

---

Научно-исследовательский институт  
«Центрпрограммсистем»

---

# Программные продукты и системы

МЕЖДУНАРОДНЫЙ НАУЧНО-ПРАКТИЧЕСКИЙ ЖУРНАЛ

№ 3, 2016

(год издания двадцать девятый)

Главный редактор  
С.В. ЕМЕЛЬЯНОВ, академик РАН

Тверь

## PROGRAMMNYE PRODUKTY I SISTEMY (SOFTWARE & SYSTEMS)

International research and practice journal

2016, vol. 29, no. 3

Editor-in-Chief  
S.V. EMEL'YANOV, *Academician of the Russian Academy of Sciences*

Tver  
Russian Federation

---

Research Institute CENTERPROGRAMSYSTEM

---

**Главный редактор**

**С.В. ЕМЕЛЬЯНОВ**, академик РАН (г. Москва, Россия)

**Научные редакторы:**

**Н.А. СЕМЕНОВ**, д.т.н., профессор ТвГТУ

(г. Тверь, Россия)

**В.Н. РЕШЕТНИКОВ**, д.ф.-м.н., профессор, МАИ

(г. Москва, Россия)

**Рецензенты:**

**А.Н. Сотников**, д.ф.-м.н., профессор МСЦ РАН (г. Москва, Россия)

**А.В. Язенин**, д.ф.-м.н., профессор ТвГУ (г. Тверь, Россия)

**И.А. Бессмертный**, д.т.н., профессор ИТМО

(г. Санкт-Петербург, Россия)

**В.Б. Тарасов**, к.т.н., доцент МГТУ им. Баумана (г. Москва, Россия)

**А.А. Рахманов**, д.т.н., профессор, Концерн «РТИ Системы» (г. Москва, Россия)

## МЕЖДУНАРОДНАЯ РЕДАКЦИОННАЯ КОЛЛЕГИЯ

**Семенов Н.А.** – д.т.н., профессор Тверского государственного технического университета, заместитель главного редактора (г. Тверь, Россия)

**Решетников В.Н.** – д.ф.-м.н., профессор Московского авиационного института (национального исследовательского университета), заместитель главного редактора (г. Москва, Россия)

**Арефьев И.Б.** – д.т.н., профессор Морской академии Польши (г. Щецин, Польша)

**Афанасьев А.П.** – д.ф.-м.н., профессор Московского физико-технического института (технического университета), заведующий Центром распределенных вычислений Института проблем передачи информации РАН (г. Москва, Россия)

**Баламетов А.Б.** – д.т.н., профессор Азербайджанского научно-исследовательского и проектно-изыскательского института энергетики (г. Баку, Азербайджан)

**Батыршин И.З.** – д.т.н., профессор Мексиканского института нефти (г. Мехико, Мексика)

**Вагин В.Н.** – д.т.н., профессор Московского энергетического института (технического университета) (г. Москва, Россия)

**Голенько В.В.** – д.т.н., профессор Белорусского государственного университета информатики и радиоэлектроники (г. Минск, Беларусь)

**Еремеев А.П.** – д.т.н., профессор Московского энергетического института (технического университета) (г. Москва, Россия)

**Котов А.С.** – кандидат наук, ассистент профессора университета Уэйна (штат Мичиган) (г. Детройт, США)

**Кузнецов О.П.** – д.т.н., профессор Института проблем управления РАН (г. Москва, Россия)

**Курейчик В.М.** – д.т.н., профессор Инженерно-технологической академии Южного федерального университета (г. Таганрог, Россия)

**Лисецкий Ю.М.** – к.т.н., генеральный директор «S&T Ukraine» (г. Киев, Украина)

**Мамросенко К.А.** – к.т.н., доцент Московского авиационного института (национального исследовательского университета), руководитель Центра визуализации и спутниковых информационных технологий НИИСИ РАН (г. Москва, Россия)

**Мейер Б.** – доктор наук, профессор, заведующий кафедрой Высшей политехнической школы – ЕТН (г. Цюрих, Швейцария)

**Нгуен Тхань Нгуи** – д.ф.-м.н., профессор, проректор Ханойского открытого университета (г. Ханой, Вьетнам)

**Николов Р.В.** – доктор наук, профессор Университета библиотековедения и информационных технологий Софии (г. София, Болгария)

**Осипов Г.С.** – д.ф.-м.н., профессор, заместитель директора Института системного анализа РАН (г. Москва, Россия)

**Палюх Б.В.** – д.т.н., профессор Тверского государственного технического университета (г. Тверь, Россия)

**Рахманов А.А.** – д.т.н., профессор, заместитель генерального директора Концерн «РТИ Системы» (г. Москва, Россия)

**Серов В.С.** – д.ф.-м.н., профессор Университета прикладных наук Оулу (г. Оулу, Финляндия)

**Сотников А.Н.** – д.ф.-м.н., профессор, Межведомственный суперкомпьютерный центр РАН (г. Москва, Россия)

**Сулейманов Д.Ш.** – академик АН Республики Татарстан, д.т.н., профессор Казанского государственного технического университета (г. Казань, Республика Татарстан, Россия)

**Тарасов В.Б.** – к.т.н., доцент Московского государственного технического университета им. Н.Э. Баумана (г. Москва, Россия)

**Таратухин В.В.** – доктор философии, управляющий директор Европейского исследовательского центра в области информационных систем (ERCIS) Вестфальского университета им. Вильгельма (г. Мюнстер, Германия)

**Хорошевский В.Ф.** – д.т.н., профессор Московского физико-технического института (технического университета) (г. Москва, Россия)

**Язенин А.В.** – д.ф.-м.н., профессор Тверского государственного университета (г. Тверь, Россия)

### АССОЦИИРОВАННЫЕ ЧЛЕНЫ РЕДАКЦИИ

Московский энергетический институт (технический университет), г. Москва, Россия

Технологический институт Южного федерального университета, г. Таганрог, Россия

Тверской государственный технический университет, г. Тверь, Россия

Научно-исследовательский институт «Центрпрограммсистем», г. Тверь, Россия

### АДРЕС РЕДАКЦИИ

Россия, 170024, г. Тверь, пр. 50 лет Октября, 3а

Телефон (482-2) 39-91-49

Факс (482-2) 39-91-00

E-mail: red@cps.tver.ru

www.swsys.ru

Подписано в печать 15.08.2016 г.

Отпечатано ООО ИПП «Фактор и К»

Россия, 170028, г. Тверь, ул. Лукина, д. 4, стр. 1

Выпускается один раз в квартал. Общее количество выпусков 115.

Год издания двадцать девятый. Формат 60×84 1/8. Объем 216 стр.

Заказ № 27. Тираж 1000 экз. Цена 257,40 руб.

Автор статьи отвечает за подбор, оригинальность и точность приводимого фактического материала.

Авторские гонорары не выплачиваются. При перепечатке материалов ссылка на журнал обязательна.

**Editor-in-chief**

**S.V. Emelyanov**, Academician of the Russian Academy of Sciences (Moscow, Russian Federation)

**Science editors:**

**N.A. Semenov**, Dr.Sc. (Engineering), Professor TSTU (Tver, Russian Federation)

**V.N. Reshetnikov**, Dr.Sc. (Physics and Mathematics), Professor MAI (Moscow, Russian Federation)

**Reviewers:**

**A.N. Sotnikov**, Dr.Sc. (Physics and Mathematics), Professor JSCC RAS (Moscow, Russian Federation)

**A.V. Yazenin**, Dr.Sc. (Physics and Mathematics), Professor TSU (Tver, Russian Federation)

**I.A. Bessmertny**, Dr.Sc. (Engineering), Professor ITMO University (St. Petersburg, Russian Federation)

**V.B. Tarassov**, PhD (Technical), Associate Professor BMSTU (Moscow, Russian Federation)

**A.A. Rakhmanov**, Dr.Sc. (Engineering), Professor, JSC RTI (Moscow, Russian Federation)

**INTERNATIONAL EDITORIAL BOARD**

**Semenov N.A.** – Dr.Sc. (Engineering), Professor of Tver State Technical University, Deputy Editor-in-Chief  
(Tver, Russian Federation)

**Reshetnikov V.N.** – Dr.Sc. (Physics and Mathematics), Professor of Moscow Aviation Institute (National Research University),  
Deputy Editor-in-Chief (Moscow, Russian Federation)

**Arefev I.B.** – Dr.Sc. (Engineering), Professor of Poland Szczecin Maritime Academy (Szczecin, Poland)

**Afanasyev A.P.** – Dr.Sc. (Physics and Mathematics), Professor of Moscow Institute of Physics and Technology,  
Head of Centre for Distributed Computing of Institute for Information Transmission Problems (Moscow, Russian Federation)

**Balametov A.B.** – Azerbaijan Scientific-Research & Design-Prospecting Power Engineering Institute (Baku, Azerbaijan)

**Batyrshin I.Z.** – Dr.Sc. (Engineering), Professor of Mexican Petroleum Institute (Mexico City, Mexico)

**Vagin V.N.** – Dr.Sc. (Engineering), Professor of Moscow Power Engineering Institute (Technical University)  
(Moscow, Russian Federation)

**Golenkov V.V.** – Dr.Sc. (Engineering), Professor of Belarusian State University of Informatics and Radioelectronics  
(Minsk, Republic of Belarus)

**Eremeev A.P.** – Dr.Sc. (Engineering), Professor of Moscow Power Engineering Institute (Technical University)  
(Moscow, Russian Federation)

**Kotov A.S.** – Ph.D. (Computer Science), Assistant Professor, Wayne State University (Detroit, MI, USA)

**Kuznetsov O.P.** – Dr.Sc. (Engineering), Professor of the Institute of Control Sciences of the Russian Academy of Sciences  
(Moscow, Russian Federation)

**Kureichik V.M.** – Dr.Sc. (Engineering), Professor of Academy of Engineering and Technology Southern Federal University  
(Taganrog, Russian Federation)

**Lisetskiy Yu.M.** – Ph.D. Tech.Sc., CEO of "S&T Ukraine" (Kiev, Ukraine)

**Mamrosenko K.A.** – Ph.D. (Engineering), Associate Professor of Moscow Aviation Institute (National Research University),  
Head of Center of Visualization and Satellite Information Technologies SRISA RAS (Moscow, Russian Federation)

**Meyer B.** – Dr.Sc., Professor, Head of Department in Swiss Federal Institute of Technology in Zurich, ETH  
(Zurich, Switzerland)

**Nguyen Thanh Nghi** – Dr.Sc. (Physics and Mathematics), Professor, Vice-Principal of Hanoi Open University (Hanoi, Vietnam)

**Nikolov R.V.** – Full Professor of the University of Library Studies and Information Technology (Sofia, Bulgaria)

**Osipov G.S.** – Dr.Sc. (Physics and Mathematics), Professor, Deputy of the Principal of Institute of Systems Analysis  
of the Russian Academy of Sciences (Moscow, Russian Federation)

**Palyukh B.V.** – Dr.Sc. (Engineering), Professor of Tver State Technical University (Tver, Russian Federation)

**Rakhmanov A.A.** – Dr.Sc. (Engineering), Professor, Deputy of the CEO of Concern "RTI Systems"  
(Moscow, Russian Federation)

**Serov V.S.** – Dr.Sc. (Physics and Mathematics), Professor of the Oulu University of Applied Sciences (Oulu, Finland)

**Sotnikov A.N.** – Dr.Sc. (Physics and Mathematics), Professor, Joint Supercomputer Center of the Russian Academy  
of Sciences (Moscow, Russian Federation)

**Suleimanov D.Sh.** – Academician of TAS, Dr.Sc. (Engineering), Professor of Kazan State Technical University  
(Kazan, Republic of Tatarstan, Russian Federation)

**Tarassov V.B.** – Ph.D. (Engineering), Associate Professor of Bauman Moscow State Technical University  
(Moscow, Russian Federation)

**Taratoukhine V.V.** – Ph.D. (Engineering), Dr.Ph., Managing Director of the Competence Centre ERP and ERCIS Lab  
Russia of the ERCIS (Muenster, Germany)

**Khoroshevsky V.F.** – Dr.Sc. (Engineering), Professor of Moscow Institute of Physics and Technology  
(Moscow, Russian Federation)

**Yazenin A.V.** – Dr.Sc. (Physics and Mathematics), Professor of Tver State University (Tver, Russian Federation)

**ASSOCIATED EDITORIAL BOARD MEMBERS**

**Moscow Power Engineering Institute (Technical University), Moscow, Russian Federation**

**Technology Institute at Southern Federal University, Taganrog, Russian Federation**

**Tver State Technical University, Tver, Russian Federation**

**Research Institute CENTERPROGRAMSYSTEM, Tver, Russian Federation**

**EDITORIAL OFFICE ADDRESS**

50 let Oktyabrya Ave. 3a, Tver, 170024, Russian Federation

Phone: (482-2) 39-91-49 Fax: (482-2) 39-91-00

E-mail: red@cps.tver.ru

www.swsys.ru

Passed for printing 15.08.2016

Printed in printing-office "Faktor i K"

Lukina St. 4/1, Tver, 170028, Russian Federation

Published quarterly. 29th year of publication

Format 60×84 1/8. Circulation 1000 copies

Prod. order № 27. Wordage 216 pages. Price 257,40 rub.

## **Вниманию авторов**

Международный журнал «Программные продукты и системы» публикует материалы научного и научно-практического характера по новым информационным технологиям, результаты академических и отраслевых исследований в области использования средств вычислительной техники. Практикуются выпуски тематических номеров по искусственному интеллекту, системам автоматизированного проектирования, по технологиям разработки программных средств и системам защиты, а также специализированные выпуски, посвященные научным исследованиям и разработкам отдельных вузов, НИИ, научных организаций.

Решением Президиума Высшей аттестационной комиссии (ВАК) Министерства образования и науки РФ международный журнал «Программные продукты и системы» внесен в Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней кандидата и доктора наук.

Информация об опубликованных статьях по установленной форме регулярно предоставляется в систему Российского индекса научного цитирования (РИНЦ), в CrossRef и готовится для передачи в международные базы цитирования.

## **Условия публикации**

К рассмотрению принимаются ранее нигде не опубликованные материалы, соответствующие тематике журнала (специализация 05.13.XX – Информатика, вычислительная техника и управление) и отвечающие редакционным требованиям.

Работа представляется в электронном виде в формате Word. При обилии сложных формул обязательно наличие статьи и в формате PDF. Формулы должны быть набраны в редакторе формул Word (Microsoft Equation или MathType). Объем статьи вместе с иллюстрациями – не менее 10 000 знаков. Диаграммы, схемы, графики должны быть доступными для редактирования (Word, Visio, Excel). Все иллюстрации для полиграфического воспроизведения представляются в черно-белом варианте. Цветные, тонированные, отсканированные, не подлежащие редактированию средствами Word рисунки и экранные формы следует присылать в хорошем качестве для их дополнительного размещения на сайте журнала в макете статьи с доступом по ссылке. (Публикация материалов с использованием гипертекста, графики, аудио-, видео-, программных средств и др. возможна в электронном издании «Программные продукты, системы и алгоритмы», сайт [www.swsys-web.ru](http://www.swsys-web.ru).) Заголовок должен быть информативным; сокращения, а также терминологию узкой тематики желательно в нем не использовать. Количество авторов на одну статью – не более 4, количество статей одного автора в номере, включая соавторство, – не более 2. Список литературы (оформленный в соответствии с ГОСТ Р 7.05–2008), наличие которого обязательно, должен включать не менее 10 пунктов.

Необходимы также аннотация (не менее 200 слов), ключевые слова (7–10) и индекс УДК. Название статьи, аннотация и ключевые слова должны быть переведены на английский язык (машинный перевод недопустим), а фамилии авторов, названия и юридические адреса организаций (если нет официального перевода), приставные списки литературы – транслитерированы по стандарту BGN/PCGN.

Вместе со статьей следует прислать отзыв-рекомендацию в произвольной форме, экспертное заключение, лицензионное соглашение, а также сведения об авторах: фамилия, имя, отчество, название и юридический адрес организации, должность, ученые степень и звание (если есть), контактный телефон, электронный адрес, почтовый адрес для отправки бесплатного авторского экземпляра журнала.

## **Порядок рецензирования**

Все статьи, поступающие в редакцию (соответствующие тематике и оформленные согласно требованиям к публикации), подлежат обязательному рецензированию в течение месяца с момента поступления.

В редакции есть устоявшийся коллектив рецензентов, среди которых члены международной редколлегии журнала, эксперты из числа крупных специалистов в области информатики и вычислительной техники ведущих вузов страны, а также ученые и специалисты НИИ «Центрпрограммсистем» (г. Тверь).

Рецензирование проводится конфиденциально. Автору статьи предоставляется возможность ознакомиться с текстом рецензии. При необходимости статья отправляется на доработку.

Рецензии обсуждаются на заседаниях рабочей группы, состоящей из членов научного совета журнала. Заседания проводятся раз в месяц в НИИ «Центрпрограммсистем» (г. Тверь), где принимается решение о целесообразности публикации статьи.

Статьи, одобренные редакционным советом, публикуются бесплатно в течение года с момента одобрения, а отправленные на доработку – с момента поступления после устранения замечаний.

Редакция международного журнала «Программные продукты и системы» в своей работе руководствуется сводом правил Кодекса этики научных публикаций, разработанным и утвержденным Комитетом по этике научных публикаций.

