_o[n]$ следует выбрать функцию Гаусса.

Построим базис Винера в следующем виде:

$$\eta_k^{M/2}[n] = \frac{2\tilde{\eta}_k^{M/2}[n]}{\sqrt{M|\tilde{\eta}_k^{M/2}[n]|^2 + M|\tilde{\eta}_{k-L}^{M/2}[n]|^2}}, \quad (30)$$

где $\tilde{\eta}_k^{M/2}[n] = \sum_{r=0}^{2L-1} g_o\left[\left(n - r \frac{M}{2}\right)_N\right] \exp\left(\frac{2\pi j}{2L} rk\right)$.

Нетрудно проверить, что при непосредственной подстановке выражения (30) в критерий ортогональности (29) мы получим равенство. Отсюда следует, что базис Вейля-Гейзенберга $\mathbf{B}[J_N]$, построенный на основе функции $g[n]$, получаемой обратным преобразованием Винера из базиса $\eta_k^{M/2}[n]$, является ортогональным.

Заметим, что получение элементов базиса Винера (27) на основе функции $g_o[n]$ фактически представляет собой дискретное преобразование Фурье (ДПФ), для реализации которого разработаны специальные быстрые алгоритмы.

Таким образом, теорема Найквиста является не только удобной формой представления условий отсутствия МКИ и МСИ OFTDM сигналов, но является важным шагом в формировании вычислительно эффективного алгоритма синтеза сигнального базиса, обладающего высоким уровнем частотно-временной локализации.

Литература

1. Прокис, Дж. Цифровая связь [Текст]: пер. с англ. / Дж. Прокис; Под ред. Д.Д. Кловского. – Москва: Радио и связь, 2000.
2. Волчков, В. П. Сигнальные базисы с хорошей частотно-временной локализацией [Текст] / В. П. Волчков // Журнал «Электросвязь». – 2007. – №2. – С. 21-25.
3. Haas, R. A time-frequency well-localized pulse for multiple carrier transmission [Текст] / R. Haas, J.C. Belfiore // Wireless Personal Communications. – 1997. – vol. 5. – pp. 1-18.
4. Muschallik, C. Improving an OFDM reception using an adaptive Nyquist windowing [Текст] // IEEE Transactions on Consumer Electronics. – 1996. – vol. 42. – no. 3. – pp. 259 – 269.
5. Mallet, S. A Wavelet Tour of Signal Processing [Текст]: Second Edition / S. Mallet. – Academic Press. – 1999.
6. Волчков, В.П., Петров Д.А. Обобщенная теорема Найквиста для OFTDM сигналов [Текст] / В.П. Волчков, Д.А. Петров // Сборник докладов. Всероссийский научно-технический семинар «Системы синхронизации формирования и обработки сигналов для связи и вещания». – 2009. – Воронеж. – С. 28-32.
7. Волчков, В.П. Оптимизация ортогонального базиса Вейля-Гейзенберга для цифровых систем связи, использующих принцип OFDM/OQAM передачи [Текст] / В.П. Волчков, Д.А. Петров // Научные ведомости БелГУ. Серия «История. Политология. Экономика. Информатика». – 2009. – №1(56). – Вып. 9/1. – С.102-110.
8. Volchkov, V.P. Orthogonal Well-Localized Weyl-Heisenberg Basis Construction and Optimization for Multicarrier Digital Communication Systems [Электронный ресурс] / V.P. Volchkov, D. A. Petrov // International Conference on Ultra Modern Telecommunications (ICUMT 2009). – Oct 12-14, 2009. – St. Petersburg, Russia. – ISBN: 978-1-4244-3941-6. – IEEE Catalog Number: CFP0963G-CDR (<http://ieeexplore.ieee.org>)
9. Петухов А.П. Периодические дискретные всплески [Текст] / А.П. Петухов // Алгебра и анализ. – 1996. – Вып. 8. – №3. – С. 151-183.



GENERALIZED WEYL-HEISENBERG BASES ORTHOGONALITY CONDITIONS FOR OFTDM SIGNALS

V. P. VOLCHKOV¹⁾

D. A. PETROV²⁾

*¹⁾ Moscow technical
university of
telecommunications
and informatics*

e-mail:

volchkovvalery@mail.ru

²⁾ M. V. Lomonosov

Moscow state university

e-mail: dapetroff@gmail.com

Article considers information transmission technology with orthogonal time-frequency multiplexing (OFTDM). OFTDM signal is formed by generalized well-localized Weyl-Heisenberg basis. Special orthogonality conditions are derived for this basis in the form of Nyquist criterion and theorem. These conditions guarantee the absence of intercarrier and intersymbol interference in Gaussian channels. In addition they allow to formulate computationally efficient orthogonalization algorithm on the base of fast Fourier transform.

Key words: OFTDM, Nyquist theorem, orthogonalization.

ОБ ЭФФЕКТИВНОСТИ МЕТОДА ОПТИМАЛЬНОЙ ФИЛЬТРАЦИИ ИЗОБРАЖЕНИЙ

Е. Г. ЖИЛЯКОВ
А. А. ЧЕРНОМОРЕЦ
В. А. ГОЛОЩАПОВА

*Белгородский
государственный
университет*

e-mail: Zhilyakov@bsu.edu.ru

В работе представлены результаты экспериментальных исследований эффективности метода оптимальной фильтрации изображений на основе частотных представлений в заданном частотном субинтервале на примере модельных изображений.

Ключевые слова: оптимальная фильтрация, изображение, трансформанта Фурье, субинтервал.