УДК 004.5

Дата подачи статьи: 22.03.16

DOI: 10.15827/0236-235X.115.005-009

## **ЕСТЕСТВЕННО-ЯЗЫКОВОЙ ПОЛЬЗОВАТЕЛЬСКИЙ ИНТЕРФЕЙС ДИАЛОГОВОЙ СИСТЕМЫ**

*(Исследование выполнено при финансовой поддержке Фонда содействия развитию малых форм предприятий в научно-технической сфере в рамках научно-исследовательского проекта № 7369ГУ2015)*

*Р.В. Посевкин, аспирант, rus\_posevkin@mail.ru;*

*И.А. Бессмертный, д.т.н., профессор, igor\_bessmertny@hotmail.com  
(Санкт-Петербургский национальный исследовательский университет  
информационных технологий, механики и оптики (Университет ИТМО),  
Кронверкский просп., 49, г. Санкт-Петербург, 197101, Россия)*

Множество людей ежедневно взаимодействуют с различными программными системами. В результате внедрения компьютерных систем во все сферы человеческой жизни все больше проявляется проблема перехода от визуальных и командных интерфейсов к естественно-языковым. В статье рассмотрены методы компьютерной лингвистики и обработки естественного языка. Представлено полное описание всех стадий обработки естественного языка, таких как морфологический, синтаксический и семантический анализ. Рассмотрен ограниченный язык как подмножество естественного языка, на котором текст хорошо воспринимается носителем естественного языка без дополнительных усилий. Подобное решение позволяет сократить время анализа естественно-языковых элементов в диалоговой системе, а также помогает избежать неоднозначностей на лингвистическом уровне. Методы для работы с естественным языком могут применяться в программных системах машинного перевода, поисковых системах, системах автоматического реферирования и в экспертных системах. Представлено формальное описание диалоговой системы в контексте марковского процесса принятия решений. Разработан прототип естественно-языкового пользовательского интерфейса, который производит преобразования пользовательского запроса на естественном языке в SQL-запрос к БД. Интерфейс взаимодействует с БД, содержащей информацию о существующих программных библиотеках и фреймворках. Таким образом, использование методов обработки естественного языка позволяет разработать естественно-языковой пользовательский интерфейс для взаимодействия с диалоговой системой.

**Ключевые слова:** диалоговая система, обработка естественного языка, пользовательский интерфейс, БД.

Существует большое количество программных систем, каждая из которых обладает характерными именно для нее принципами взаимодействия. В результате свободное использование системы затрудняется и увеличивается время, необходимое на обучение пользователя работе с ней.

Одним из вариантов решения проблемы взаимодействия с компьютерной системой является использование естественного языка, более привычного для пользователя.

Такое решение имеет следующие преимущества:

- минимальная подготовка пользователя для работы с системой;
- высокая скорость и простота формирования произвольных запросов к системе.

Простота при работе с естественно-языковым пользовательским интерфейсом достигается путем применения пользователем языка, используемого в ежедневной коммуникации [1].

Для естественно-языкового пользовательского интерфейса диалоговых систем возможно использование ограниченной лексики и грамматики языка. При этом не происходит серьезного ухудшения функциональности и производительности вопросно-ответной системы.

Ограниченный естественный язык – это подмножество естественного языка, на котором текст воспринимается носителем естественного языка без дополнительных усилий. При этом не требуется изучение ограниченной версии языка для

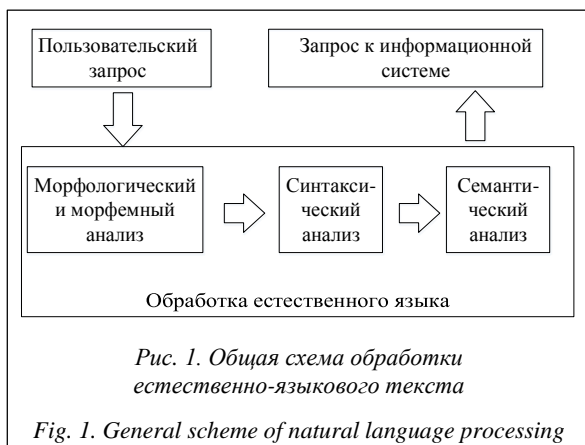
составления текстов. Данный язык обладает сокращенным набором лексики и грамматики, что позволяет сократить время анализа естественно-языковых элементов в диалоговой системе, а также избежать неоднозначностей на лингвистическом уровне [2].

Цикл работы естественно-языкового пользовательского интерфейса начинается с ввода пользователем текста сообщения на естественном языке. На основе введенного текста строится его формальное описание. При этом все предшествующие результаты анализа используются при анализе последующих запросов, что позволяет системе сохранять нить диалога с пользователем и разрешать неоднозначные и спорные моменты, связанные с различным использованием одних и тех же слов в разных предметных областях [3].

В процессе обработки естественно-языкового текста последовательно выполняется морфологический, синтаксический и семантический анализ. Общая схема работы системы представлена на рисунке 1.

На первом этапе обработки пользовательского запроса осуществляется морфологический и морфемный анализ. Для каждого слова предложения в формальной записи строятся отношения, задающие соответствия для значений грамматических категорий [4].

В результате морфологического анализа определяются морфологические характеристики каждого слова, такие как падеж, склонение, часть речи



и т.д. Количество и наличие морфологических характеристик слов и допустимых значений зависят от конкретного языка. Однако некоторые характеристики, например часть речи, существуют во многих языках.

Для проведения морфологического анализа текста существуют следующие подходы: четкая морфология, нечеткая морфология, вероятностный подход.

При обработке естественно-языковых текстов на русском языке наиболее часто применяется подход, основанный на четкой морфологии и базирующийся на словаре Зализняка. В этом словаре описаны основные словоформы для слов русского языка. Каждой такой словоформе сопоставляется определенный код. Существует система правил, в соответствии с которой для слова можно построить все остальные формы [5]. В качестве исходных данных используются начальная словоформа и соответствующий ей код. В случае использования четкого подхода в процессе морфологического анализа необходимо иметь в наличии словарь всех словоформ и слов для данного языка. Словарь на входе принимает форму слова, а на выходе представляет морфологические характеристики конкретной словоформы.

Возможно построение словаря на основе словаря Зализняка по следующему алгоритму. Сначала перебираются все слова, содержащиеся в словаре. Для каждого из этих слов выделяются все допустимые словоформы и сохраняются в формируемом словаре. В процессе морфологического анализа конкретного слова достаточно найти его в словаре. В результате будут получены точные значения всех морфологических характеристик анализируемого слова.

При морфемном анализе для каждого слова выделяются морфемы: приставка, корень, суффикс, окончание. В словаре морфем русского языка для каждого слова указано разделение на составные части. При этом в словаре не указываются типы каждой из составных частей. Таким образом, достоверно неизвестно, какая из частей является суффиксом, а какая корнем.

Совокупность всех возможных корней слов русского языка представляет собой открытое множество. В то же время множество всех приставок, суффиксов и окончаний ограничено. Также известен порядок следования морфем в составе слова на русском языке: приставки, затем корни, суффиксы и окончания. Таким образом, с помощью словаря морфем русского языка можно построить словарь, содержащий информацию как о разбиении слова на составные части, так и о типе каждой из морфем. В результате для проведения морфемного анализа слова достаточно использовать разработанный словарь.

В случае отсутствия слова в словаре возможен непосредственный анализ на основе стандартного строения слов русского языка (приставка – корень – суффикс – окончание) и множества всех приставок, суффиксов и окончаний [6].

На этапе синтаксического анализа выделяются отношения синтаксических связей внутри предложения. Далее определяются главные и второстепенные члены предложения, тип предложения. Синтаксический анализ выполняется поэтапно: при описании формальной структуры предложения используется полученная на этапе морфологического анализа информация. На данном этапе используются синтаксические и лексические правила анализируемого языка.

Далее следует наиболее сложная часть обработки естественно-языкового текста – этап семантического анализа, который можно рассматривать как обработку семантической сети. В свою очередь, семантическая сеть представляет собой результат анализа естественно-языкового текста на предыдущих этапах и присутствующих в системе знаний о предметной области и языке программной системы. На данном этапе анализируются соотношения лингвистической конструкции и хранимых в памяти системы конструкций, предназначенных для выявления отношений соответствия.

В процессе построения семантической модели слова учитывается его многозначность. Смысл слова рассматривается как множество различных значений. Вся совокупность этих значений в целом или каждое из них реализуется в определенном контексте. Результатом семантического анализа является конструкция запроса к информационной системе.

Система преобразования естественно-языкового текста в запрос к БД является частью диалоговой системы. Рассмотрим диалоговую систему в контексте марковского процесса принятия решений. Данный процесс представлен в виде диалога, конечная цель которого – получение пользователем списка библиотек за наименьшее количество возможных итераций. В рассматриваемом примере пользователь формирует запрос с использованием значений «язык программирования» и «тип лицензий» в качестве фильтра.

Марковский процесс принятия решений описывается в терминах пространства состояний, набора действий и стратегии. Состояния диалоговой системы затрагивают все ресурсы, с которыми взаимодействует пользователь [7].

Полный набор состояний диалоговой системы может включать в себя:

- одно начальное состояние;
- различные комбинации списков библиотек, формируемых в зависимости от начальных значений и используемого фильтра;
- специальное состояние, обозначающее окончание работы текущей диалоговой сессии.

В данном примере рассматриваются такие состояния диалоговой системы, как список библиотек, выбранных с применением фильтров «язык программирования» и «тип лицензии».

В качестве возможных действий с диалоговой системой рассматриваются взаимодействие пользователя с системой (например, ввод пользователем запроса к системе, предоставление системой ответа на пользовательский запрос, уведомления), взаимодействие системы с внешними ресурсами (например, отправка запроса к БД), внутренняя обработка данных [8].

В предлагаемом примере возможен следующий набор действий с диалоговой системой.

1. Пользовательский запрос к системе на получение списка программных библиотек для определенного языка программирования.
2. Пользовательский запрос к системе на получение списка программных библиотек, распространяемых в соответствии с запрашиваемым типом лицензии.
3. Пошаговые запросы, когда пользователь запрашивает список программных библиотек для конкретного языка программирования, а затем отбирает из них только библиотеки с определенным типом лицензии (или в обратном порядке).
4. Финальное действие, завершающее текущую сессию диалога с программной системой.

Когда действие  $a$  производится над системой в состоянии  $s$ , система переходит в состояние  $s'$ . В подобной ситуации вероятность перехода описывается формулой

$$P_T(s(t+1) = s' | s(t) = s, a(t) = a). \quad (1)$$

Диалоговая сессия представляет собой длину пути при переходе пользователя из начального в конечное состояние диалоговой системы. При этом диалоговая стратегия отображает, какое действие будет совершено для перехода в каждое последующее состояние программной системы.

Производительность системы можно измерить с помощью целевой функции  $C$ , где веса  $C_i$  представляют собой расстояние до достижения конечной цели. Таким образом, эффективность взаимодействия пользователя с диалоговой системой рассчитывается по формуле

$$C = \sum C_i. \quad (2)$$

В целом веса в марковском процессе принятия решений описываются условным распределением  $P_C(c(t) = c | s(t) = s, a(t) = a)$ . (3)

Совокупность из четырех состояний системы, набора вариантов взаимодействий с системой, вероятностей перехода и распределения весов определяют марковский процесс принятия решений. Для рассматриваемого примера целевая функция представлена как

$$C = W_i \times \langle \# \text{ interactions} \rangle + W_e \times \langle \# \text{ errors} \rangle + W_f \times \langle \# \text{ incomplete values} \rangle. \quad (4)$$

Целевая функция включает в себя три компонента: первый – ожидаемая продолжительность диалога, второй – предполагаемое количество ошибок в полученных данных, третий – предполагаемое расстояние до получения пользователем целевого результата от диалоговой системы.

В зависимости от соответствия целевой функции диалоговой модели возникает связь веса  $c$  с совершаемым взаимодействием  $a$  пользователя с системой, находящейся в состоянии  $s$ .

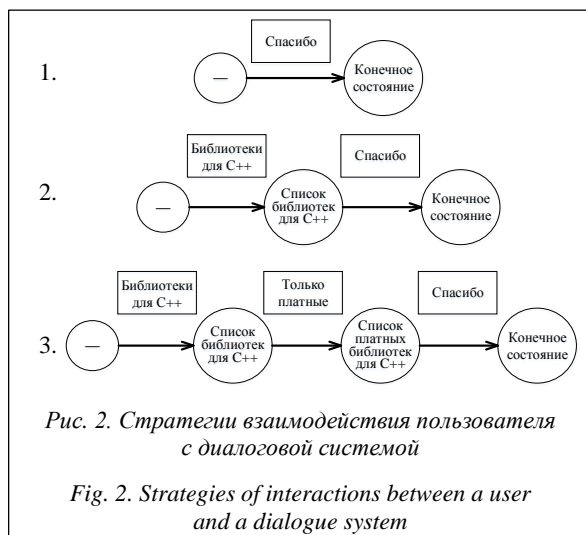
Вес включает в себя любое из первых трех действий с диалоговой системой:  $W_i + W_e \times \text{number of errors}$ . Рассмотрим понятие веса ошибок для определения значений с применением фильтров по отдельности (для действий 1 и 2) – это  $p_1$ , и совместно (для действия 3) – это  $p_2$ , при этом  $p_2 > p_1$ . Тогда ожидаемый вес для взаимодействий 1 и 2 будет представлен как  $W_i + W_e \times p_1$ , для действия 3 –  $W_i + 2 \times W_e \times p_2$ . Для действия 4 (закрытие диалоговой сессии) вес зависит от состояния диалоговой системы, в котором произошло взаимодействие. Так, для начального состояния вес будет выражен как  $W_i + 2 \times W_f$ . Для случаев, когда был применен только один из фильтров («язык программирования» или «тип лицензии»), вес определяется как  $W_i + W_f$ . В случае применения обоих фильтров вес представлен как  $W_i$ .

На рисунке 2 отображены три различные стратегии пользовательского взаимодействия с диалоговой системой. Оптимальная стратегия стремится к минимизации целевой функции. В данном случае  $C_1 = W_i + 2 \times W_f$ . В соответствии со стратегией 1, когда пользователь первым же действием закрывает диалоговую сессию, оптимальной является ситуация со слишком большим весом ошибок:  $p_1 > (W_f - W_i)/W_e$ .

В стратегии 2 пользователь запрашивает список библиотек с использованием одного из фильтров, а затем закрывает диалоговую сессию. В данном случае  $C_2 = 2 \times W_i + 2 \times p_2 \times W_e$ .

В стратегии 3 пользователь запрашивает список программных библиотек с использованием как фильтра «язык программирования», так и «тип лицензии» и после этого закрывает диалоговую сессию. В текущей ситуации целевая функция имеет следующий вид:  $C_3 = 3 \times W_i + 2 \times p_1 \times W_e$ . Стратегия 3 оптимальна, когда верно выражение  $p_2 - p_1 > W_i/2 \times W_e$ .





Существуют различные методы вычисления оптимальной стратегии с учетом корректных параметров модели (вероятности перехода и распределения весов), среди которых алгоритм итераций по критерию или итерации по стратегиям. Данные подходы основаны на динамическом программировании, которое может быть использовано благодаря марковскому характеру представленной модели [9]. Подходы основаны на следующем определении: оптимальное значение  $V(s)$  состояния диалоговой системы  $s$  представляет собой минимальный ожидаемый вес, формируемый начиная с момента, когда система перешла из состояния  $s$ , и до тех пор, пока не достигла финального (конечного) состояния. Оптимальное значение функции уникально и может быть определено как решение системы уравнений:

$$V(s) = \min_a \left( \langle C(s, a) \rangle + \sum_{s'} P_T(s'|s, a) V(s') \right), \quad (5)$$

где  $\langle C(s, a) \rangle$  – ожидаемый вес для взаимодействия  $a$  в состоянии  $s$  [10].

В рамках исследований разработан прототип программной системы, реализующей естественно-языковой пользовательский интерфейс к структурированному источнику данных.

В качестве исходных данных представлена БД MySQL, которая содержит информацию о существующих программных библиотеках и фреймворках. Тестовая БД содержит в себе таблицу, имеющую следующие поля: уникальный идентификатор, имя, ссылка на официальный сайт или репозиторий, дата создания, автор, тип лицензии, тип записи (библиотека или фреймворк), размер, список зависимостей, язык программирования, наличие активного сообщества разработчиков.

Значением, получаемым на выходе программной системы, является SQL-запрос к БД. Одно из необходимых условий для работы системы – наличие структуры данных, описывающей содержимое БД. В частности, необходимо описание таблицы и входящих в нее полей.

Сложные предложения на естественном языке не используются в пользовательском запросе к базе в рамках естественно-языкового пользовательского интерфейса. Несмотря на то, что требуется создание лингвистического анализатора в виде преобразователя, состоящего из двух видов абстракции – морфологической и синтаксической [11], каждый из уровней абстракции должен иметь в наличии компонентную модель с набором правил и библиотек и определенный образ запросов в виде морфологических и синтаксических структур.

Также не требуется фаза семантического анализа за счет того, что известна предметная область. Таким образом, лингвистический анализатор можно назвать лингвистическим преобразователем (транслятором).

Добиться приемлемого уровня обработки естественного языка возможно без использования полного грамматического анализа. Достаточно извлечь наиболее информативные части предложения, такие как ключевые слова, фразы и фрагменты. Следующим шагом является создание морфологических и синтаксических моделей с использованием морфологических и синтаксических методов анализа.

Например, пользовательский запрос на естественном языке «Показать библиотеки для C++» будет преобразован в SQL-запрос к БД `SELECT Name, Url FROM Data WHERE Type='library' AND Language='cpp'`.

Таким образом, естественно-языковой пользовательский запрос на русском языке преобразуется в SQL-запрос, который в дальнейшем может быть отправлен к БД.