Введение

Актуальность выполнения исследований обусловлена тем, что одно из направлений интенсивного развития современных информационно-телекоммуникационных систем определяется тенденцией использования наиболее естественных для человека форм информационного обмена, и, прежде всего, в виде визуальных данных. Обработка изображений является важной задачей. Существующие методы выделения (фильтрации) значимых характеристик отдельных компонент изображения, некоторых периодических структур изображений не являются оптимальными с точки зрения аппроксимации трансформант Фурье в заданных частотных интервалах, в которых осуществляется фильтрация. Поэтому актуальной проблемой является создание математических моделей и методов фильтрации, позволяющих адекватно учитывать энергетические характеристики изображений в выбранных частотных интервалах. В работе [1] разработан и теоретически обоснован новый метод оптимальной линейной фильтрации изображений на основе частотных представлений, который является оптимальным в том смысле, что спектр получаемого в результате фильтрации изображения имеет наименьшее среднеквадратическое отклонение от спектра фильтруемого изображения в заданном двумерном частотном субинтервале, а вне этого субинтервала имеет наименьшее отклонение от нуля. В данной работе изложены результаты экспериментальных исследований эффективности указанного метода. В основе исследований лежит метод вычислительных экспериментов с использованием модельных данных.

Теоретические основы метода оптимальной фильтрации изображений на основе частотных представлений

О частотных представлениях при обработке изображений [2] можно говорить, поскольку изображение, как функция с конечной или неограниченной областью определения, может быть представлено в виде суммы синусов и косинусов различных частот, умноженных на некоторые весовые коэффициенты

$$f_{ik} = \frac{1}{4\pi^2} \sum_{u=1}^M \sum_{v=1}^N F(u, v) e^{ju(i-1)} e^{jv(k-1)}, \quad i = 1, 2, \dots, M, \quad k = 1, 2, \dots, N, \quad (1)$$

где $F(u, v)$ – значение трансформанты Фурье функции

$$F(u, v) = \sum_{i=1}^M \sum_{k=1}^N f_{ik} e^{-ju(i-1)} e^{-jv(k-1)}, \quad u = 1, 2, \dots, M, \quad v = 1, 2, \dots, N. \quad (2)$$

Метод оптимальной фильтрации изображений на основе частотных представлений [1,3] позволяет для нахождения результатов Y_{Ω} фильтрации изображения $\Phi = (f_{ik}), i=1, 2, \dots, M, k=1, 2, \dots, N$, в частотной двумерной области Ω построить вычислительную процедуру, не вычисляя при этом трансформанту Фурье. Для фильтрации изображений предложено использовать следующий вариационный принцип: транс-



форманта Фурье $Z(u, v)$ результата фильтрации Y_Ω , является оптимальной в смысле евклидовой нормы ее отклонения в заданном частотном интервале Ω от трансформанты Фурье $F(u, v)$ исходного изображения Φ и от нуля – вне данного интервала, т.е.

$$\iint_{(u,v) \in \Omega} |F(u, v) - Z(u, v)|^2 dudv + \iint_{(u,v) \notin \Omega} |Z(u, v)|^2 dudv \Rightarrow \min. \quad (3)$$

Указанному вариационному принципу соответствует следующее изображение Y_Ω

$$Y_\Omega = A^T \cdot \Phi \cdot B, \quad (4)$$

где матрицы $A=(a_{i1i2})$, $i_1, i_2=1, 2, \dots, M$, и $B=(b_{k1k2})$, $k_1, k_2=1, 2, \dots, N$ (субполосные матрицы [2]), определяются соотношениями

$$a_{i_1 i_2} = \begin{cases} \frac{\sin(\alpha_2(i_1 - i_2)) - \sin(\alpha_1(i_1 - i_2))}{\pi(i_1 - i_2)}, & i_1 \neq i_2, \\ \frac{\alpha_2 - \alpha_1}{\pi}, & i_1 = i_2, \end{cases} \quad b_{k_1 k_2} = \begin{cases} \frac{\sin(\beta_2(k_1 - k_2)) - \sin(\beta_1(k_1 - k_2))}{\pi(k_1 - k_2)}, & k_1 \neq k_2, \\ \frac{\beta_2 - \beta_1}{\pi}, & k_1 = k_2. \end{cases} \quad (5)$$

Значения

$$0 \leq \alpha_1, \alpha_2, \beta_1, \beta_2 \leq \pi \quad (6)$$

задают границы частотного субинтервала Ω .

Важным свойством определенным таким образом результата фильтрации является то, что указанный метод не допускает растекания энергии двумерного сигнала за пределы заданной частотной области.

Концептуальные основы экспериментальных исследований

В ходе вычислительных экспериментов проводился сравнительный анализ оценки доли энергии изображения Y_{Opt} , полученного в результате оптимальной фильтрации,

$$\delta_{\Omega Opt} = \frac{\|Y_{Opt}\|^2 - P_\Omega(Y_{Opt})}{\|Y_{Opt}\|^2} = 1 - \frac{P_\Omega(Y_{Opt})}{\|Y_{Opt}\|^2}, \quad (7)$$

где $\|Y_{Opt}\|^2$ – энергия изображения Y_{Opt} в области определения трансформанты Фурье, $P_\Omega(Y_{Opt})$ – доля энергии изображения Y_{Opt} в частотном субинтервале Ω , определяемая на основании соотношения

$$P_\Omega(Y_{Opt}) = \text{tr}ec(A_\Omega^T Y_{Opt} B_\Omega Y_{Opt}^T). \quad (8)$$

с соответствующими характеристиками изображений, полученных при реализации известных частотных фильтров [4]:

- идеальный фильтр (результат фильтрации Y_{Ideal}),

$$H_{Ideal}(u, v) = \begin{cases} 1, & u \in [-\pi, \pi], \quad v \in [-\beta_2, -\beta_1] \cup [\beta_1, \beta_2], \\ 0, & \text{в противном случае.} \end{cases} \quad (9)$$