В результате разработанная программная система реализует естественно-языковой пользовательский интерфейс к БД. Подобная система может применяться в областях, где требуется обеспечить доступ пользователя к информации из определенной предметной области. Например, возможна реализация естественно-языкового интерфейса для аналитической программной системы. Введя запрос на естественном языке, пользователь получает отчет, выборку по данным за определенный временной период. В результате, применяя методы обработки естественно-языкового текста, можно создать программный интерфейс, обеспечивающий взаимодействие пользователя с программной системой на естественном языке.

### Литература

1. Bessmertny I. On constructing intellectual systems in ternary logic. Programming and Computer Software, 2014, vol. 40, no. 1, pp. 43–46.
2. Deshpande A.K., Devalle P.R. Natural language query processing using probabilistic context free grammar. Intern. Journ. of Advances in Engineering & Technology. 2012, vol. 3, no. 2, pp. 568–573.
3. Житко В.А. Пользовательский интерфейс интеллектуальных вопросно-ответных систем // NB: Кибернетика и программирование. 2012. № 1. С. 23–30.



4. Крайванова В.А. Модель естественно-языкового интерфейса для систем управления сложными техническими объектами и оценка эффективности алгоритмов на ее основе // Управление большими системами. М.: Изд-во ИПУ РАН, 2009. Вып. 26. С. 158–178.
5. Damjanovic D., Agatonovic M., and Cunningham H. FREyA: An interactive way of querying Linked Data using natural language. Proc. 8th Intern. conf. on the Semantic Web: ESWC 2011. Springer Berlin Heidelberg, 2012, pp. 125–138.
6. Селезнев К. Обработка текстов на естественном языке // Открытые системы. 2003. № 12; URL: <http://www.osp.ru/os/2003/12/183694/> (дата обращения: 27.02.2016).
7. Li F., Jagadish H.V. NaLIR: An interactive natural language interface for querying relational databases. Proc. 2014 ACM SIGMOD Intern. conf. on Management of data, ACM, 2014, pp. 709–712.
8. Dezhao S., Schilder F., Smiley C., Brew C., Zielund T., Bretz H., Martin R., Dale C., Pomerville S., Duprey J., Miller T., and Harrison J. TR Discover: a natural language interface for querying and analyzing interlinked datasets. Proc. 14th Intern. conf. on the Semantic Web: ISWC 2015, Springer Intern. Publ., 2015, pp. 21–37.
9. Kesavan S., Giudice N.A. Indoor scene knowledge acquisition using a natural language interface. SKALID 2012 – Spatial Knowledge Acquisition with Limited Information Displays, 2012, pp. 1–6.
10. Levin E., Pieraccini R., Eckert W. Learning dialogue strategies within the Markov decision process framework. Automatic Speech Recognition and Understanding IEEE Proc., 1997, pp. 72–79.
11. Елисеева О.Е. Естественно-языковой интерфейс интеллектуальных систем: учеб. пособие. Минск: Изд-во БГУИР, 2009. С. 84–85.

DOI: 10.15827/0236-235X.115.005-009

Received 22.03.16

**NATURAL LANGUAGE USER INTERFACE OF A DIALOGUE SYSTEM**

(Acknowledgements. The research was partially supported by FASIE (Fund for Assistance to Small Innovative Enterprises), research project no. 7369GU2015)

<sup>1</sup>Posevkin R.V., Postgraduate Student, [rus\\_posevkin@mail.ru](mailto:rus_posevkin@mail.ru)

<sup>1</sup>Bessmertny I.A., Dr.Sc. (Engineering), Professor, [igor\\_bessmertny@hotmail.com](mailto:igor_bessmertny@hotmail.com)

<sup>1</sup>The National Research University of Information Technologies, Mechanics and Optics, Kronverksky Ave. 49, St. Petersburg, 197101, Russian Federation

**Abstract.** Many people interact with software every day. Due to computer systems expansion to all fields of people's activity the problem of transition from visual and command interfaces to natural language user interfaces is becoming more and more relevant. The article describes computational linguistics and natural language processing methods. It considers all phases of natural language processing such as morphological, syntactic and semantic analysis. Limited language is represented as a subset of natural language. This language allows avoiding linguistic uncertainty and decreases natural language elements processing time. Translation system is a part of a dialogue system. Methods of natural language manipulation are applied in machine translation software systems, search engines, data exchange, automatic abstracting and expert systems. The paper represents a formal definition of a dialogue system as a Markov decision process. The authors developed a prototype of natural language user interface for a structured data source. It converts natural language user's query into SQL query to database. User interface interacts with a database that contains information about existent program libraries and frameworks. Consequently, using natural language processing methods makes it possible to develop natural language user interface providing a capability to interact with a dialogue system.

**Keywords:** dialogue system, natural language processing, user interface, database.

**References**

1. Bessmertny I. On constructing intellectual systems in ternary logic. *Programming and Computer Software*. 2014, vol. 40, no. 1, pp. 43–46.
2. Deshpande A.K., Devalle P.R. Natural language query processing using probabilistic context free grammar. *Int. Journ. of Advances in Engineering & Technology*. 2012, vol. 3, no. 2, pp. 568–573.
3. Zhitko V.A. Intelligent user interface for question answering systems. *NB: Kibernetika i programmirovaniye* [Cybernetics and Programming]. 2012, no. 1, pp. 23–30 (in Russ.).
4. Krayvanova V.A. Model of natural-language interface for complex technical objects control systems and efficiency evaluation of model-based algorithms. *Upravlenie bolshimi sistemami* [Large-scale Systems Control]. Moscow, Institute of Control Sciences Publ., 2009, vol. 26, pp. 158–178 (in Russ.).
5. Damjanovic D., Agatonovic M., Cunningham H. FREyA: An interactive way of querying Linked Data using natural language. *The Semantic Web: ESWC 2011 Workshops*. Springer Berlin Heidelberg Publ., 2012, pp. 125–138.
6. Seleznev K. Text processing on a natural language. *Otkrytye sistemy* [Open Systems]. 2003, no. 12. Available at: <http://www.osp.ru/os/2003/12/183694/> (accessed February 27, 2016) (in Russ.).
7. Li F., Jagadish H.V. NaLIR: An interactive natural language interface for querying relational databases. *Proc. 2014 ACM SIGMOD Int. Conf. on Management of Data*. ACM Publ., 2014, pp. 709–712.
8. Dezhao S. TR Discover: A Natural Language Interface for Querying and Analyzing Interlinked Datasets. *The Semantic Web-ISWC 2015*. Springer Int. Publ., 2015, pp. 21–37.
9. Kesavan S., Giudice N.A. Indoor Scene Knowledge Acquisition using a Natural Language Interface. *SKALID 2012 – Spatial Knowledge Acquisition with Limited Information Displays*. 2012, pp. 1–6.
10. Levin E., Pieraccini R., Eckert W. Learning dialogue strategies within the Markov decision process framework. *Automatic Speech Recognition and Understanding IEEE Proc.* 1997, pp. 72–79.
11. Елисеева О.Е. *Естественно-языковой интерфейс интеллектуальных систем* [Natural Language Interface of Intellectual Systems]. Study guide. Minsk, BGUIR Publ., 2009, pp. 84–85.

УДК 004.89

Дата подачи статьи: 01.03.16

DOI: 10.15827/0236-235X.115.010-014

## **ИСПОЛЬЗОВАНИЕ ВЕРОЯТНОСТНОГО ВЫВОДА В СЛАБОФОРМАЛИЗОВАННЫХ БАЗАХ ЗНАНИЙ**

*(Работа выполнена в рамках темы НИР № 615869**«Методы проектирования ключевых систем информационной инфраструктуры»)**Е.А. Полещук, аспирант, eapoleschuk@corp.ifmo.ru;**А.В. Платонов, аспирант, avplatonov@corp.ifmo.ru**(Санкт-Петербургский национальный исследовательский университет  
информационных технологий, механики и оптики (Университет ИТМО),  
Кронверкский просп., 49, г. Санкт-Петербург, 197101, Россия)*

В статье рассматривается процесс вероятностного вывода в слабоформализованных базах знаний. В качестве такой базы выбрана графическая модель представления знаний – семантическая сеть. Выбор обусловлен удобством представления автоматически извлеченных данных в виде графа со связями, а также удобством дальнейшего использования (чтения, изменения и поиска ошибок) данного графа. Также проводится сравнение широко используемого на данный момент продукционного подхода с предложенным, указываются основные недостатки продукционного подхода, которые необходимо учитывать при разработке подобных систем извлечения знаний.

Целью исследования является извлечение новых знаний из автоматически полученных данных. Для достижения этой цели на графических моделях обычно производится логический вывод. Поскольку модель, а также способ получения данных (в данном случае автоматически или полуавтоматически) накладывают ограничения на механизм вывода, алгоритм вынужден работать в условиях неопределенности. Отсюда следует, что стандартные механизмы логического вывода, предусмотренные для данной модели, становятся неактуальными.

В статье предлагается использовать вероятностный вывод и, следовательно, вероятностный язык логического программирования для его реализации. Также делается сравнение нескольких существующих языков вероятностного логического программирования, таких как PRISM, ICL и ProbLog. По результатам сравнения производится выбор языка вероятностного программирования для осуществления вывода. В качестве такого языка выбран язык ProbLog (в частности система ProbLog2), являющийся вероятностным расширением языка Prolog.

**Ключевые слова:** вероятностный вывод, семантические сети, слабоформализованные базы знаний, ProbLog, вероятностное логическое программирование, неопределенность данных, неопределенность отношений.

Одним из популярных подходов к представлению знаний является продукционный подход. Продукционная модель – это модель, в которой знания представлены с помощью правил «ЕСЛИ – ТО» [1]. Эта модель довольно просто интерпретируется человеком, так как правила в ней схожи с предложениями на естественном языке. Положительными сторонами данной модели являются также простота создания и интерпретации отдельных правил (ЕСЛИ X – человек, ТО X смертен), изменения правил и пополнение базы правилами, довольно простой и прозрачный механизм логического вывода.

Однако, помимо сильных сторон, у данной модели имеются и слабые:

- не ясно, каким образом правила соотносятся между собой;
- довольно сложно составить целостную картину знаний по отдельным правилам;
- продукционная модель отличается от структуры знаний человека;
- в логическом выводе отсутствует гибкость.

Таким образом, задачи, решаемые с помощью продукционной модели, должны быть небольшими (оперировать небольшим объемом знаний) и не требующими гибкого вывода [2].

Ответ на интересующий вопрос можно получить в БЗ, основанных на продукционной модели, с помощью логического вывода (прямого или обрат-

ного). Обратный логический вывод осуществляет поиск от целей к данным. При прямом выводе происходит поиск от данных к целям [3]. Прямой вывод используется тогда, когда по исходным данным требуется вывести все следствия (задача интерпретации или прогнозирования), тогда как обратный вывод применяется в случае необходимости проверки определенной гипотезы на соответствие фактам (задача диагностики). Подробные примеры приведены в [3].

Так как продукционная БЗ, как и программа на языке Prolog, состоит из правил, соответственно, можно создавать программы, написанные на этом языке и осуществляющие логический вывод в этой БЗ. Такое соответствие на примере задачи поиска всех братьев, если известны имена людей и родственные отношения между ними, проиллюстрировано в таблице 1.

Продукционную модель можно представить в виде графа (рис. 1 и 2).

Однако следует отметить, что логический вывод плохо работает с данными в условиях неопределенности. Далее речь пойдет о вариантах решения этой проблемы.

Построение более или менее больших БЗ требует очень больших затрат ресурсов, так как это в основном выполняется вручную. Поэтому логичным кажется использование методов машинного обучения для построения таких БЗ [4]. Эти методы,

соответственно, обладают определенной точностью, то есть мы не совсем уверены в правильности данных, полученных автоматически. Отсюда следует вероятностная природа полученных данных, то есть явление неопределенности.

Таблица 1

### Соответствие программы на Prolog продукционной БЗ

Table 1

### Prolog program matching to a rule-oriented knowledge base

Продукционная БЗ	Программа на Prolog
<b>Факты:</b> Джейк – отец Тома Джейк – отец Джима Том – отец Тима Петр является дядей Тома	<b>Факты:</b> parent(jake, tom) parent(jake, jim) parent(tom, tim) uncle(peter, tom)
<b>Правила:</b> 1. Человек1 и Человек2 братья, если существует ЧеловекРодитель, который является родителем для Человека1 и Человека2, и при этом Человек1 – это не Человек2. 2. Человек1 и Человек2 братья, если у Человека1 есть ребенок, и Человек2 дядя этого ребенка	<b>Правила:</b> 1. brother(Man1, Man2) :- parent(Parent, Man1), parent(Parent, Man2), Man1 <> Man2. 2. brother(Man1, Man2) :- parent(Man1, Child), uncle(Man2, Child)
<b>Запрос:</b> Кто из людей являются братьями?	<b>Запрос:</b> askMainQuestion() :- brother(X, Y)
	<b>Ответ:</b> Jake Peter, Tom Jim, Jim Tom

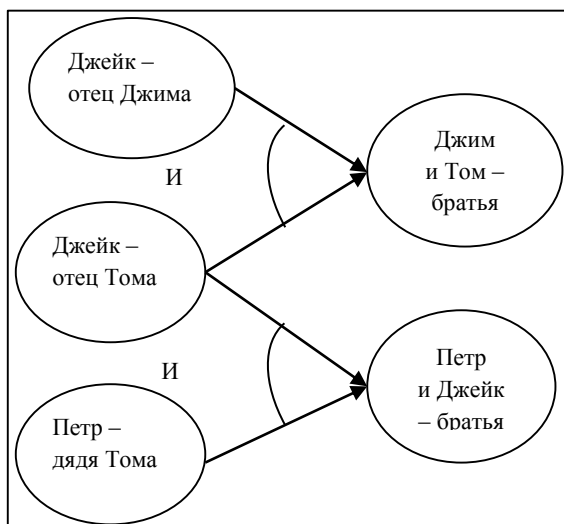


Рис. 1. Представление базы правил БЗ для поиска братьев в виде графа И/ИЛИ, где дуга обозначает отношение И

Fig. 1. Representation of the knowledge base's rules base to search for the brothers in the form of a graph "AND / OR" where the arc means "AND" relation

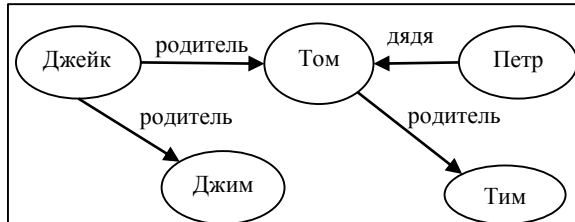


Рис. 2. Граф отношений БЗ

Fig. 2. Knowledge base relationships graph

### Графические модели представления знаний

Поскольку самым удобным способом представления автоматически извлеченных данных является граф, сформируем структуру из автоматически извлеченных данных. В данной работе в качестве графической модели БЗ примем семантическую сеть [5].

Семантические сети (рис. 3) представляют собой информационную модель предметной области, графическое представление БЗ (ориентированный граф) [6]. Они способны представлять объекты (находятся в вершинах графа), их категории и отношения (являются дугами графа).

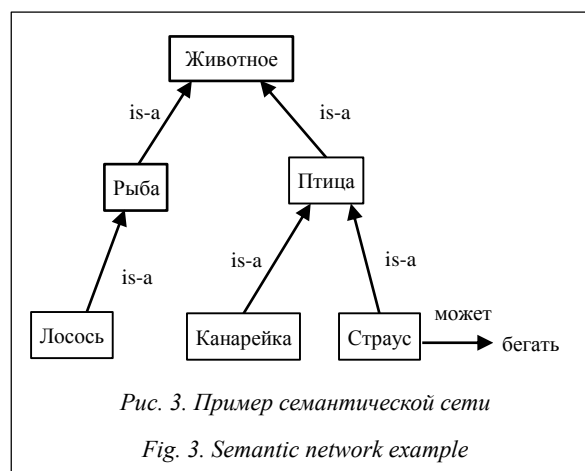


Рис. 3. Пример семантической сети

Fig. 3. Semantic network example

### Вывод в графических моделях

Учитывая вероятностную природу данных, получим слабоформализованную БЗ (здесь в существовании связей между элементами БЗ нет уверенности, поскольку данные получены автоматически). Представим полученную БЗ в виде взвешенного графа. Причем веса в графе обозначают степень уверенности алгоритма в правильности полученных данных. Однако так как данные оказываются неточными, вывод по ним стоит строить соответствующий. Задачу вывода в таких системах решает вероятностный вывод.

Вероятностный вывод – более общий по отношению к логическому выводу, он позволяет оценивать степень уверенности в том, что полученный результат с нечетко определенными данными имеет место быть.

Пусть дана графическая модель  $X$ . Хотим получить ответ на вероятностный запрос вида  $p(y/z)$ , где  $Y$  – целевая переменная, а  $Z$  – входная переменная.  $Y$  и  $Z$  являются непересекающимися подмножествами  $X$ . Эта задача известна как задача вероятностного вывода.

Для объяснения механизма вероятностного вывода приведем пример программы на языке вероятностного программирования. На сегодня существует множество языков вероятностного логического программирования, например PRISM, ICL и ProbLog. В данной работе используется язык ProbLog (Probabilistic Prolog), синтаксис которого схож с наиболее распространенным на данный момент языком Prolog [7].

В программе на языке ProbLog каждому факту присваивается какая-либо вероятность. Далее, как и в Prolog, определяются правила, по которым можно производить вывод. Результатом исполнения программы является оценка вероятности запрошенного факта. Стандартным примером программы на ProbLog является программа Smokers, которая показывает, каким образом отношения между людьми влияют на их привычку курить (табл. 2).

Соответствие программы на языке ProbLog и семантической сети

Таблица 2

Table 2

ProbLog program matching to a semantic network

Семантическая сеть	Программа на ProbLog
	<b>Факты:</b> <code>0.8::stress(ann).</code> <code>0.6::influences(ann,bob).</code> <code>0.2::influences(bob,carl).</code> <b>Правила:</b> <code>smokes(X) :- stress(X).</code> <code>smokes(X) :- influences(Y,X),</code> <code>smokes(Y).</code> <b>Запрос:</b> <code>query(smokes(carl))</code>
	<b>Ответ:</b> <code>probability(smokes(carl))</code> <code>= 0,096</code>

Входным множеством переменных в данном примере является множество фактов:

$Z = \{0.8::\text{stress}(\text{ann}), 0.6::\text{influences}(\text{ann},\text{bob}), 0.2::\text{influences}(\text{bob},\text{carl})\}$ ,  
а целевой переменной  $Y$  является переменная в запросе:  $Y = \{\text{smokes}(\text{carl})\}$ .

Символ `:-` в правилах обозначает импликацию, направленную влево, то есть выражение `«smokes(X) :- stress(X).»` следует читать так: если человек испытывает стресс, то он курит. Каждая строка заканчивается точкой (как в предложении на естественном языке), запятые обозначают совместное применение выражений, то есть «И».

На данный момент задача вероятностного вывода становится наиболее актуальной, так как расчет количества неточных данных и методы логического вывода работают уже не так хорошо. На эту тему существует много работ, в частности, в [8, 9] рассматривается вывод на основе конвертации исходной программы во взвешенную булеву формулу (пропозиционную), а также во взвешенную конъюнктивную нормальную форму. В работе [10] рассматривается вывод в графических моделях и, в частности, в семантических сетях, что наиболее близко к данной работе.

### Сравнение языков вероятностного логического программирования

Сравним ProbLog с другими языками вероятностного программирования, а также с его предшественником – языком Prolog. Ранее было отмечено, что на данный момент существует несколько языков вероятностного вывода: PRISM [11], ICL [12], ProbLog и другие.

ProbLog поддерживает ограниченный набор функций языка Prolog для выражения моделей в вероятностной логике. Главное различие между языком ProbLog and Prolog в том, что Prolog является полноценным языком вероятностного программирования, тогда как ProbLog – язык логического представления. Это означает, что не все функции языка Prolog, связанные с частью программирования (управляющие конструкции и ввод/вывод), не поддерживаются в языке ProbLog.

В таблице 3 представлены основные характеристики каждого языка и основная, соответствующая этому языку система.

По сравнению с другими языками вероятностного логического программирования ProbLog более выразителен по отношению к правилам, доступным в программе. Это, в частности, относится к PRISM и ICL. Эти языки выдвигают требования об ацикличности правил [13]. В языке ProbLog программа может быть зациклена, например: `smokes(X) :- smokes(Y), influences(Y,X)`. Такие типы циклических правил часто используются для задач классификации и анализа социальных сетей [14]. Также PRISM требует правила с унифицированной (все одинаковые) головой, которые должны иметь взаимоисключающие тела (все разные), то есть одновременно может быть истинно только одно такое тело. ProbLog [15, 16] не имеет такого ограничения, таким образом, правила с унифицированной головой могут иметь пересекающиеся тела (где голова

правила – это часть правила до символа «:-», а тело – после). Например, тела двух правил сигнализации пересекаются: и кража, и землетрясение могут заставить сигнализацию сработать, однако они также могут произойти одновременно:

```
сигнализация
0.1::burglary.           person(mary).
0.2::earthquake.        person(john).
0.7::hears_alarm(X) :- person(X).
alarm :- burglary.
alarm :- earthquake.
calls(X) :- alarm, hears_alarm(X).
```

Таблица 3

**Обзор основных характеристик некоторых языков вероятностной логики и соответствующих им систем (реализации)**

Table 3

**Review of key features of some probabilistic logic languages and corresponding implementation systems**

Язык	ProbLog		PRISM	ICL
Система	ProbLog1	ProbLog2	PRISM	AILog2
Циклические правила	+	+	-	-
Правила перекрытия (дублирования)	+	+	-	+
Индуктивные определения	+	+	+	+
Доказательства на произвольных атомах	-	+	-	+
Множественные запросы	-	+	-	-

В заключение отметим, что в работе рассмотрен процесс вероятностного вывода в слабоформализованных БЗ. В качестве базы выбрана семантическая сеть, поскольку в графической модели удобно представлять автоматически извлеченные данные. Также проведено сравнение продукционного подхода с представлением БЗ в виде семантической сети.

В данной статье предложено использовать вероятностный вывод в графической модели и, следовательно, вероятностный язык логического программирования для его реализации. Также срав-

ниваются несколько существующих языков вероятностного логического программирования, в результате для реализации механизма вероятностного вывода выбран язык ProbLog.

### Литература

1. Katerinenko R.S., Bessmertny I.A. A method for acceleration of logical inference in the production knowledge model. Programming and Computer Software, 2011, vol. 37, no. 4, pp. 197–199.
2. Уэно Х., Исидзука М. Представление и использование знаний. М.: Мир, 1989. 220 с.
3. Довбуш Г.Ф. Логическое программирование. ПГУПС, 2011. URL: [http://ivc.clan.su/\\_fr/1/FLP\\_19.pdf](http://ivc.clan.su/_fr/1/FLP_19.pdf) (дата обращения: 28.02.2016).
4. Petasis G., Karkaletsis V., Paliouras G., Krithara A., Zavisanos E. Ontology population and enrichment: state of the art. Proc. 14th Intern. Conf., Philadelphia, Pennsylvania, USA, 2011, vol. 6050, pp. 134–166.
5. Bessmertny I.A. Knowledge Visualization Based on Semantic Networks. Programming and Computer Software, 2010, vol. 36, no. 4, pp. 197–204.
6. Roussopoulos N.D. A semantic network model of data bases. TR no. 104, DCS, Univ. of Toronto Publ., 1976.
7. Бессмертный И.А. Искусственный интеллект. СПб: Изд-во ИТМО, 2010. 132 с.
8. Fierens D., Van den Broeck G., Renkens J., et. al. Inference and learning in probabilistic logic programs using weighted boolean formulas. DCS, TPLP, 2015, vol. 15, no. 3, pp. 358–401.
9. Fierens D., Van den Broeck G., Thon I., Gutmann B., De Raedt L. Inference in probabilistic logic programs using weighted CNF's. Proc Conf. AUA1, 2011, pp. 211–220.
10. Сукач Е.И., Ратобильская Д.В., Мережа В.Л. Реализация вывода в семантической сети с использованием вероятностно-алгебраического моделирования // Открытые семантические технологии проектирования интеллектуальных систем. OSTIS- 2011: тр. Междунар. науч.-технич. конф. Минск: Изд-во БГУИР, 2011. С. 241–246.
11. Sato T., Kameya Y. PRISM: a language for symbolic-statistical modeling. Intern. Joint Conf. on Artificial Intelligence, 1997, vol. 15, pp. 1330–1339.
12. Poole D. The independent choice logic and beyond. In probabilistic inductive logic programming – theory and applications. LNCS, Springer, Berlin/Heidelberg, 2008, vol. 4911, pp. 222–243.
13. Sato T. and Kameya Y. Parameter learning of logic programs for symbolic-statistical modeling. JAIR, 2001, vol. 15, pp. 391–454.
14. Russell S.J., Norvig P. Artificial intelligence. A modern approach. 2nd ed. Upper Saddle River, NJ, Prentice Hall Publ., 2003, 1110 p.
15. De Raedt L., Kimmig A., and Toivonen H. ProbLog: A probabilistic Prolog and its application in link discovery. IJCAI, 2007, pp. 2462–2467.
16. Fierens D., Broeck G.V., Renkens J., et. al. Inference and learning in probabilistic logic programs using weighted Boolean formulas. Theory and Practice of Logic Programming, 2014, vol. 15, no. 3, pp. 358–401.

DOI: 10.15827/0236-235X.115.010-014

Received 01.03.16

### PROBABILISTIC INFERENCE IN WEAKLY FORMALIZED KNOWLEDGE BASES

(Acknowledgements. The work has been done as a part of research topic no. 615869

“Designing methods of information infrastructure key systems”)

<sup>1</sup>Poleschuk E.A., Postgraduate Student, [epoleschuk@corp.ifmo.ru](mailto:epoleschuk@corp.ifmo.ru)

<sup>1</sup>Platonov A.V., Postgraduate Student, [avplatonov@corp.ifmo.rut](mailto:avplatonov@corp.ifmo.rut)

<sup>1</sup>The National Research University of Information Technologies, Mechanics and Optics, Kronverksky Ave. 49, St. Petersburg, 197101, Russian Federation

**Abstract.** The article introduces the process of probabilistic inference in weakly formalized knowledge bases. Semantic network is chosen as a graphical model of knowledge representation due to a convenient representation of automatically ex-

tracted data as a graph with links. The paper also contains a comparison of widely used production model approach with the proposed one. The authors describe main disadvantages of a production model approach that should be considered for developing such knowledge extraction systems.

The aim of the work is new knowledge extraction from an automatically built knowledge base. Usually logical inference is used to achieve this aim in graphical models. In our case a domain model as well as a process of knowledge building (in particular by automatic or semi-automatic methods) restrict logical inference mechanism, so this algorithm is forced to work in conditions of uncertainty. Thus, standard logical inference algorithms provided for such model become irrelevant.

The article proposes using probabilistic inference for the task and consequently using probabilistic inference programming language. The paper contains a comparison of several modern probabilistic logic programming languages like PRISM, ICL and ProbLog. The authors select a probabilistic logic programming language based on the results of this comparison. To implement probabilistic inference in a weakly formalized knowledge base we have selected ProbLog language (ProbLog2 in particular) that is a probability extension of Prolog.

**Keywords** probabilistic inference, semantic networks, weakly formalized knowledge bases, ProbLog, probabilistic logic programming, data uncertainty, relationships uncertainty.

### References

1. Katerinenko R.S., Bessmertny I.A. A method for acceleration of logical inference in the production knowledge model. *Programming and Computer Software*. 2011, vol. 37, no. 4, pp. 197–199.
2. Ueno Kh., Ishizuka M. *Knowledge Representation and Usage*. 1987 (Russ.ed.: Moscow, Mir Publ., 1989).
3. Dovbush G.F. *Logicheskoe programmirovaniye* [Logical Programming]. PGUPS Publ., 2011. Available at: [http://ivc.clan.su/\\_fr/1/FLP\\_19.pdf](http://ivc.clan.su/_fr/1/FLP_19.pdf) (accessed February 28, 2016).
4. Petasis G., Karkaletsis V., Paliouras G., Krithara A., Zavitsanos E. Ontology Population and Enrichment: State of the Art. J.H. Moore, T. Soule (Eds.). *Proc. 14th int. conf.* Philadelphia, Pennsylvania, USA, 2011, vol. 6050, pp. 134–166.
5. Bessmertny I.A., Knowledge Visualization Based on Semantic Networks. *Programming and Computer Software*. 2010, vol. 36, no. 4, pp. 197–204.
6. Roussopoulos N.D. A semantic network model of data bases. *TR no. 104*. DCS, Univ. of Toronto, 1976.
7. Bessmertny I.A.. *Iskusstvennyy intellekt* [Artificial Intelligence]. St. Petersburg, SPbGU ITMO Publ., 2010, 132 p.
8. Fierens D., Van den Broeck G., Renkens J. *Inference and Learning in Probabilistic Logic Programs using Weighted Boolean Formulas*. DCS, TPLP, 2015, vol. 15, no. 3, pp. 358–401.
9. Fierens D., Broeck G.D.V., Thon I. Inference in Probabilistic Logic Programs using Weighted CNF's. *Proc Conf. AUAI*. 2011, pp. 211–220.
10. Sukach E.I., Ratobylskaya D.V., Merezha V.L. Semantic network inference implementation using probabilistic algebraic modeling. *Otkrytye semanticheskie tekhnologii proektirovaniya intellektualnykh system. OSTIS-2011: Mezhdunar. nauch.-tekhn. konf.* [Proc. Int. Science and Technical Conf. on Open Semantic Technologies for Intelligent Systems (OSTIS-2011)]. Minsk, 2011, BSUIR Publ., pp. 241–246 (in Russ.).
11. Sato T., Kameya Y. PRISM: a language for symbolic-statistical modeling. *Intern. Joint Conf. on Artificial Intelligence*, 1997, vol. 15, pp. 1330–1339.
12. Poole D. The independent choice logic and beyond. *Probabilistic Inductive Logic Programming – Theory and Applications*. De Raedt L., Frasconi P., Kersting K., Muggleton S. (Eds.). LNCS, Springer, Berlin/Heidelberg Publ., 2008, vol. 4911, 222–243.
13. Sato T., Kameya Y. Parameter learning of logic programs for symbolic-statistical modeling. *JAIR*. 2001, vol. 15, pp. 391–454.
14. Russell S.J., Norvig P. *Artificial intelligence: a modern approach*. 2nd ed., Upper Saddle River, NJ, Prentice Hall Publ., 2003.
15. De Raedt L., Kimmig A., Toivonen H. *ProbLog: A probabilistic Prolog and its application in link discovery*. *IJCAI*. 2007, pp. 2462–2467.
16. Fierens D., Broeck G.V., Renkens J. Inference and learning in probabilistic logic programs using weighted Boolean formulas. *Theory and Practice of Logic Programming*. 2014, vol. 15, no. 3, pp. 358–401.

### Примеры библиографического описания статьи

1. Полещук Е.А., Платонов А.В. Использование вероятностного вывода в слабоформализованных базах знаний // Программные продукты и системы. 2016. Т. 29. № 3. С. 10–14.
2. Полещук Е.А., Платонов А.В. Использование вероятностного вывода в слабоформализованных базах знаний // Программные продукты и системы. 2016; DOI: 10.15827/0236-235X.115.010-014.
3. Poleschuk E.A., Platonov A.V. Probabilistic inference in weakly formalized knowledge bases. *Programmnye produkty i sistemy* [Software & Systems]. 2016, vol. 29, no. 3, pp. 10–14 (in Russ.); DOI: 10.15827/0236-235X.115.010-014.

## ОБ ЭФФЕКТИВНОСТИ НАСЛЕДОВАНИЯ ТАБЛИЦ В СУБД PostgreSQL

В.Е. Сорокин, к.т.н., ведущий научный сотрудник, sorokinve@yandex.ru  
(НИИ «Центрпрограммсистем», просп. 50 лет Октября, 3а, г. Тверь, 170024, Россия)

Одним из следствий расширения использования объектно-ориентированного программирования явилось дополнение реляционных СУБД объектными чертами. На кроссплатформенную свободно распространяемую в исходных кодах объектно-реляционную СУБД PostgreSQL возлагаются большие надежды в импортозамещении инфраструктурного ПО. В ней сочетаются классические для реляционных СУБД принципы ACID с поддержкой сложных типов данных и наследованием таблиц, востребованных при моделировании объектов в БД. При работе со встроенным в PostgreSQL сложным типом jsonb достигнута характерная для объектно-ориентированных СУБД производительность. Для случаев предпочтительного использования базовых и композитных типов данных при моделировании в БД объектов реализован механизм наследования таблиц. В статье проведен сопоставительный анализ этого механизма с наиболее известными методами объектно-реляционного преобразования. Планируемая разработчиками PostgreSQL, но еще не реализованная поддержка единой индексации физических таблиц одной иерархии наследования отмечается ими в качестве основного фактора, ограничивающего применение наследования таблиц ввиду отсутствия поддержки в них уникальной и ссылочной целостности и снижения производительности запросов. Предлагаемое в статье в дополнение к существующему механизму наследования таблиц создание одной общей таблицы для иерархии наследуемых таблиц позволяет обеспечить целостность данных. Эксперименты по выполнению различных запросов свойств объектов, моделируемых в БД по классической модели объектно-реляционного преобразования и по предлагаемому дополненному наследованию таблиц, подтвердили высокую эффективность последнего, что позволяет рекомендовать его для широкого применения, особенно с перспективой появления единой индексации иерархии наследуемых таблиц.

**Ключевые слова:** объектно-реляционная БД, СУБД, PostgreSQL, объектно-реляционное преобразование, наследование таблиц, целостность данных, SQL-запрос, эффективность.

Активное развитие в течение нескольких десятилетий *объектно-ориентированного программирования* (ООП) как парадигмы программирования (совокупности идей, понятий и устоявшейся системы научных взглядов, определяющих стиль написания программ) не могло не коснуться СУБД. Это привело как к созданию объектно-ориентированных СУБД, позволяющих работать с объектами БД подобно объектам в языках ООП, так и к дополнению традиционных реляционных СУБД объектно-ориентированными чертами [1, 2]. Разработан широкий спектр средств объектно-реляционного преобразования как промежуточного слоя между объектно-ориентированными программами и реляционными СУБД, например Hibernate и LINQ to SQL [3, 4].

Объектно-ориентированные СУБД, такие как ObjectStore, Jasmine, CouchDB, MongoDB, Caché, предназначены прежде всего для хранения и высокопроизводительной обработки данных сложной структуры, включая произвольные документы (так называемые документо-ориентированные СУБД). Как правило, они обеспечивают долговременное хранение в БД объектов, созданных на определенных языках программирования, таких как C++ и Java. Им присущи как преимущества, так и недостатки, характерные для ООП и объектного моделирования в целом [5]. Принципиальным является отсутствие строгого и общепризнанного определения ООП и, как следствие, соответствующей формальной теории и доказуемости таких важнейших свойств объектной модели, как непротиворечивость и целостность данных [1].