- фильтр Баттерворта порядка $n=2$ (результат фильтрации $Y_{Butterw}$),

$$H_{Butterw}(u, v) = \frac{1}{1 + \left(\frac{D_1(u, v)}{D_0}\right)^{2n} + \left(\frac{D_2(u, v)}{D_0}\right)^{2n}}, \quad (10)$$

где D_0 – частота среза, $D_1(u, v)$, $D_2(u, v)$ – расстояние между точкой (u, v) и осями симметрии l_1, l_2 фильтра,

- Гауссов фильтр (результат фильтрации $Y_{Gaussian}$),

$$H_{Gaussian}(u, v) = \exp\left(-\frac{D_1^2(u, v)}{2D_0^2}\right) + \exp\left(-\frac{D_2^2(u, v)}{2D_0^2}\right), \quad (11)$$

где D_0 – частота среза, $D_1(u, v)$, $D_2(u, v)$ – расстояния между точкой (u, v) и осями симметрии l_1, l_2 фильтра.

Построение фильтров осуществлено на основании параметров частотного субинтервала, в котором выполнялась оптимальная фильтрация.

Соответствующие оценки доли энергии результатов фильтрации (изображения Y_{Ideal} , $Y_{Butterw}$, $Y_{Gaussian}$), которая сосредоточена вне выбранного частотного субинтервала Ω , имеют следующий вид

$$\delta_{\Omega Ideal} = 1 - \frac{P_{\Omega}(Y_{Ideal})}{\|Y_{Ideal}\|^2}, \quad (12)$$

$$\delta_{\Omega Butterw} = 1 - \frac{P_{\Omega}(Y_{Butterw})}{\|Y_{Butterw}\|^2}, \quad (13)$$

$$\delta_{\Omega Gaussian} = 1 - \frac{P_{\Omega}(Y_{Gaussian})}{\|Y_{Gaussian}\|^2}. \quad (14)$$

Отметим, что вычисления как относительной доли «просачивания» энергии за пределы частотного субинтервала оказались доступными, благодаря разработанному в [1] методу вычисления долей энергии (8).

Для получения большей наглядности при оценивании различий оптимальной фильтрации и фильтрации с помощью известных фильтров представляется естественным построить графики соответствующих характеристик выходных последовательностей фильтров.

Сравнительный анализ в работе выполнен, исследуя параметры фильтров и результаты их применения в различных подобластях двумерной области (подобласти получены при фиксированных значениях одной из независимых переменных частотного пространства). Однако, данный факт не снижает важности полученных результатов в виду известного свойства преобразования Фурье (преобразование Фурье можно записать так, что переменные в нем оказываются разделенными).

Результаты вычислительных экспериментов

Экспериментальные исследования целесообразно осуществлять на основе обработки модельных сигналов. Генерирование значений модельного изображения (рис. 1) $\Phi = (f_{ik})$, $i = 1, 2, \dots, 512$, $k = 1, 2, \dots, 512$, размерностью 512×512 пикселей, осуществляется на основе соотношения

$$f_{ik} = a_1 \sin(2\pi f_1 k) + a_2 \sin(2\pi f_2 k) + a_3 \sin(2\pi f_3 k) + a_4 \sin(2\pi f_4 k) + a_5 \sin(2\pi f_5 k), \quad (15)$$

$$i = 1, 2, \dots, 512, \quad k = 1, 2, \dots, 512,$$

где

$$f_1 = 0.0193; \quad f_2 = 0.0208; \quad f_3 = 0.0279, \quad f_4 = 0.0295, \quad f_5 = 0.0311,$$

$$a_1 = 3.2; \quad a_2 = 2.9; \quad a_3 = 3.5, \quad a_4 = 2.1, \quad a_5 = 2.2.$$

Такие значения частот позволяют осуществить исследование различных аспектов фильтрации и, прежде всего, влияния энергии исходного сигнала за пределами частотного субинтервала на ее результаты.

Фильтрация модельного сигнала (15) на основании метода оптимальной фильтрации изображений на основе частотных представлений осуществлялась в следующих частотных субинтервалах:

$$\Omega_i : \{\alpha_1 = 0, \quad \alpha_2 = \pi, \quad \beta_1 = \beta_{01} + \Delta\beta(i-1), \quad \beta_2 = \beta_{02} + \Delta\beta(i-1)\}, \quad i=1, 2, \dots, 15, \quad (16)$$

где $\Delta\beta$ – величина смещения границ очередного частотного субинтервала.

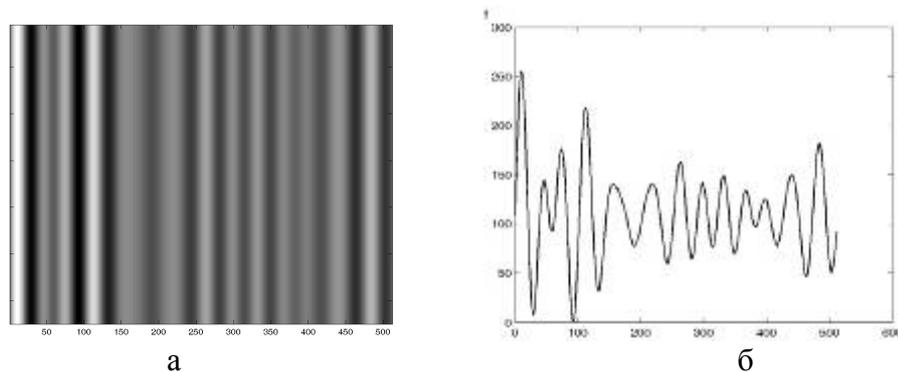


Рис. 1. Модельное изображение:
а – в виде изображения, б – профиль

В ходе экспериментов были использованы следующие значения границ частотного субинтервала

$$\beta_{01} = 0.035\pi, \beta_{02} = 0.04\pi, \Delta\beta = 0.0025\pi. \tag{17}$$

В табл. 1 представлены значения долей энергии выходных последовательностей оптимального и других анализируемых фильтров, находящихся вне заданного частотного интервала, вычисленные согласно (7), (12)-(14).

Таблица 1

Оценка долей энергии результатов фильтрации вне частотного субинтервала Ω

№	Границы частотного субинтервала Ω		$\delta_{\Omega Opt}$	$\delta_{\Omega Ideal}$	$\delta_{\Omega Butterw}$	$\delta_{\Omega Gaussian}$
	β_1	β_2				
1	0.035π	0.04π	0.13211	0.28323	0.27313	0.27189
2	0.0375π	0.0425π	0.3492	0.30822	0.31119	0.31036
3	0.04π	0.045π	0.13101	0.55406	0.55503	0.55472
4	0.0425π	0.0475π	0.34099	0.72638	0.55996	0.60419
5	0.045π	0.05π	0.41356	0.86981	0.87022	0.87001
6	0.0475π	0.0525π	0.23196	0.56816	0.57263	0.57082
7	0.05π	0.055π	0.34246	0.58755	0.6305	0.62484
8	0.0525π	0.0575π	0.12423	0.22294	0.22436	0.22396
9	0.055π	0.06π	0.33214	0.63152	0.6312	0.63129
10	0.0575π	0.0625π	0.40426	0.85964	0.7762	0.8095
11	0.06π	0.065π	0.1968	0.41986	0.4204	0.42016
12	0.0625π	0.0675π	0.24863	0.57319	0.57095	0.57158
13	0.065π	0.07π	0.4328	0.92163	0.92157	0.92155
14	0.0675π	0.0725π	0.47163	0.69156	0.69313	0.69208
15	0.07π	0.075π	0.20767	0.7969	0.68474	0.7815

Легко видеть, что относительные доли «просачивания» энергии за пределы частотного субинтервала у выходных последовательностей идеального, Гауссова фильтров и фильтра Баттерворта существенно больше.

Как показано в ходе экспериментов выходная последовательность оптимального фильтра определяется только трансформантой Фурье в выбранном частотном субинтервале. В отличие от этого, на выходную последовательность идеального, Гауссова фильтров и фильтра Баттерворта влияет энергия входного изображения, сосредоточенная вне выбранного частотного интервала. Это различие легко обнаружить на рис. 2, на котором представлены абсолютные значения трансформант Фурье выходных последовательностей оптимального и других фильтров.

Далее представлены результаты фильтрации модельного изображения (рис. 1) в частотном субинтервале

$$\Omega: \{\alpha_1 = 0, \alpha_2 = \pi, \beta_1 = 0.04\pi, \beta_2 = 0.045\pi\}.$$

На рис. 2 приведены абсолютные значения трансформант Фурье при $u=0.22\pi$, $v \in [0.03\pi, 0.075\pi]$ исходного изображения и результатов фильтрации, полученных при применении оптимального и других фильтров в заданном частотном субинтервале Ω .

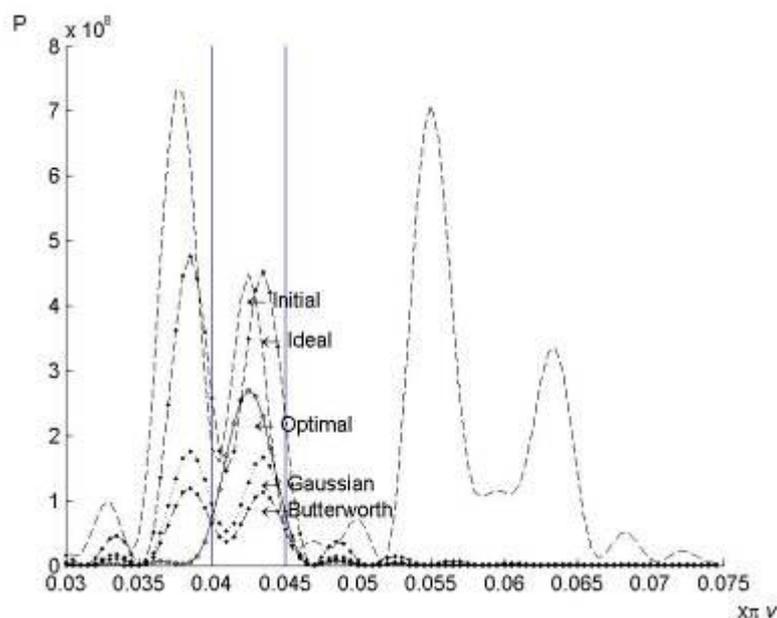


Рис. 2. Трансформанты Фурье (абсолютные значения) в частотном субинтервале

$\alpha_1=0, \alpha_2=\pi, \beta_1 = 0.04\pi, \beta_2 = 0.045\pi$ при $u=0.22\pi$. ----- – исходного изображения; выходных последовательностей фильтров:

—○— – оптимальный, -▲- - идеальный, -◆- - Баттерворта, -◇- - Гауссов,

На рис. 3 влияние дополнительной энергии исходного изображения, сосредоточенной вне выбранного частотного субинтервала у его левой границы (рис. 2), проявляется в наличии существенных изменений яркости пикселей изображений (рис. 3в, 3д, 3ж), являющихся результатами фильтрации с помощью фильтров, отличных от оптимального фильтра, и в наличии биений на графиках соответствующих профилей изображений.