Многие наиболее известные изначально реляционные СУБД, такие как Oracle Database, IBM DB2, Informix, Microsoft SQL Server и PostgreSQL, после дополнения объектно-ориентированными свойствами стали классифицироваться как объектно-реляционные. В их основе лежат разрабатывавшаяся с 1970 года английским математиком Э.Ф. Коддом реляционная теория и сформулированные им в 1985 году знаменитые «12 правил Кодда», которым должна удовлетворять классическая реляционная СУБД [1]. Однако реляционная модель данных не адекватна многим сложным структурам данных и, соответственно, неэффективна в работе с ними прежде всего из-за необходимости выполнения многочисленных соединений JOIN в SQL-запросах. Считается, что реляционные СУБД высокопроизводительны на глобальных запросах по большим участкам БД, а объектно-ориентированный доступ более эффективен с малыми объемами данных, в том числе сложноструктурированных. Для работы с такими данными объектно-реляционные СУБД поддерживают некоторые технологии, реализующие объектно-ориентированный подход: объекты, классы и наследование реализованы в структуре БД и языке запросов. Объем и эффективность таких решений существенно отличаются в различных СУБД.

Судя по представительскому составу докладчиков конференции по СУБД PostgreSQL [6] и участников круглого стола другой аналогичной конференции [7], в России наблюдается устойчивый рост ее популярности и большое внимание к ней со стороны органов государственной власти и госкорпо-



раций. Минкомсвязи России издал 2 июня 2015 года «Протокол экспертной оценки проектов по импортозамещению инфраструктурного программного обеспечения», в котором проект «СУБД PostgreSQL и связанные с ним решения» занимает 1-е место по направлению «Системы управления базами данных». Все вышеизложенное подтверждает актуальность рассмотрения объектных черт данной СУБД.

### PostgreSQL как объектно-реляционная СУБД

В качестве классической реляционной СУБД PostgreSQL соответствует стандартам ANSI SQL-92 и SQL-99, начиная с которого определяется наследование типов, поддерживает многие из возможностей стандарта SQL:2011, а также реализует многочисленные собственные расширения и предоставляет широкие возможности пользовательского расширения. Она классифицируется как кроссплатформенная, имеющая реализации для множества UNIX-подобных платформ и Microsoft Windows, свободно распространяемая в исходных кодах (лицензии PostgreSQL и BSD) объектно-реляционная СУБД. В ней полноценно соблюдаются принципы ACID: atomicity (атомарность) – вносимые транзакцией в БД изменения полностью принимаются либо полностью отвергаются; consistency (согласованность) – транзакции переводят БД из одного согласованного состояния в другое; isolation (изолированность) – параллельные транзакции не оказывают влияние на результат выполнения транзакции; durability (надежность) – аппаратные сбои не приводят к потере данных [8, 9].

Ее основными объектными чертами являются поддержка наряду с базовыми типами данных композитных типов данных, массивов и сложных типов (hstore, xml, json, jsonb), а также наследование таблиц (несколько отличающееся от стандартного наследования типов) и возможность создания методов благодаря синтаксической эквивалентности записей вида tab.func и func(tab). На механизме наследования таблиц реализовано секционирование (partitioning) таблиц [8, 9].

Как отмечалось, нормализуемая модель атрибутов сущностей обладает нетривиальностью схемы для сложных структур данных и, как следствие, невысокой эффективностью запросов к ним. Развитие от этой модели в направлении поддержки сложных типов, все более эффективно хранящихся и индексируемых, позволяет PostgreSQL при работе со специфическими для объектно-ориентированных СУБД данными получать соизмеримую эффективность. Последний (на момент написания статьи) встроенный в PostgreSQL сложный тип jsonb [8], определяемый как структурированные декомпозированные двоичные json-данные, дополняет гибкость и универсальность стандартных (ECMA-404

The JSON Data Interchange Standard, JSON RFC-7159) текстовых json-данных, поддерживающих документо-ориентированную модель, возможностью использования GIN-индексов в операторах поиска.

Так, на рассмотренных в [10] примерах показано, что PostgreSQL 9.4 при использовании типа jsonb с GIN-индексами на данных объектной БД, функционирующей под управлением объектно-ориентированной СУБД MongoDB 2.6.0, сопоставима с последней по большинству параметров. Размер таблиц составил 1.3 Гб и 1.8 Гб, суммарно размер индексов 975 Мб и 487 Мб в PostgreSQL и MongoDB соответственно. Время загрузки данных типов json и jsonb в PostgreSQL и объектного типа в MongoDB составило порядка 37 сек., 43 сек. и 13 мин. соответственно. Среднее время выполнения поискового оператора @> содержания – 10 сек. при последовательном сканировании json, 8.5 мс по общему GIN-индексу jsonb, 0.7 мс по GIN-индексу с путями в jsonb и 1.0 мс по B-tree-индексу в MongoDB. Появление языка JQuery-запросов к данным типа jsonb с поисковым оператором @@ соответствия и текстовым типом данных jquery таких запросов с поддержкой GIN-индексов при использовании встроенного простого оптимизатора запросов JQuery позволило PostgreSQL извлекать контрольные данные за время порядка 0.5 сек. по сравнению с 7 сек. их извлечения в MongoDB.

Основными недостатками объектной модели, реализуемой сложными типами данных, по сравнению с нормализуемой моделью атрибутов сущностей, реализуемой базовыми и композитными типами данных, являются ограниченность механизмов поддержания целостности данных, избыточность данных и большая конкурентность при параллельном доступе. Для поддержания целостности данных (спецификации схемы) могут использоваться ограничения, вызывающие функции от данных сложного типа, проверяющие, например, наличие заданного ключа в hstore, правильность сформированного xml документа или заданного пути в json. Наибольшие возможности при этом предоставляет оператор @@ соответствия в языке JQuery-запросов к данным типа jsonb, однако проверочные выражения для сложных типов могут быть весьма нетривиальными. Принципиальным недостатком является отсутствие поддержки ссылочной целостности. Полноценное поддержание целостности данных сложного типа в настоящий момент возможно только созданием наборов соответствующих триггеров и написанием вызываемых ими оригинальных триггерных функций.

Естественно, что отсутствие нормализации в общем случае приводит к значительному дублированию данных в объектной модели. Кроме того, это увеличивает накладные расходы при корректировке таких данных, а также увеличивает вероят-

ность блокировок или конфликтов изменения (при пессимистической или оптимистической стратегии) при параллельном доступе к ним. Например, при хранении в БД информации обо всех жителях большого города как объектов, в которых содержится адрес их проживания, переименование этого города потребует изменения всех этих объектов БД с их одновременной блокировкой на запись. Хранение сложного объекта в одном поле записи БД приводит к невозможности одновременного изменения различными пользователями любых его свойств. Это существенный недостаток сложных типов по сравнению с композитными и тем более базовыми типами данных. Именно поэтому наследование таблиц с композитными и базовыми типами данных в нормализуемой модели атрибутов сущностей может оказаться существенно эффективнее использования сложных типов в интенсивно изменяющихся БД с высоким уровнем конкурентного доступа даже для сложных структур данных.

#### Варианты моделирования объектов реляционными СУБД

Общепринятой технологией программирования, связывающей реляционные БД с концепциями ООП и создающей в реляционной БД «виртуальную объектную» БД, способствуя тем самым преодолению семантического разрыва между объектной и реляционной моделями данных, является объектно-реляционное преобразование (ORM – object-relational mapping). Его основу составляют преобразования коллекций, компонентов и наследников. Для нас интерес представляют преобразования наследников. На примере одной из наиболее популярных ORM-библиотек Hibernate рассмотрим три основные стратегии такого преобразования [3, 11]: одна таблица на подкласс (table per subclass), одна на классовую иерархию (table per class hierarchy) и одна на конкретный класс (table per concrete class).

В первой стратегии для каждого класса в иерархии создаются различные таблицы, имеющие один и тот же столбец идентификатора объекта как первичный ключ, который в каждой таблице подкласса является также внешним ключом ссылки на таблицу его суперкласса. При определении имени класса в HQL-запросе Hibernate автоматически выполняет SQL JOIN для восстановления всей необходимой информации из таблиц иерархии. При неопределенном имени класса в HQL-запросе наряду с запрошенными столбцами формируется дополнительный динамический столбец *clazz*, используемый Hibernate для определения и размещения восстановленного объекта. При получении информации из более чем одной таблицы требуется выполнение соединений JOIN в SQL-запросах. Принципиальным недостатком этой стратегии является

обеспечение уникальности ключа для каждого конкретного класса иерархии на уровне Hibernate, а не БД ввиду отсутствия в таблицах столбца, подобного *clazz*. Это допускает нарушение целостности данных при работе с БД приложений, не использующих библиотеку Hibernate. Например, в такой БД возможны объекты различных конкретных классов одной иерархии, не являющиеся наследниками друг друга и имеющие один и тот же идентификатор.

Во второй стратегии единственная таблица имеет необходимые столбцы для хранения всех свойств классовой иерархии и технический столбец *discriminator*, определяющий класс объекта. При определенном имени класса для получения всей необходимой информации достаточно фильтрации по соответствующему значению столбца *discriminator*. Кроме сильной разреженности хранящихся в такой таблице данных, существенным недостатком является возможная при значительной иерархии большая сложность поддержания целостности БД. На одну таблицу должны быть наложены ограничения, характерные для каждого класса иерархии. Столбцы, определенные в подклассах, не могут быть NOT NULL, что резко сужает использование этого важнейшего механизма спецификации схемы БД.

Третья стратегия наиболее специфична для Hibernate. В ней создается по одной таблице на каждый конкретный класс и не создаются таблицы для абстрактных классов. Вся необходимая информация по каждому классу извлекается из его таблицы. Однако при этой стратегии в БД не существует уникальной таблицы, которую можно использовать для создания ограничения внешнего ключа. Кроме поддержания иерархии классов, на Hibernate возлагается недопущение совпадения идентификационных свойств, распределенных между различными таблицами внутри одной классовой иерархии. С этой целью в таблицы вводятся столбцы дискриминатора и имени столбца первичного ключа другой (проверяемой) таблицы. Стандартный SQL не допускает ограничителей ссылок одновременно из нескольких таблиц в данный столбец. Поэтому возможно лишь создание процедуры, которая проверяет наличие соответствующих данных в проверяемой таблице. Такой метод интеграции модели БД может оказаться очень сложным в применении.

Стратегия *table per class hierarchy* рекомендуется не только в Hibernate, но и во многих других ORM-средствах, например, в реляционном конструкторе объектов LINQ to SQL от Microsoft [4]. В то же время некоторые специалисты считают, что создание в реляционных БД таблиц типов, которые объединяют несколько типов в одной таблице, – плохая проектная идея. В таблицу вводятся сложные ограничения в зависимости от отнесения записи к конкретному типу. Добавляется нежела-

тельная связность управления в БД, обусловленная необходимостью приложений добавить логику управления к выборке SQL в зависимости от значения столбца-дискриминатора, обычно имени типа, просто для различения столбцов, которые относятся к данному типу [12]. Дополнительно отметим, как правило, снижение компактности физического хранения данных таких таблиц, влекущее увеличение требуемой памяти и снижение производительности.

Стратегия *table per subclass* идеологически наиболее близка процедуре реляционной нормализации. Это обусловило ее широкое распространение в различных средствах объектно-реляционного преобразования, в которых, в отличие от Hibernate, как правило, требуется определение столбца-дискриминатора в таблице суперкласса, что позволяет поддерживать целостность БД независимо от такого средства. Для этого вместо создания ограничителей ссылок одновременно из нескольких таблиц в данный столбец с целью поддержания уникальности идентификаторов объектов внутри одной классовой иерархии достаточно использовать первичные ключи и механизм ссылочной целостности (по составному ключу из столбцов идентификатора и дискриминатора).

При простом наследовании, когда класс связан с единственным предком через генерализацию, в [12] сущность стратегии *table per subclass* называют прямым отображением классов, наряду с которым рассматривают отображение классов с помощью распространения атрибутов на подклассы. Распространение наиболее близко стратегии *table per concrete class*, когда вся необходимая информация по каждому классу извлекается из единственной таблицы. Однако при распространении атрибутов ценой простоты получения данных является денормализация схемы БД со значительной избыточностью данных, приводящая к аномалиям включения, обновления и удаления. Поэтому даже в варианте применения с данными одного класса и достаточности доступа к одной таблице при обновлении в ней данных нужно также обновлять те же столбцы в соответствующей строке таблицы суперкласса. Это означает введение дополнительного триггера.

Таким образом, каждая стратегия ORM-преобразования имеет свои сильные и слабые стороны и при различных условиях и целях наиболее эффективными могут оказаться различные стратегии.

#### **Сравнение наследования с объектно-реляционными преобразованиями в PostgreSQL**

Описанные стратегии ORM могут быть положены в основу моделирования объектов в БД под управлением PostgreSQL как реляционной СУБД. Однако в PostgreSQL реализован собственный ме-

ханизм наследования таблиц, не совпадающий полностью ни с одной данной стратегией. В нем, как в первой стратегии, для каждого класса в иерархии создаются различные таблицы и, как в третьей стратегии, вся необходимая информация по каждому классу извлекается из его таблицы. В отличие от первой стратегии в таблице подкласса отсутствует столбец внешнего ключа ссылки на таблицу его суперкласса, а в отличие от третьей стратегии таблицы создаются не только для конкретных, но и для абстрактных классов. Имеется подобный техническому столбцу *discriminator* определяющий класс объекта записи системный столбец *tableoid* типа объектного идентификатора, в настоящий момент технически поддерживаемый беззнаковым четырехбайтовым целым. В отличие от второй стратегии в таблицах отсутствуют избыточные столбцы для хранения свойств объектов других классов [8]. По данным свойствам можно считать реализацию наследования таблиц в PostgreSQL на уровне самых эффективных стратегий ORM.

Поддержание иерархии наследования таблиц возлагается на СУБД и определяется в спецификации схемы БД, что делает защиту от нарушения целостности данных в аспекте наследования независимой от среды разработки приложений, работающих с БД. При этом допускается множественное наследование при отсутствии конфликта типов между таблицами суперклассов таблицы их подкласса. Иерархия наследования таблиц поддерживается в SQL-командах запроса и изменения данных, а также модификации схемы БД. Указанием опции *ONLY* в их нотации или заданием конфигурационного параметра совместимости с синтаксисом предыдущих версий можно ограничить их действие только конкретной физической таблицей без поддержки наследования. План выполнения запроса данных из иерархии таблиц представляет собой объединение запросов данных из всех физических таблиц этой иерархии, то есть в общем случае время выполнения запроса линейно зависит от количества таблиц вниз по иерархии относительно указанной в запросе таблицы.

Команды технического обслуживания и настройки БД, такие как *REINDEX* и *VACUUM*, работают с отдельными физическими таблицами и не поддерживают рекурсию по иерархии наследования таблиц. Самым серьезным ограничением текущей реализации наследования таблиц в PostgreSQL является применение индексов только к отдельным физическим таблицам. Следствием этого являются невозможность применения ограничений уникальности и внешних ключей к иерархии таблиц, а также возможное во многих случаях снижение производительности выполнения запросов к иерархии таблиц. Несмотря на предположения в документации нескольких последних версий СУБД PostgreSQL об устранении в будущем данной проблемы, она остается основным фактором,

ограничивающим применение наследования таблиц [8].

Наиболее эффективным средством преодоления указанного недостатка автору представляется создание одной таблицы на иерархию наследования таблиц, содержащей столбцы первичного ключа и класса объекта записи. При наличии в такой таблице записей, соответствующих всем записям таблиц иерархии, она может служить для создания необходимых индексов по иерархии таблиц, включая уникальные, частичные и по выражениям или функциям, для реализации ограничений уникальности и внешних ключей. Триггеры вставки и удаления записей в таблицах иерархии, поддерживающие соответствие записей в такой таблице, должны быть максимально просты, а изменения записей в ней запрещены. С этой целью достаточно придерживаться неизменных первичных ключей, используя неявную идентификацию по сгенерированным ключам, когда имеется потенциальный ключ, который может измениться, поскольку неизменная явная идентификация возможна далеко не всегда. Класс объекта записи неизменен. Для гарантии целостности таких данных требуется обеспечить неизменность значений отдельных столбцов таблиц. Для этого необходимо создать триггеры изменения записей в таблицах иерархии, контролирующие неизменность значений этих столбцов. Поддержание неизменности значений столбцов, как NOT NULL, так и NULL, после определения значения достаточно распространено, и автор считает целесообразным включение в SQL соответствующего ограничения CONST наряду с такими ограничениями, как NOT NULL и CHECK.

Разумеется, предлагаемая общая таблица на иерархию наследования таблиц может содержать столбцы не только первичного ключа и класса объекта записи, что может позволить реализовать более сложную целостность данных или бизнес-логику. Оценка целесообразности таких решений подобна созданию отдельных таблиц для реализации нетривиальных ограничений, накладываемых на БД в целом (утверждений ASSERTION), определяется спецификой каждого конкретного случая и выходит за рамки предмета статьи.