Выводы

Результаты вычислительных экспериментов как с модельными изображениями, так и с реальными фотоснимками в цифровом виде показывают, что относительные доли «просачивания» энергии за пределы указанного частотного субинтервала у выходных последовательностей частотных фильтров, основанных на применении ДПФ, существенно больше, чем у разработанного метода.

В ходе экспериментов показано, что энергия входного изображения, сосредоточенная вне выбранного частотного интервала, влияет на выходную частотных фильтров, основанных на применении ДПФ, в значительно большей степени, чем на результат, полученный на основании метода оптимальной фильтрации. Данное влияние проявляется как на графиках значений соответствующих трансформант Фурье, так и на изображениях, полученных в результате фильтрации.

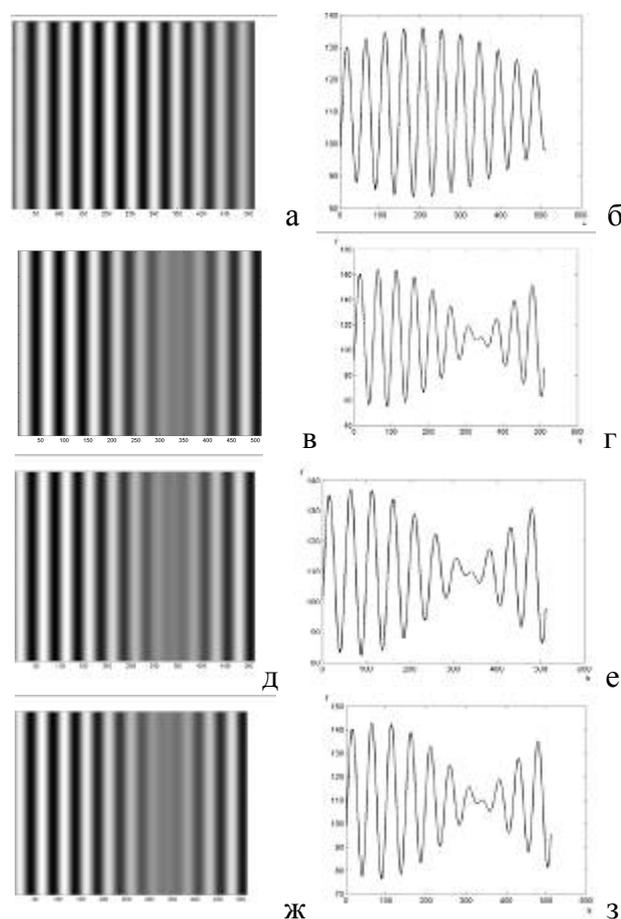


Рис. 3. Результаты фильтрации (изображение и профиль) на основании фильтров:
а, б – оптимальный, в, г – идеальный, д, е – Баттворта, ж, з – Гауссов

Литература

1. Жилияков, Е.Г. Вариационные алгоритмы анализа и обработки изображений на основе частотных представлений [Текст] / Е.Г. Жилияков, А.А. Черноморец. – Белгород: Изд-во БелГУ, 2008. – 146 с.
2. Жилияков, Е.Г. Методы анализа и построения функций по эмпирическим данным на основе частотных представлений [Текст] / Е.Г. Жилияков. – Белгород: Изд-во БелГУ, 2007. – 160с.
3. Жилияков, Е.Г. Оптимальная фильтрация изображений на основе частотных представлений [Текст] / Е.Г. Жилияков, А.А. Черноморец // Вопросы радиоэлектроники. – Сер. ЭВТ, 2008. – Вып. 1. – С. 118-131.
4. Гонсалес, Р. Цифровая обработка изображений [Текст] / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2006. – 1072 с.

ON EFFICIENCY OF IMAGE OPTIMAL FILTRATION METHOD

E. G. ZHILYAKOV
A. A. CHERNOMORETS
V. A. GOLOSCHAPOVA

Belgorod State University

e-mail:
Zhilyakov@bsu.edu.ru

Using model image, the experimental research results of efficiency of the image optimal filtration method on the basis of frequency representations in the specified frequency subintervals are given in the work.

Key words: optimal filtration, image, Furie transformant, subinterval.

СВЕДЕНИЯ ОБ АВТОРАХ

- Алексеев В.Б.*** – доктор физико-математических наук, профессор, заведующий кафедрой математической кибернетики Московского государственного университета им. М.В. Ломоносова
- Бром А.Е.*** – кандидат экономических наук, преподаватель кафедры промышленной логистики Московского государственного технического университета им. Н.Э. Баумана
- Болтенков В. И.*** – кандидат экономических наук, декан экономического факультета, доцент кафедры мировой экономики Белгородского государственного университета
- Варновский Н.П.*** – старший научный сотрудник института проблем информационной безопасности Московского государственного университета им. М.В. Ломоносова
- Владька М.В.*** – кандидат педагогических наук, доцент, заведующая кафедрой финансов и кредита Белгородского государственного университета
- Волчков В.П.*** – доктор технических наук, профессор Московского технического университета связи и информатики
- Геппенер В.В.*** – доктор технических наук, профессор кафедры математического обеспечения и применения ЭВМ, ведущий научный сотрудник научно-инженерного центра (научно-исследовательский отдел) Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» им. В.И. Ульянова-Ленина
- Голощанова В.А.*** – аспирант кафедры информационно-телекоммуникационных систем и технологий Белгородского государственного университета
- Дмитриенко В.Д.*** – доктор технических наук, профессор кафедры вычислительной техники и программирования Национального технического университета «Харьковский политехнический институт»
- Жиляков Е.Г.*** – доктор технических наук, профессор, заведующий кафедрой информационно-телекоммуникационных систем и технологий Белгородского государственного университета
- Заковоротный А.Ю.*** – старший преподаватель кафедры вычислительной техники и программирования Национального технического университета «Харьковский политехнический институт»