Для оценки эффективности реализованного механизма наследования таблиц в СУБД PostgreSQL производилось сравнение среднего времени выполнения аналогичных запросов свойств около 10 моделируемых объектов в двух БД, условно называемых наследуемой и реляционной. Первая БД (наследуемая) спроектирована как иерархия наследуемых таблиц с одной добавленной таблицей, содержащей столбцы первичного ключа и класса объекта, как предложено выше. Во второй БД (реляционной) подобно стратегии table per subclass воплощена классическая реляционная нормализация. Свойства объектов запрашиваются по интервалам значений как индексируемых, так и неиндексируе-

мых столбцов, по объектам как только конкретных классов, так и с учетом их подклассов. Используется простая модель трех уровней иерархии наследования из одного суперкласса и трех его подклассов, у каждого из которых также по три подкласса. Объект любого из классов имеет уникальный целочисленный идентификатор и равное уровню его иерархии количество целочисленных свойств – наследуемые и одно собственное – для максимального упрощения модели и исключения влияния факторов, не зависящих от принципов проектирования БД.

Реализация классической реляционной модели заключается в создании отдельной таблицы для каждого класса, содержащей столбцы идентификатора *id*, дискриминатора *dtr* и собственного целочисленного свойства. Идентификатор является первичным ключом, пара (*id*, *dtr*) уникальна. На дискриминатор накладывается ограничение в виде списка значений дискриминатора собственного класса и всех его дочерних подклассов. Наследование моделируется внешними ключами в виде пары (*id*, *dtr*). Все свойства конкретного объекта содержатся в связанных по внешним ключам записях таблиц (от 1 до 3 таблиц в зависимости от уровня иерархии класса объекта).

Реализация иерархии наследуемых таблиц заключается в создании отдельной таблицы для каждого класса, между которыми установлено соответствующее наследование. Таблица супертупа содержит столбцы идентификатора *id*, являющегося первичным ключом, дискриминатора *dtr* и собственного целочисленного свойства. Остальные таблицы, кроме указания на таблицу, от которой они наследуются, и наследования ограничения первичного ключа, содержат столбец собственного целочисленного свойства, в результате чего все свойства конкретного объекта содержатся в записи его таблицы. Для обеспечения такой же целостности данных, как в классической реляционной модели, а именно уникальности идентификатора во всей иерархии таблиц и, соответственно, возможности внешней ссылки сразу на определенное подмножество этих таблиц, создается дополнительная таблица. Эта таблица имеет столбцы идентификатора *id*, являющегося первичным ключом, и дискриминатора *dtr* и содержит записи со всеми значениями пары (*id*, *dtr*), являющейся уникальной, из всех таблиц иерархии, в которые добавляется внешний ключ ссылки на эту таблицу в виде пары (*id*, *dtr*).

В ходе эксперимента в обеих описанных БД, установленных на системе с процессором Intel Core i5-4430 3.0 GHz и 16 GB оперативной памяти с настройками СУБД PostgreSQL 9.5 по умолчанию, последовательно моделировалось с шагом в 1 порядок от 10 тысяч до 100 миллионов идентичных объектов, с примерно одинаковым количеством объектов во всех классах. Для моделирования

10 тысяч объектов в наследуемой БД потребовалось 20 тысяч записей, занимающих 108 страниц дисковой памяти, а в реляционной БД – соответственно около 26 тысяч записей на 147 страницах. Для моделирования 100 тысяч объектов – 200 тысяч записей на 1 052 страницах в наследуемой БД и около 262 тысяч записей на 1 420 страницах в реляционной БД. Дальнейший рост значений этих показателей пропорционален количеству моделируемых объектов и для 100 миллионов объектов составляет 200 миллионов записей на 1 050 тысячах страниц в наследуемой БД и около 262 миллионов записей на 1 414 тысячах страниц в реляционной БД. Принципиально важным является потабличное распределение количества записей и дисковой памяти. Они практически одинаковы для таблиц свойств объектов нижнего уровня иерархии в реляционной БД и всех таблиц (кроме дополнительной, созданной исключительно для поддержания целостности данных и не задействованной в запросах) в наследуемой БД. По сравнению с ними в 4 раза больше количество записей и занимаемой дисковой памяти для таблиц свойств объектов среднего уровня иерархии и еще почти в 4 раза для таблицы свойств объектов верхнего уровня иерархии в реляционной БД.

На каждом шаге роста количества моделируемых объектов на обеих БД выполнялось по несколько запросов, возвращающих свойства около 10 объектов, каждого из 10 видов запросов. Отметим, что выборочно выполнявшийся возврат свойств другого количества объектов (от 1 до 100) на время выполнения запросов принципиально не влиял. В запросе 1-го вида возвращались свойства объектов конкретного (в данном эксперименте единственного) класса верхнего уровня иерархии, выбираемых по диапазону значений индексируемого столбца, отфильтровывающему около 10 объектов. В качестве индексируемых столбцов во всех запросах использовались идентификаторы. В запросе 3-го вида в отличие от запроса 1-го вида возвращались свойства объектов конкретного класса среднего уровня иерархии, 5-го вида – конкретного класса нижнего уровня иерархии, 7-го вида – любого класса иерархии (от конкретного единственного класса верхнего уровня иерархии), 9-го вида – любого класса иерархии от конкретного класса среднего уровня иерархии. Запросы одного вида могут быть реализованы несколькими различающимися по синтаксису запросами. Все запросы четного вида отличаются от предшествующего запроса нечетного вида только тем, что для фильтрации вместо индексируемого столбца используется неиндексируемый столбец.

На рисунке 1 приведены полученные в результате эксперимента графики зависимостей среднего времени выполнения запроса свойств 10 объектов, выбираемых по диапазону значений индексируемого столбца, от общего количества объектов, мо-

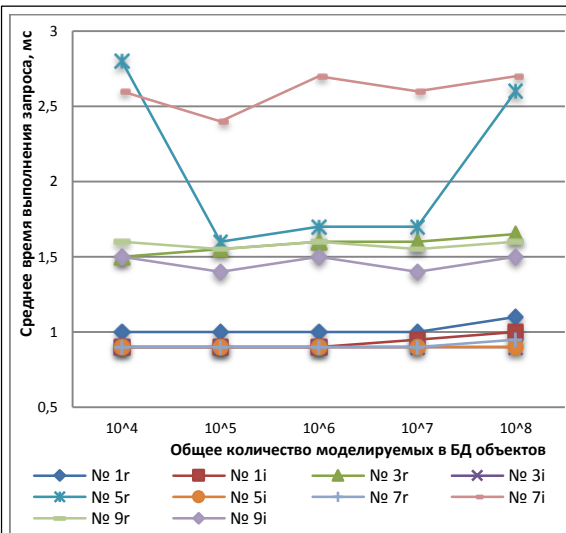


Рис. 1. Выполнение запросов по значениям индексируемого столбца

Fig. 1. Query execution on indexing column values

делируемых в БД, для различных нечетных видов запросов. Номер графика состоит из вида запроса с символом *r* для реляционной БД и *i* для наследуемой БД.

Главная особенность представленных на рисунке 1 графиков – в целом несущественная зависимость времени выполнения запросов от количества записей в запрашиваемых таблицах. Прежде всего это объясняется максимально широким применением индексов в планах выполнения запросов. При этом обращает на себя внимание существенное сначала снижение, а затем увеличение времени выполнения запроса в реляционной БД свойств объектов конкретного класса нижнего уровня иерархии (№ 5r) с увеличением количества записей в таблицах. В запросах этого вида в реляционной БД выполняются соединения 3 таблиц по идентичным значениям ключей и фильтрация по ключевым столбцам. В них возможны различный порядок соединения таблиц и выбор таблицы, по значению ключевого столбца которой выполняется фильтрация. В этом смысле запросы данного вида наиболее вариативны, и в эксперименте специально использовались различные их реализации при различных количествах моделируемых объектов.

Проведенный с помощью команды EXPLAIN отображения плана выполнения запроса анализ выбранных планировщиком-оптимизатором методов доступа и соединения данных показал, что наиболее эффективные реализации запроса № 5r соответствуют основному принципу: чем к меньшему количеству данных осуществляется обращение, тем быстрее выполняется запрос. В исполнение этого принципа в них наибольшая часть ненужных строк отсекалась на самых ранних стадиях выполнения запроса, а соединение таблиц осуществлялось в направлении уменьшения их селективности (уве-

личения числа отфильтрованных строк к общему числу строк). Максимально используется возможность получения данных прямо из индекса без обращения к таблице, поскольку количество страниц памяти, занимаемое индексами, существенно меньше, чем занимаемое таблицами [13]. В наименее эффективных реализациях запроса № 5г к нарушению этого принципа приводили фильтрация строк не по таблице с наибольшим их количеством и соединение таблиц не в направлении уменьшения их селективности. При выполнении наиболее эффективных реализаций запроса № 5г для всех количеств моделируемых в БД объектов график запроса № 5г располагается над графиком запроса № 3г подобно их средней части.

Наблюдаемое на большинстве графиков незначительное увеличение времени выполнения запроса при уменьшении количества моделируемых в БД объектов объясняется выбором не самого оптимального плана выполнения запроса, несмотря на поддерживаемую актуальность статистики, обновляемой командой ANALYZE. В этих случаях, как правило, во вложенных циклах выполнялось соединение по хэшу с предшествующим индексным сканированием вместо несколько более эффективных двойных вложенных циклов с индексным сканированием. Подобные неточности в работе построителя оптимального плана запроса иногда случаются. Как правило, их причина в неактуальности, недостаточной детализации статистики (устанавливается одной из форм команды ALTER TABLE) или неточности стоимостных параметров настройки сервера.

В данном случае незначительная неоптимальность выбранных планов выполнения некоторых запросов качественно не влияет на полученный результат. Время выполнения всех запросов, кроме запроса свойств любого класса иерархии, в наследуемой БД не больше, чем в реляционной. При этом время выполнения запросов свойств объектов конкретного класса в наследуемой БД меньше, чем в реляционной, и эта разница увеличивается с понижением уровня иерархии класса. Последнее объясняется увеличением количества соединений таблиц в запросах в реляционной БД и обращением к единственной таблице в наследуемой БД. В то же время свойства любого класса иерархии, то есть общие свойства всех классов, в реляционной БД хранятся в одной таблице, а в наследуемой – во всех таблицах, кроме дополнительной. Этим объясняется большая эффективность реляционной БД именно в запросе свойств любого класса иерархии, поскольку при суммарно одинаковом количестве записей как в единственной, так и в нескольких таблицах время поиска по индексу в каждой из таблиц соизмеримо и в нескольких таблицах практически линейно зависит от количества таблиц. Однако ожидаемая поддержка единых индексов для физических таблиц иерархии наследуемых таб-

лиц [8] делает и эти запросы в наследуемой БД не менее эффективными, чем в реляционной. Кроме того, если в выражении фильтра запроса свойств любого класса иерархии в наследуемой БД присутствуют данные, ограничивающие класс иерархии, перечень таблиц, к которым происходит обращение, может быть ограничен установлением значения `partition` для параметра `constraint_exclusion`. Этот параметр определяет, учитывает ли планировщик табличные проверочные ограничения для оптимизации запросов. При этом из плана выполнения запроса исключаются обращения к таблицам, проверочные ограничения СHECK которых противоречат условиям запроса. Постоянное выполнение таких проверок часто снижает эффективность выполнения запроса из-за увеличения времени построения плана его выполнения. Но, поскольку эти ограничения типичны для наследуемых и секционированных таблиц и запросов к ним, значение `partition` параметра `constraint_exclusion`, при котором проверки выполняются только для дочерних наследуемых таблиц и UNION ALL подзапросов, можно считать оптимальным при наследовании таблиц [8]. Индексы по соответствующим столбцам для выполнения таких проверок не нужны, поскольку в них не используются значения столбцов в записях таблиц.

В тех случаях, когда необходимо сократить время выполнения запроса, возможны следующие действия. Прежде всего требуется учитывать многоверсионность записей в PostgreSQL, которая может приводить к такой замусоренности таблиц, что даже таблица с небольшим количеством записей будет сканироваться заметное время. Поэтому выполнение команды VACUUM после значительного изменения данных крайне целесообразно. Практически любой запрос может выполняться по нескольким различным планам. Команда EXPLAIN отображения плана выполнения запроса позволяет проанализировать выбранные планировщиком-оптимизатором методы доступа, соединения и агрегации данных в зависимости от заданных стоимостных констант планировщика, разрешающих или запрещающих различные методы доступа, и прочих конфигурационных параметров сервера. Если известно, что существует более эффективный план выполнения запроса, можно постараться добиться его выбора планировщиком соответствующими перестроением текста запроса и/или изменениями констант и параметров. Изменить метод доступа или метод соединения прежде всего можно, запретив выбранный планировщиком метод. Например, если вместо более эффективного в каком-то случае индексного сканирования планировщик предпочел последовательное сканирование, то достаточно установить значение параметра `enable_seqscan` в `off` при установленном в `on` значении параметра `enable_indexscan`. Если известен более эффективный порядок соединения, необходимо переписать



запрос с заданным скобками порядком соединения и установить для параметра `join_collapse_limit` значение 1 [13].

На рисунке 2 приведены аналогичные графики для объектов, выбираемых по диапазону значений неиндексируемого столбца, то есть для четных видов запросов.

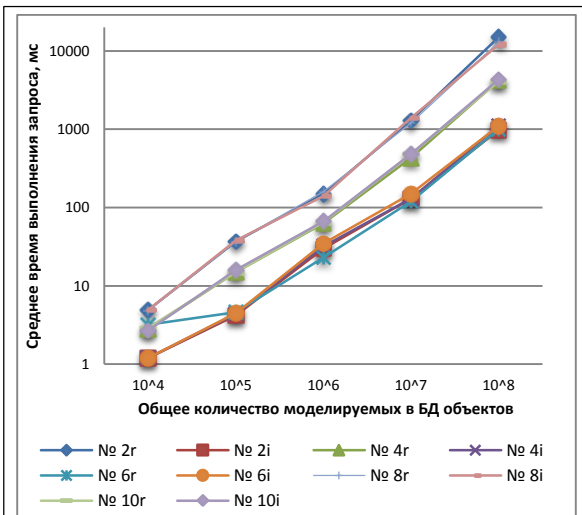


Рис. 2. Выполнение запросов по значениям неиндексируемого столбца

Fig. 2. Query execution on non-indexing column values

Практически линейный характер зависимости времени выполнения запроса от количества записей при фильтрации по неиндексируемым столбцам для всех видов запросов объясняется необходимостью выполнения последовательного сканирования таблиц. Обычные для рисунка 1 времена выполнения запросов достижимы для запросов с фильтрацией по неиндексируемым столбцам только при небольшом (не превышающем сотни тысяч) количестве записей. Это подтверждает практическую невозможность эффективной работы с любой БД без индексации таблиц с большим количеством записей и объясняется тем, что, чем к меньшему количеству данных осуществляется обращение, тем быстрее выполняется запрос. Количество страниц памяти, занимаемое индексами, как правило, существенно меньше, чем занимаемое таблицами, поэтому, если это возможно, следует получать данные прямо из индекса, без обращения к таблице. Многоколоночные (составные) индексы позволяют получать непосредственно из индекса большее количество данных и осуществлять одновременно доступ и сортировку по их ключам [13]. Для повышения эффективности использования составных индексов при выполнении SQL-запросов применяется совместная оптимизация логического выражения раздела выборки `WHERE` и списка выражений раздела сортировки `ORDER BY` команды запроса `SELECT` со списком выражений составного индекса [14].

Однако при небольшом количестве записей фильтрация по неиндексируемым столбцам может быть достаточно эффективной. При этом наблюдаются следующие закономерности. Время выполнения запросов свойств конкретного класса нижнего уровня иерархии, любого класса иерархии от конкретного класса среднего уровня иерархии и любого класса иерархии (от конкретного единственного класса верхнего уровня иерархии) в реляционной и наследуемой БД практически одинаково и существенно растет в порядке перечисления самих запросов. В наследуемой БД в несколько раз быстрее, чем в реляционной, выполняется запрос свойств конкретного класса среднего уровня иерархии и почти на порядок быстрее выполняется запрос свойств конкретного класса верхнего уровня иерархии. Во всех этих случаях определяющее влияние на время выполнения запросов оказывает количество последовательно сканируемых записей. Запрос любого из рассматриваемых видов выполняется в наследуемой БД практически не медленнее, чем в реляционной, а некоторых – существенно быстрее.

Таким образом, текущая реализация механизма наследования таблиц в СУБД PostgreSQL позволяет (при использовании для поддержания целостности данных предлагаемого проектного решения на основе создания для иерархии наследования таблиц одной таблицы со столбцами первичного ключа и класса объекта записи) получить эффективность работы с БД, на большинстве видов запросов не уступающую или превосходящую эффективность стратегий объектно-реляционных преобразований. Несомненно, что планируемая разработчиками СУБД поддержка создания единых индексов для физических таблиц иерархии наследуемых таблиц еще более упростит проектирование БД и повысит эффективность работы с ними. Простой переход от описанного проектного решения к наследованию таблиц с едиными индексами служит обоснованием рекомендации предложенного, основанного на текущей реализации наследования таблиц проектного решения при моделировании объектов в БД, работающих под управлением СУБД PostgreSQL.

#### Литература

1. Дейт К.Дж. Введение в системы баз данных; [пер. с англ.]. М.: Вильямс, 2005. 1328 с.
2. Кузнецов С.Д. Три манифеста баз данных: ретроспектива и перспективы. CIT Forum. URL: <http://www.citforum.ru/database/articles/manifests/> (дата обращения: 15.03.2016).
3. Hibernate упрощает преобразование наследования. Изучение трех упрощенных (для выполнения) стратегий для отображения классовых иерархий. IBM developerWorks. URL: <http://www.ibm.com/developerworks/ru/java/library/j-hibernate/> (дата обращения: 27.04.2016).
4. Наследование классов данных (реляционный конструктор объектов). MSDN Library. URL: <https://msdn.microsoft.com/ru-ru/library/bb531247.aspx> (дата обращения: 27.04.2016).
5. Буч Г., Максимчук Р.А., Энгл М.У., Янг Б.Дж., Коналден Дж., Хьюстон К.А. Объектно-ориентированный анализ и проектирование с примерами приложений (UML 2). Объектная



модель. М.: Вильямс, 2010. 720 с.; URL: [www.williams-publishing.com/PDF/978-5-8459-1401-9/part.pdf](http://www.williams-publishing.com/PDF/978-5-8459-1401-9/part.pdf) (дата обращения: 25.04.2016).

6. Вторая Рос. конф. по вопр. разработки и эксплуатации PostgreSQL: PG Day'15 Russia. СПб., 2015. URL: <http://pgday.ru/ru/2015/schedule> (дата обращения: 05.04.2016).

7. Междунар. конф. разработчиков и пользователей PostgreSQL: PgConf.Russia 2016. М., 2016. URL: <https://pgconf.ru/2016> (дата обращения: 05.04.2016).

8. PostgreSQL 9.5.0 Documentation. The PostgreSQL Global Development Group. 1996–2016. URL: <http://www.postgresql.org/docs/9.5/postgresql-9.5-A4.pdf> (дата обращения: 20.01.2016).

9. Коротков А. Концепции PostgreSQL. PG Day'15 Russia. URL: <http://pgday.ru/files/papers/12/pgday.2015.alexander.korotkov.pg.concepts.pdf> (дата обращения: 05.04.2016).

10. Коротков А., Сигаев Ф. Авторский взгляд на слабо-структурированные данные в PostgreSQL. PG Day'15 Russia. URL: <http://pgday.ru/files/papers/33/pgday.2015.korotkov.sigaev.nosql.pdf> (дата обращения: 05.04.2016).

11. HIBERNATE – Relational Persistence for Idiomatic Java. URL: <http://samsongh.com/lib/hibernate/> (дата обращения: 27.04.2016).

12. Мюллер Р. Дж. Базы данных и UML. Проектирование. М.: ЛОРИ, 2002. 420 с.

13. Фролков И. Оптимизация запросов в PostgreSQL. PgConf.Russia 2016. URL: <https://pgconf.ru/media/2016/02/19/Фролков.pdf> (дата обращения: 05.04.2016).

14. Сорокин В.Е. Метод искусственного соответствия SQL-запросов индексам реляционных баз данных // Программные продукты и системы. 2013. № 2. С. 47–54.

DOI: 10.15827/0236-235X.115.015-023

Received 25.05.16

### ABOUT TABLE INHERITANCE EFFECTIVENESS IN PostgreSQL DBMS

<sup>1</sup>Sorokin V.E., Ph.D. (Engineering), Leading Researcher, [sorokinve@yandex.ru](mailto:sorokinve@yandex.ru)

<sup>1</sup> (R&D Institute Centerprogramsystem, 50 let Oktyabrya Ave. 3a, Tver, 170024, Russian Federation)

**Abstract.** One of the consequences of extending the use of object-oriented programming was the addition of a relational DBMS the object features. On cross-platform open source object-relational DBMS PostgreSQL hopes in import substitution infrastructure software. It combines classic relational database ACID principles with complex data types and table inheritance support that are in demand in the simulation objects in the database. When working with the built-in complex type jsonb PostgreSQL is reached object-oriented DBMS performance. For the case of the preferred use of the base and composite data types for modeling objects in the database table inheritance is implemented. The article presents a comparative analysis of this mechanism with the most known methods of object-relational mapping. PostgreSQL developers planned but not yet implemented support single indexing for physical tables of one inheritance hierarchy is marked by them as the main factor restricting the use of inheritance tables due to the lack of support in unique and referential integrity and reduce the performance of the queries. The proposed in addition to existing table inheritance mechanism creating one common table for the hierarchy inherited tables allows you to ensure data integrity. Performed experiments on the execution different queries the properties of the objects modeled in the DB by the classical models of object-relational mapping and the proposed amended inheritance tables confirmed the high effectiveness of the latter, which allows to recommend it for widespread use, especially with the prospect of a single indexing hierarchy of inherited tables.

**Keywords:** object-relational database, database management system, PostgreSQL, object-relational mapping, table inheritance, data integrity, SQL-query, effectiveness.

### References

1. Date C.J. *An Introduction to Database Systems*. 8-th ed. Addison Wesley Publ., 2004, 1328 p.
2. Kuznetsov S.D. *Tri manifesta baz dannykh: retrospektiva i perspektivy* [The three databases manifests: retrospection and prospects]. CIT Forum. Available at: <http://www.citforum.ru/database/articles/manifests/> (accessed Mart 15, 2016).
3. Hibernate simplifies inheritance mapping. *IBM developerWorks*. Available at: <http://www.ibm.com/developerworks/java/library/j-hibernate/> (accessed April 27, 2016).
4. Data Class Inheritance (O/R Designer). *MSDN Library*. Available at: <https://msdn.microsoft.com/ru-ru/library/bb531247.aspx> (accessed: April 27, 2016).
5. Booch G., Maksimchuk R.A., Engle M.W., Young B.J., Conallen J., Houston K.A. *Object-Oriented Analysis and Design with Applications. Ch. 2. The Object Model*. Addison Wesley Longman Publ., 2004. Available at: [www.williams-publishing.com/PDF/978-5-8459-1401-9/part.pdf](http://www.williams-publishing.com/PDF/978-5-8459-1401-9/part.pdf) (accessed April 25, 2016).
6. *Proc. 2nd Official Russ. Conf. Devoted to Problems of Development and Operations of PostgreSQL. PG Day'15 Russia*. St. Petersburg, July 15–17, 2015. Available at: <http://pgday.ru/en/2015/schedule> (accessed April 05, 2016) (in Russ.).
7. *Proc. Intern. PostgreSQL Conf. PgConf. Russia 2016*. Moscow, 3–5 February 2016. Available at: <https://pgconf.ru/en/2016> (accessed April 05, 2016) (in Russ.).
8. The PostgreSQL Global Development Group. *PostgreSQL 9.5.0 Documentation*. 3064 p. Available at: <https://www.postgresql.org/docs/9.5/static/> (accessed January 20, 2016).
9. Korotkov A. PostgreSQL concepts. *Proc. PG Day'15 Russia*. Available at: <http://pgday.ru/files/papers/12/pgday.2015.alexander.korotkov.pg.concepts.pdf> (accessed April 05, 2016) (in Russ.).
10. Korotkov A., Sigaev F. An author's view on loosely structured data in PostgreSQL. *Proc. PG Day'15 Russia*. Available at: <http://pgday.ru/files/papers/33/pgday.2015.korotkov.sigaev.nosql.pdf> (accessed April 05, 2016) (in Russ.).
11. HIBERNATE – Relational Persistence for Idiomatic Java. Available at: <http://samsongh.com/lib/hibernate/> (accessed April 27, 2016).
12. Muller R.J. *Database design for smarties. Using UML for data modeling*. Morgan Kaufmann Publ., 1999, 442 p.
13. Frolov I. *Optimizatsiya zaprosov v PostgreSQL* [PostgreSQL query optimization]. PgConf.Russia 2016. Available at: <https://pgconf.ru/media/2016/02/19/%D0%A4%D1%80%D0%BE%D0%BB%D0%BA%D0%BE%D0%B2.pdf> (accessed April 05, 2016) (in Russ.).
14. Sorokin V.E. A method of artificial matching of SQL query to relational databases indexes. *Programmnye produkty i sistemy* [Software and Systems], 2013, no. 2, pp. 47–54 (in Russ.).

УДК 519.816

Дата подачи статьи: 10.02.16

DOI: 10.15827/0236-235X.115.024-028

## **О НЕКОТОРЫХ ПРОБЛЕМАХ ПРЕДМЕТНОЙ ОБЛАСТИ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ**

*О.В. Тиханыхев, к.т.н., старший научный сотрудник, tow65@yandex.ru  
(27 Центральный научно-исследовательский институт Минобороны России,  
1-й Хорошевский пр-д, 5, г. Москва, 123007, Россия)*

В статье проанализировано существующее состояние описания предметной области систем поддержки принятия решений относительно как подходов к классификации таких систем, так и наличия их официального определения. Сделан вывод о том, что сложившаяся ситуация замедляет процессы создания и внедрения в практику управления систем поддержки принятия решений. С целью упорядочения предметной области проведено формализованное описание процесса принятия управленческих решений и введены понятия малого и большого циклов принятия решения. При этом под малым циклом понимается поиск оптимального решения в условиях заданной целевой функции и ограничений, под большим – поиск нетривиального решения, которое невозможно получить без модификации целевой функции или снятия части ограничений. На основе данного подхода предложено уточнить определение автоматизированной системы принятия решений и классификацию таких систем. В части уточнения классификации основным показателем в работе является деление по функционалу относительно реализуемого системой цикла принятия решения. В рамках новой классификации предложено взамен применяемого в настоящее время класса «относительно уровня пользователя» использовать класс «по функционалу», в состав которого включить информационные, расчетно-информационные и интеллектуальные системы. Данная классификация более полно соответствует современному пониманию автоматизированной системы поддержки принятия решений и позволит ускорить процесс внедрения таких систем в практику управления.

**Ключевые слова:** система поддержки принятия решений, экспертная система, классификация систем поддержки принятия решений, система информационной поддержки, интеллектуальная система, цикл принятия решения.

Повышение эффективности управления является одной из важнейших задач с точки зрения достижения успеха в любых сферах человеческой деятельности. Именно управление объединяет в единую систему разнородные средства и ресурсы, обеспечивая их рациональное применение в едином цикле. Типовой цикл управления, описывающий общий подход к организации этого процесса, включает этапы целеполагания, оценки имеющихся возможностей и условий обстановки, принятия решения и разработки плана действий, формирования управляющих воздействий (команд), контроля выполнения задач.

Выбор цели действий (этап целеполагания) обычно возлагается на человека и в связи с его неформализованностью математическими методами и техническими средствами практически не поддерживается. При оценке возможности достижения цели и выработке рациональных управляющих воздействий в рамках реализации цикла управления обычно используются специализированные средства и методы поддержки принятия решений.

При этом степень использования и типаж этих средств и методов могут быть различными: от применения простейших субъективных подходов, использующих, например, коллективные методы принятия решений, до объективных методов, базирующихся на точных математических методиках. Спектр программных средств, реализующих объективные математические методы поддержки принятия решений, достаточно широк: от отдельных программ обработки информации и решения расчетных задач до автоматизированных систем под-

держки принятия решений (СППР) и экспертных систем (ЭС).

Наиболее перспективным подходом специалистами признается использование в процессе управления программных средств, реализующих принципы СППР и ЭС. Эти системы в той или иной степени основаны на использовании знаний экспертов.

ЭС относятся к числу интеллектуальных вычислительных систем и являются практической реализацией методов и идей искусственного интеллекта. Такие системы представляют собой программы для ЭВМ, решающие задачи примерно так же, как решает их эксперт в реальной обстановке. Это позволяет накапливать, систематизировать и использовать знания и профессиональный опыт экспертов, которые выполняют конкретные задачи наилучшим образом и в первую очередь в тех областях, где задачи и их решения слабо формализованы или совсем не формализованы.

Компьютерные системы аналогичного назначения – СППР (Decision Support System – DSS) – предназначены для оказания помощи пользователям (управленческому персоналу, аналитикам, лицам, ответственным за принятие решений на различных уровнях) в неструктурируемых или слабо структурируемых ситуациях выбора [1]. Такие системы выступают в роли помощника, позволяющего расширить способности человека, но не заменяющего его мнение или систему предпочтений.

ЭС и СППР как конкретные программные продукты в ряде случаев внешне могут выглядеть одинаково. Однако они имеют различие в своей це-

левой направленности: СППР призваны помочь человеку в решении стоящей перед ним проблемы, а ЭС – заменить человека при решении проблемы.

В областях деятельности, где предполагается высокая личная ответственность *лица, принимающего решения* (ЛПР), например, в системах государственного и военного управления, основным средством являются именно СППР, предоставляющие руководителям больше возможностей по формированию качественного решения.

Несмотря на важность организации поддержки принятия решений и на наличие широкого спектра средств решения этой задачи, данная проблема до сих пор актуальна. Причин этому несколько, но основной из них, на взгляд автора, является неопределенность предметной области: отсутствие узаконенного определения СППР и адекватной классификации таких систем. Эти причины обуславливают проблемы при задании и выполнении работ по созданию и внедрению СППР.

### Анализ состояния предметной области

В настоящее время наиболее часто используется определение СППР из работы [1]: «СППР – это интерактивные автоматизированные системы, помогающие лицу, принимающему решения, использовать данные и модели для решения слабоструктуризованных проблем».

Данное определение является достаточно точным, но очень общим. Под него в той или иной степени подойдет практически любая АСУ, и этим могут пользоваться недобросовестные разработчики программных продуктов, выдавая их за СППР. Проблему усугубляет и то, что даже приведенное определение не утверждено ни в каких нормативных документах, официального определения понятия СППР в настоящее время просто не существует. Таким образом, необходимо уточнить и узаконить определение СППР как «системы систем», объединяющей в себе автоматизированные средства сбора и анализа информации, расчетные задачи и модели, обеспечивающие формирование вариантов управляющих решений в процессе общения с пользователем на языке, близком к естественному.

Не меньше проблем в данной предметной области порождает и отсутствие единой классификации СППР. В настоящее время классификация базируется на переводных понятиях, используемых в ведущих зарубежных государствах.

По существующей классификации **на уровне пользователя** СППР делят на пассивные, активные и кооперативные [2]. Пассивной СППР (passive DSS) называют систему, которая обеспечивает процесс принятия решения информационно, но не формирует предложение, какое конкретно решение принять. Такие СППР иногда называют системами информационной поддержки принятия решений.

Благодаря простоте реализации, в настоящее время это наиболее распространенный тип СППР. Активная СППР (active DSS) может обосновать, какое из множества сформированных решений следует выбрать по тем или иным критериям. Кооперативная (диалоговая или итерационная) СППР позволяет ЛПР изменять, пополнять или улучшать решения, предлагаемые системой, путем их корректировки.

**На концептуальном уровне** СППР различают по признаку управляющих воздействий: управляемые сообщениями (Communication-Driven DSS), управляемые данными (Data-Driven DSS), управляемые документами (Document-Driven DSS), управляемые знаниями (Knowledge-Driven DSS), управляемые моделями (Model-Driven DSS) [3]. Впрочем, в связи со схожестью современных форматов представления графических, тестовых, аудио и других электронных данных деление по этому признаку является достаточно условным.

**На техническом уровне** различают СППР предприятия и персональную СППР [4]. СППР предприятия (Enterprise Resource Planning System) использует большие хранилища информации и обслуживает группу пользователей. В настоящее время имеются практические примеры программной реализации таких ERP-систем. К ним можно отнести коммерческие продукты разработки компаний SAP (Systeme, Anwendungen und Produkte in der Datenverarbeitung) и Oracle. Автор лично, сотрудничая с одним из ведущих системных интеграторов России, имел опыт эксплуатации ERP-системы SAP R/3, которая представляет собой полноценную СППР информационно-расчетного типа. Персональная СППР (Personal Information System) – это малая система, обслуживающая одного пользователя. Хотя, с формальной точки зрения, любая СППР является персональной, так как, какое бы количество персонала ни взаимодействовало с системой, конечное решение принимает единственный человек – ЛПР, на которого в итоге и работает вся система.

**В зависимости от актуальности данных**, с которыми эти системы работают, СППР условно делят на оперативные и стратегические [1]. Оперативные СППР (Executive Information Systems) предназначены для немедленного реагирования на изменения текущей ситуации в управлении объектами. Стратегические СППР (Customer Synchronized Resource Planning) ориентированы на анализ значительных объемов разнородной информации, собираемых из различных источников, и выработку перспективных управляющих решений. Впрочем, и такое деление достаточно условно, так как оценка оперативности используемых данных – в значительной степени субъективный показатель, определяемый, скорее, требованиями управляемой системы, чем самой СППР. Справедливости ради необходимо отметить, что управляющие системы – это вообще не СППР в классическом понимании:

точнее было бы называть их системами поддержки принятия и контроля выполнения решений. Но в рамках данной статьи не ставится задача предложить новый тип систем, поэтому данный аспект можно опустить.

Когда в состав СППР добавляется база знаний, систему иногда называют интеллектуальной [5]. В ряде случаев СППР, включающие в свой состав частные экспертные системы, относят к классу экспертных СППР [6–8].

Анализ показывает, что существующая классификация не в полной мере соответствует требованиям разработки и практического применения СППР. Типичным примером неточности классификации может служить использование перевода англоязычного термина «на уровне пользователя» (user level) для разделения систем по принципам взаимодействия с пользователем, хотя данный термин по сути подменяет понятие «по функциональным возможностям». Еще один пример: разделение СППР на обычные и интеллектуальные по наличию в них компонентов баз знаний, то есть по структурному признаку, на взгляд автора, не вполне корректно. В то же время является фактом, что интеллектуальные и экспертные СППР уже существуют и активно развиваются, и не учитывать их в общей классификации нельзя. Таким образом, вопрос уточнения классификации СППР является актуальным и своевременным.

### Формализация процесса принятия решения

В целях уточнения классификационных признаков СППР предлагается рассмотреть формализованное описание процесса принятия решения, предложенное в работах [9–11].