- Захаров В.А.** – кандидат физико-математических наук, доцент факультета вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова
- Игнатъев А.П.** – соискатель Белгородского государственного университета, ЗАО «РК-ТЕЛЕКОМ»
- Искренко Д.И.** – кандидат экономических наук, специалист филиала ОАО «Национальный банк “Траст”», г. Волгоград
- Калугин В.А.** – доктор экономических наук, доцент, начальник отдела инновационной деятельности УНИД Белгородского государственного университета
- Капалин В.И.** – доктор технических наук, профессор кафедры кибернетики Московского государственного института электроники и математики
- Каширина Е.И.** – ассистент, старший лаборант кафедры финансово-экономического инжиниринга Ростовского государственного экономического университета «РИНХ»
- Клионский Д. М.** – аспирант кафедры математического обеспечения и применения ЭВМ Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» им. В.И. Ульянова-Ленина
- Кондрашов Ю.В.** – адъюнкт кафедры информационно-коммуникационных технологий в автоматизированных системах боевого управления Ставропольского военного института связи ракетных войск
- Корневский Н.А.** – доктор технических наук, профессор, заведующий кафедрой биомедицинской инженерии Курского государственного технического университета
- Корчагина Н.А.** – кандидат экономических наук, заведующая кафедрой мировой экономики Астраханского государственного университета
- Кузюрин Н.Н.** – доктор физико-математических наук, начальник отдела Института системного программирования Российской академии наук
- Куриленко Н.Н.** – ассистент кафедры экономики и статистики Белгородского государственного университета

- Кулешов В.Л.** – кандидат технических наук, преподаватель кафедры ракетного и реактивного вооружения Тульского артиллерийского инженерного института
- Лазарев В.В.** – научный сотрудник Федерального государственного унитарного предприятия «Всероссийский научно-исследовательский институт по осушению месторождений полезных ископаемых, защите инженерных сооружений от обводнения, специальным горным работам, геомеханике, геофизике, геологии и маркшейдерскому делу» (ФГУП ВИОГЕМ)
- Мешечкина Р.П.** – доктор экономических наук, профессор, декан факультета таможенного дела Белгородского университета потребительской кооперации
- Михелев В.М.** – кандидат технических наук, доцент кафедры математического и программного обеспечения информационных систем Белгородского государственного университета
- Овчинникова О.П.** – доктор экономических наук, профессор Орловской региональной академии государственной службы
- Омаров Р.Р.** – аспирант кафедры математической кибернетики Московского государственного университета им. М.В. Ломоносова
- Орешко Н.И.** – кандидат технических наук, ведущий научный сотрудник научно-инженерного центра Санкт-Петербургского электротехнического университета
- Парфенова Е.Н.** – старший преподаватель кафедры менеджмента организации Белгородского государственного университета
- Петров Д.А.** – аспирант физического факультета кафедры математики Московского государственного университета им. М. В. Ломоносова
- Петровский А.Б.** – доктор технических наук, заведующий лабораторией методов и систем поддержки принятия решений Института системного анализа Российской академии наук
- Поворознюк А.И.** – кандидат технических наук, доцент, профессор кафедры вычислительной техники и программирования Национального технического университета «Харьковский политехнический институт»
- Попова Е.В.** – специалист отдела организации научных исследований Магнитогорского государственного университета



- Привалов А.Н.** – кандидат технических наук, доцент, член-корреспондент Академии информатизации образования, начальник кафедры математического, программного и информационного обеспечения АСУ Тульского артиллерийского инженерного института
- Смирнова З.В.** – аспирант Московской академии рынка труда и информационных технологий, ЗАО «РК-ТЕЛЕКОМ»
- Титова И.Н.** – ассистент кафедры экономики и статистики Белгородского государственного университета
- Тубольцев М.Ф.** – кандидат технических наук, доцент кафедры математического и программного обеспечения информационных систем Белгородского государственного университета
- Филист С.А.** – доктор технических наук, профессор кафедры биомедицинской инженерии Курского государственного технического университета
- Хавина И.П.** – кандидат технических наук, доцент кафедры вычислительной техники и программирования Национального технического университета «Харьковский политехнический институт»
- Червяков Н.И.** – доктор технических наук, профессор, заслуженный деятель науки и техники Российской Федерации, академик МАИ, профессор кафедры информационно-коммуникационных технологий в автоматизированных системах боевого управления Ставропольского военного института связи ракетных войск
- Черноморец А.А.** – кандидат технических наук, доцент кафедры информационно-телекоммуникационных систем и технологий Белгородского государственного университета
- Чурсин Г.В.** – аспирант Курской государственной сельскохозяйственной академии
- Шокуров А.А.** – кандидат физико-математических наук, старший научный сотрудник Института системного программирования Российской академии наук
- Шутеев В.А.** – аспирант кафедры промышленной логистики Московского государственного технического университета им. Н.Э. Баумана

ИНФОРМАЦИЯ ДЛЯ АВТОРОВ

Уважаемые коллеги!

Материалы необходимо высылать в 2-х экземплярах:

- по адресу: 308015, г. Белгород, ул. Победы, 85, Белгородский государственный университет;
- по электронной почте редакторам разделов: «Актуальные вопросы отечественной истории» – shatohin@bsu.edu.ru (Шатохин Иван Тихонович – заместитель главного редактора); «Актуальные вопросы всеобщей истории» – bolgov@bsu.edu.ru (Болгов Николай Николаевич); «Актуальные вопросы политологии» – Shilov@bsu.edu.ru (Шилов Владимир Николаевич – заместитель главного редактора); «Актуальные проблемы экономики» – Lomovceva@bsu.edu.ru (Ломовцева Ольга Алексеевна – заместитель главного редактора); ответственный секретарь серии журнала – vasilenko_v@bsu.edu.ru (Василенко Виктория Викторовна); сайт журнала: <http://unid.bsu.edu.ru/unid/res/pub/index.php>.

Статьи, отклоненные редколлегией, к повторному рассмотрению не принимаются. Материалы, присланные без соблюдения правил, редколлегией не рассматриваются.

ТРЕБОВАНИЯ К ОФОРМЛЕНИЮ СТАТЕЙ СЕРИИ «ИСТОРИЯ. ПОЛИТОЛОГИЯ. ЭКОНОМИКА. ИНФОРМАТИКА» ЖУРНАЛА «НАУЧНЫЕ ВЕДОМОСТИ БЕЛГУ»

В материалы включается следующая информация:

- | | | |
|---|---|---|
| 1) УДК научной статьи; | } | <i>на русском
и английском
языках</i> |
| 2) аннотация статьи (не более 1200 знаков); | | |
| 3) ключевые слова; | | |
| 4) сведения об авторах (Ф.И.О., должность с указанием места работы (без сокращений), ученая степень, ученое звание, почтовый адрес, адрес электронной почты (если имеется), контактные телефоны); | | |
| 5) внешняя рецензия доктора наук (для аспирантов и кандидатов наук); | } | <i>на русском
языке</i> |
| 6) текст статьи; | | |
| 7) ссылки. | | |

Технические требования к оформлению текста

1. Текст набирается в Microsoft Word 2000/2003. Лист – А4, портретный. Без переносов.

2. Поля:

- правое – 1,5 см;
- левое – 3,0 см;
- нижнее – 2,0 см;
- верхнее – 2,0 см.

3. Шрифт:

- гарнитура: текст – **Georgia**; УДК, название, ФИО автора – **Impact**;
- размер: в тексте – **11 пт**; в таблице – **9 пт**; в названии – **14 пт**.

4. Абзац:

- отступ 1,25 мм, выравнивание – по ширине;
- межстрочный интервал – одинарный.



5. Ссылки:

- номер ссылки размещается перед знаком препинания (перед запятой, точкой);
- нумерация – автоматическая, сквозная;
- текст сноски внизу каждой страницы;
- размер шрифта – 9 пт.

6. Объем статей: до **8 страниц (Georgia, 11 пт)**.

7. Формулы набираются в «Редакторе формул» Word, допускается оформление формул только в одну строку, не принимаются формулы, выполненные в виде рисунков, формулы отделяются от текста пустой строкой.

8. Требования к оформлению статей, таблиц, рисунков приведены в прил. 1, 2, 3.



Приложение 1. Оформление статьи

УДК 65.01

КЛЮЧЕВЫЕ ВЫЗОВЫ РАЗВИТИЮ РЕГИОНА В УСЛОВИЯХ ГЛОБАЛИЗАЦИИ РОССИЙСКОЙ ЭКОНОМИКИ*

А. В. ИВАНОВ¹⁾
Л. Н. ПЕТРОВ²⁾

¹⁾ *Департамент экономического
развития Белгородской
области*

²⁾ *Белгородский
государственный
университет*

e-mail: bor@bsu.edu.ru

При выборе пути инновационного развития необходимо учитывать возможные риски и ограничения социально-экономического развития, продуцированные перспективами постепенного вступления России в единое мировое экономическое пространство. В работе рассмотрены ключевые вызовы развитию России и регионов на долгосрочную перспективу.

Ключевые слова: глобализация, вызовы развитию, риски и ограничения социально-экономического развития, региональная политика.

В последние годы в российском обществе обозначился явный дефицит долгосрочного (на 10-15 и более лет) видения перспектив развития национальной экономики¹.

KEY CHALLENGES TO REGION DEVELOPMENT IN CONDITIONS OF GLOBALIZATION OF THE RUSSIAN ECONOMY

A. V. IVANOV¹⁾
L. N. PETROV²⁾

¹⁾ *Department of Economic
Development, Belgorod Region*

²⁾ *Belgorod State University*

e-mail: bo@bsu.edu.ru

Choosing a way of innovative development it is necessary to take into account the risks and restrictions of socio-economic development, produced by prospects of the gradual introduction of Russia into the whole world economic space. There considered key challenges to development of Russia and its regions for the long-term prospect.

Key words: globalization, challenges to development, risks and restrictions of socio-economic development, regional policy.

СВЕДЕНИЯ ОБ АВТОРАХ

Иванов А.В.

— кандидат экономических наук, доцент кафедры экономики и права Белгородского государственного университета
308015, г. Белгород, ул. Победы, 85, Белгородский государственный университет;
e-mail: dizelsnab@mail.ru, тел. 33-22-44

¹ Караганов С.А. XXI век и интересы России // Современная Европа. 2004. №3. С. 6; Айналов Д.В. Эллинистические основы византийского искусства. СПб., 1900. С. 2.



Приложение 2. Оформление таблиц

1. Каждая таблица должна быть пронумерована справа, иметь заголовок, расположенный по центру.

Таблица 1

Рейтинговая оценка ЦФО за 1999-2004 гг.

Регионы	1999 г.	2000 г.	2001 г.	2002 г.	2003 г.	2004 г.	В среднем за	
							1999-2001 гг.	2002-2004 гг.
РФ	1,3222	1,5091	1,3470	1,4661	1,5940	1,6954	1,3928	1,5852
ЦФО	1,5028	1,9389	1,7210	1,6149	1,6888	1,6930	1,7209	1,6656

2. Таблицы не должны выходить за границы полей страницы слева и справа.

Таблица 1

Рейтинговая оценка ЦФО за 1999-2004 гг.

Регионы	1999 г.	2000 г.	2001 г.	2002 г.	2003 г.	2004 г.	В среднем за	
							1999-2001 гг.	2002-2004 гг.
РФ	1,3222	1,5091	1,3470	1,4661	1,5940	1,6954	1,3928	1,5852
ЦФО	1,5028	1,9389	1,7210	1,6149	1,6888	1,6930	1,7209	1,6656

3. Если таблица располагается на 2-х страницах, ее столбцы должны быть пронумерованы на каждой новой странице, так же, как на первой.

Таблица 1

Рейтинговая оценка ЦФО за 1999-2004 гг.

Регионы	1999 г.	2000 г.	2001 г.	2002 г.	2003 г.	2004 г.	В среднем за	
							1999-2001 гг.	2002-2004 гг.
1	2	3	4	5	6	7	8	9
РФ	1,3222	1,5091	1,3470	1,4661	1,5940	1,6954	1,3928	1,5852
ЦФО	1,5028	1,9389	1,7210	1,6149	1,6888	1,6930	1,7209	1,6656

Таблица, расположенная на первой странице.

Продолжение табл. 1

1	2	3	4	5	6	7	8	9
Белгородская область	1,2620	0,4169	2,2612	1,0176	1,2012	0,6413	1,3134	0,9534
Брянская область	0,9726	0,4817	0,5612	1,8653	0,9064	1,6898	0,6718	1,4872

Таблица, расположенная на следующей странице.

Приложение 3. Оформление графических объектов

1. Изображение каждого графического объекта должно иметь номер и заголовок, расположенные по центру рисунка.

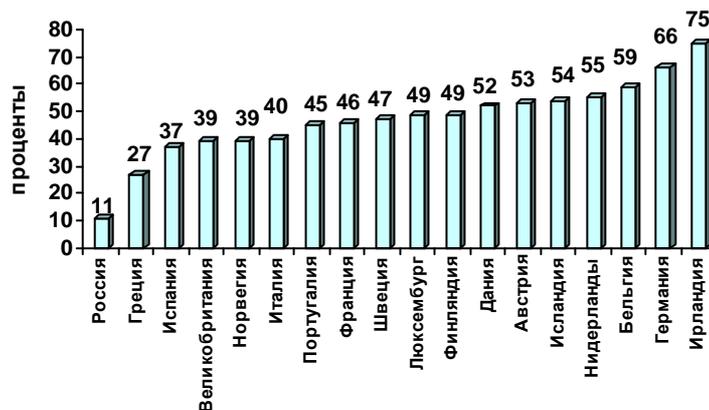


Рис. 1. Уровень инновационной активности в России, странах ЕС, Норвегии, Исландии

2. Изображение графического объекта должно быть в виде рисунка или сгруппированных объектов.

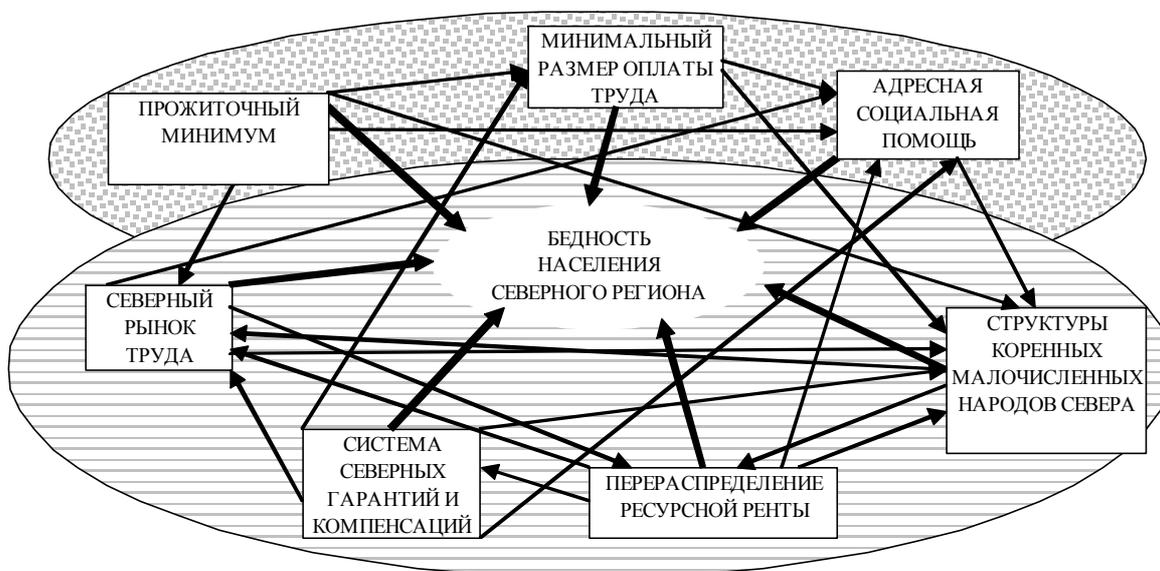


Рис. 2. Институциональная среда существования бедности населения северного региона России

3. Изображение графического объекта не должно выходить за пределы полей страницы.

4. Изображение графического объекта не должно превышать одной страницы.

За публикацию рукописи в журнале «Научные ведомости Белгородского государственного университета» плата с авторов не взимается.