В формализованном виде деятельность ЛПР заключается в формировании вариантов управляющих воздействий и выборе наиболее приемлемых из них. Как правило, формируемые решения являются взаимоисключающими, то есть альтернативами. Для обеспечения управленческой деятельности ЛПР от СППР требуется сформировать некоторое множество альтернатив:  $R = \{A_j, S_j\}$ , где  $A$  – альтернатива;  $j$  – уровень предпочтения (ранг) альтернативы;  $S$  – описание альтернативы, пояснение, почему именно ей отдается предпочтение над предыдущими.

При этом функционирование описанной системы должно осуществляться по следующим основным этапам:

- формирование множества альтернатив  $R$ , которое заключается в последовательном формировании целевой функции и поиске некоторых ее рациональных параметров, определяющих стратегию применения средств и их ресурсов в рамках установленных ограничений;

- создание системы ограничений, формирующей границы области допустимых решений (ОДР),

которые могут быть заданы вышестоящим органом управления, обуславливаясь обстановкой или задаваться непосредственно оператором СППР;

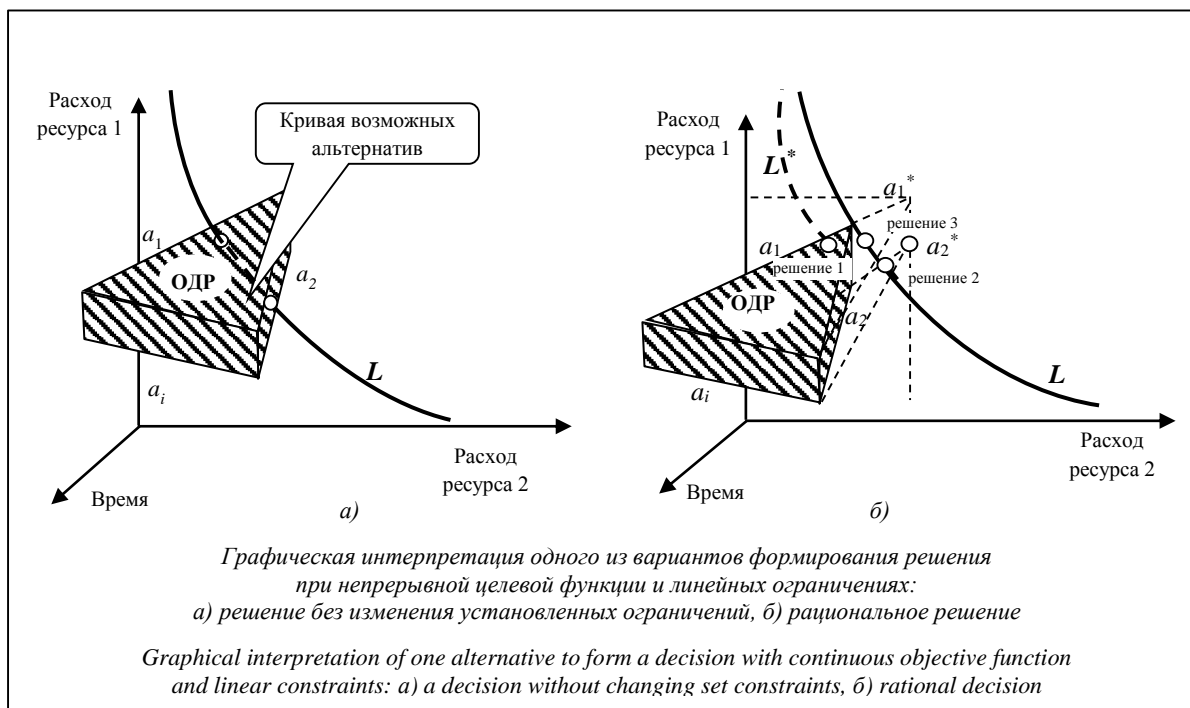
- вербальное или графическое описание альтернатив, которое не только упрощает принятие окончательного решения, но и повышает доверие пользователя к результатам решения задач и моделирования.

В частном случае, когда задача имеет однозначное решение в рамках поставленных условий и сформулированных ограничений, поддержку ЛПР может успешно осуществлять простая информационно-расчетная система активного или кооперативного типа по существующей классификации. Если она сформирует несколько вариантов решений (альтернатив), у ЛПР останется задача сравнения альтернатив и выбора из них лучшей относительно заданных параметров. Такой вариант решения без изменения установленных ограничений можно назвать решением по малому циклу (см. рисунок).

В то же время в условиях ограниченного количества сил и средств, наличия временных ограничений часто возникает ситуация, когда поставленная задача не имеет решения в рамках заданных условий, то есть множество альтернатив  $R$  оказывается пустым. Тогда, чтобы найти рациональное решение, необходимо либо изменять некоторые характеристики целевой функции, либо переводить одно или несколько ограничений в ранг управляемых параметров и уже с их помощью искать вариант, обеспечивающий нахождение рационального решения (см. рисунок). Такую задачу можно называть поиском решения по большому циклу. Задача в такой постановке является многопараметрической, и ее решение без применения интеллектуальной СППР невозможно.

В формализованном виде постановка задачи поиска решения СППР по большому циклу представляет собой поиск новых значений ограничений  $a_i^*$ , при которых задача поиска оптимума  $L$  имеет решение, или новых параметров функции  $L^*$ , обеспечивающих нахождение решения при существующих ограничениях  $a_i$ . В рамках решения сформулированной задачи СППР позволит сформировать группу решений, изучив которые, ЛПР сможет сделать выводы о том, что необходимо выполнить для обеспечения достижения поставленной цели.

Справедливости ради следует отметить, что на рисунке представлена несколько упрощенная интерпретация постановки принятия решения. На практике ограничения, вероятнее всего, будут нелинейными, а целевая функция представлена не кривой, а «трубкой поведенческих траекторий», формирующейся в условиях неопределенности. Но в любом случае постановка задачи будет аналогична вышеописанной, а сама задача сведется к поиску параметров целевой функции или ограничений, обеспечивающих получение рационального решения.



Решение задач по малому циклу является достаточно тривиальным – систем, реализующих такой цикл, создано довольно много. Примером практической реализации автоматизированной СППР малого цикла может служить обычный автомобильный GPS-навигатор. Это программно-техническое средство имеет все признаки СППР [1]: общается с пользователем на языке, близком к естественному, решает оптимизационную задачу формирования альтернативных управляющих воздействий по заданным пользователем условиям, автоматически собирает исходную информацию для собственного функционирования. А задать условия, решение для которых потребуется по большому циклу, пользователь не сможет физически. Ведь выбор пунктов для поиска пути между ними осуществляется по списку, формируемому из матрицы связности дорожной сети, только для тех точек, между которыми есть дороги.

Намного сложнее создать СППР большого цикла – интеллектуальную систему, позволяющую искать управленческие альтернативы там, где они не существуют в заданных условиях. Такая система должна не просто обрабатывать информацию по заданным алгоритмам, но и оперировать знаниями и формировать алгоритмы обработки данных. Чтобы показать отличие предлагаемого подхода к описанию интеллектуальных систем от обычных СППР, достаточно представить тот же автомобильный навигатор, который может прокладывать путь там, где нет дороги, и, проанализировав обстановку, не выдавать сообщение о том, что проехать нельзя, а формировать предложения: проехать через поле, проходимое в данный сезон, идти вброд через реку или использовать другой вид транспорта. Вот это и будет разницей в возможностях

СППР малого и большого циклов. Именно такие системы предлагается называть интеллектуальными СППР.

### Предложения по уточнению классификации СППР

Анализ формализованного представления принятия решения позволяет выделить основные значимые признаки систем, обеспечивающих поддержку этого процесса. Анализ данных признаков позволяет не только сформулировать определение СППР, но и синтезировать предложения по уточнению их классификации.

Во-первых, предлагается устранить избыточность системы классификации, уточнив классификационные признаки или отменив деление по признакам «на техническом уровне», «в зависимости от актуальности данных» и «на концептуальном уровне».

Во-вторых, заменить определение признака «относительно принципа взаимодействия с пользователем» на деление «по функционалу», разделив последний на три подкласса:

- информационные СППР, обеспечивающие сбор и агрегирование больших объемов информации, необходимой для принятия решений;
- информационно-расчетные системы, обеспечивающие формирование набора альтернатив управляющих воздействий и их ранжирование;
- интеллектуальные СППР, обеспечивающие поиск управляющих воздействий при решении нетривиальных задач.

В заключение сделаем следующие выводы. Проблема упорядочения предметной области давно назрела, перешла из области теоретических

исследований в практическую и требует скорейшего решения. Как показал анализ состояния проблемы, для приведения в порядок предметной области поддержки принятия решений необходимо выполнить следующие мероприятия:

- уточнить и официально утвердить определение системы поддержки принятия решений, однозначно отражающее ее основные свойства;
- уточнить классификацию систем поддержки принятия решений.

Предлагаемое в статье уточнение классификации СППР позволит усовершенствовать понятийный аппарат в области поддержки принятия решений и в итоге обеспечить разработку и внедрение таких систем в деятельность органов управления.

### Литература

1. Ларичев О.И., Петровский А.В. Системы поддержки принятия решений. Современное состояние и перспективы их развития // Итоги науки и техники: Сер. Технич. кибернетика. М.: Изд-во ВИНТИ, 1987. Т. 21. С. 131–164.
2. Haettenschwiler P. Neues anwenderfreundliches Konzept der Entscheidungs-unterstützung. Gutes Entscheiden in Wirtschaft, Politik und Gesellschaft. Zurich: Hochschulverlag AG, 1999, pp. 189–208.

3. Power D.J. What is a DSS? The On-Line Executive Journ. for Data-Intensive Decision Support, 1997, vol. 1, no. 3.

4. Power D.J. A brief history of decision support systems. DSSResources.COM, 2003. URL: <http://DSSResources.com/history/history.html> (дата обращения: 13.12.2015).

5. Turban E. Decision support and expert systems: management support systems. Englewood Cliffs, NJ, Prentice Hall Publ., 1995, 885 p.

6. Кравченко Т.К., Середенко Н.Н. Создание систем поддержки принятия решений: интеграция преимуществ отдельных подходов // Искусственный интеллект и принятие решений. 2012. № 1. С. 39–46.

7. Рыбина Г.В., Блохин Ю.М., Иващенко М.Г. Интеллектуальная технология построения интегрированных экспертных систем // Искусственный интеллект и принятие решений. 2011. № 3. С. 48–67.

8. Рыбина Г.В., Иващенко М.Г. Методы и программные средства интеллектуальной поддержки разработки интегрированных экспертных систем // Программные продукты и системы. 2006. № 6. С. 27–38.

9. Тиханьчев О.В. Автоматизация поддержки принятия решений. М.: Эдитус, 2015. 96 с.

10. Тиханьчев О.В. Системы поддержки принятия решений – перспективное направление развития автоматизации управления войсками (силами) // Военная мысль. 2012. № 8. С. 45–51.

11. Tikhanychev O.V. Decision-making support systems: prospects for troops control automation. Military Thought, 2012, vol. 21, iss. 3, pp. 74–83.

DOI: 10.15827/0236-235X.115.024-028

Received 10.02.16

### ON SOME PROBLEMS IN THE DECISION SUPPORT DOMAIN

<sup>1</sup>Tikhanychev O.V., Ph.D. (Engineering), Senior Researcher, tow65@yandex.ru

<sup>1</sup> 27 Central Research Institute of the Ministry of Defense of Russia,

1<sup>st</sup> Khoroshevsky proezd 5, Moscow, 123007, Russian Federation

**Abstract.** The article analyzes the current state of a decision support domain: both in terms of classification principles of such systems and their official definition. This situation complicates the process of creation and application of decision support systems. To normalize this domain the author formalizes a description of a decision-making process and formulates the concepts of “small” and “extended” decision-making cycles. “Small” cycle refers to the process of finding an optimal solution in terms of a set objective function and the restrictions. “Big” cycle includes searching for a non-trivial solution, which cannot be obtained without modifying the objective function or restriction cancellation. Based on this approach the author proposes to clarify the definition of an automated decision-making system and its classification. The main indicator in clarifying the classification is functional defined by the type of a decision-making cycle. Within a new classification it is proposed to replace the currently used class “relative to the user level” by the class “by the functional”, which includes information, calculation and information and intelligent systems. The proposed classification completely corresponds to the modern understanding of automated decision-making support system and can accelerate the process of implementing such systems in management practice.

**Keywords:** decision-making support systems, expert system, classification, intelligent assistants, intellectual system, decision-making cycle.

### References

1. Larichev O.I., Petrovsky A.V. Decision-making support systems. Current state and prospects. *Itoги nauki i tekhniki. Tekhnicheskaya kibernetika* [Science and Engineering Results. Series Technical Cybernetics]. Moscow, VINITI Publ., 1987, vol. 21 (in Russ.).
2. Haettenschwiler P. *Neues anwenderfreundliches Konzept der Entscheidungs-unterstützung. Gutes Entscheiden in Wirtschaft, Politik und Gesellschaft* [New User-friendly Concept of Decision Support. Good Decisions in Business, Politics and Society]. Zurich, Hochschulverlag AG, 1999, pp. 189–208 (in Ger.).
3. Power D.J. What is a DSS? *The On-Line Executive Journal for Data-Intensive Decision Support*. 1997, vol. 1, no. 3.
4. Power D.J. A Brief History of Decision Support Systems. *DSSResources.COM*. 2003. Available at: <http://DSSResources.com/history/history.html> (accessed December 13, 2015).
5. Turban E. *Decision support and expert systems: management support systems*. Englewood Cliffs, N.J., Prentice Hall Publ., 1995, 885 p.
6. Kravchenko T.K., Seredenko N.N. Decision support systems development: integration of the some different approaches. *Iskusstvenny intellekt i prinyatie resheny* [Artificial Intelligence and Decision Making]. 2012, no. 1, pp. 39–46 (in Russ.).
7. Rybina G.V., Ivashchenko M.G., Blokhin Yu.M. Intelligent integrated expert systems technology. *Iskusstvenny intellekt i prinyatie resheny* [Artificial Intelligence and Decision Making]. 2011, no. 3, pp. 48–75 (in Russ.).
8. Rybina G.V., Ivashchenko M.G. Methods and software means for intelligent support of integrated expert systems development. *Programmnye produkty i sistemy* [Software and Systems]. 2006, no. 6, pp. 27–38 (in Russ.).
9. Tikhanychev O.V. *Avtomatizatsiya podderzhki prinyatiya resheny* [Decision-making Support Automation]. Moscow, Editus Publ., 2015, 96 p.
10. Tikhanychev O.V. Decision-Making Support Systems: Prospects for Troops Control Automation. *Voennaya mysl* [Military Thought]. 2012, no. 8, pp. 45–51 (in Russ.).
11. Tikhanychev O.V. Decision-Making Support Systems: Prospects for Troops Control Automation. *Voennaya mysl* [Military Thought]. 2012, vol. 21, iss. 3, pp. 74–83.

УДК 519.68

Дата подачи статьи: 01.07.16

DOI: 10.15827/0236-235X.115.029-035

## **ВЫЧИСЛИТЕЛЬНЫЕ ПРОЦЕДУРЫ ОБРАБОТКИ НЕТОЧНОЙ ВРЕМЕННОЙ ИНФОРМАЦИИ**

*В.В. Фомин, д.т.н., профессор, v\_v\_fomin@mail.ru;*

*С.В. Мальцев, аспирант, maltsevsergey@live.com*

*(Российский государственный педагогический университет им. А.И. Герцена,  
наб. реки Мойки, 48, г. Санкт-Петербург, 191186, Россия)*

В статье представлены результаты обширного исследования, обусловленного практической потребностью повышения эффективности методов представления, алгоритмов извлечения, обработки нечетких темпоральных данных. Акцент делается на онтологических моделях процессов прогнозирования и интерпретации при взаимодействии с неточной временной информацией. Предлагаются методология хранения данных, а также модели и унифицированные процедуры извлечения нечеткой информации. Непосредственно затрагивается вопрос о прагматике алгоритмов представления темпоральных данных через новый инструмент так называемой возможностной темпоральной информации в аналитических системах, где параметры неточности и оперативности являются критическими условиями в принятии бизнес-решений. Предлагаются решения проблем хранения и извлечения неточной временной информации за счет построения модели обработки нечетко-темпоральных отношений между временными точками и/или интервалами. Представлена концепция анализа и представления нечеткой темпоральности в рамках развития модели темпоральных реляционных БД с учетом критериев неточности, неоднозначности, неопределенности. Разработана алгебра возможных отношений точек и/или интервалов, в ее рамках определены понятия возможностной временной точки и возможностного интервала, введены соответствующие значения возможностных временных точек и интервалов. Представлен фрагмент формализованного алгоритма реализации алгебры на примере расширения одного из отношений. Предлагается подход к разработке и моделированию нечеткого темпорального домена и отношений между возможностными интервалами и моментами времени. Дано описание реализованной библиотеки алгоритмов обработки отношений возможностных временных точек и/или интервалов, а также соответствующих процедур и функций на платформе стандарта SQL. Представлены результаты внедрения библиотеки алгоритмов при реализации коммерческого проекта аналитической системы для телекоммуникационной компании, а также определены направления дальнейшего развития прикладного инструментария, основанного на теории нечетко-темпоральной модели.

**Ключевые слова:** темпоральные БД, временная неопределенность, программы и алгоритмы.

В общем смысле под *темпоральностью* понимается связь фактов, событий или явлений и их временных показателей. Факты в явном виде могут не иметь никакой принадлежности к временным показателям, но в неявном виде всегда существует и может быть определена их зависимость от времени. Можно сказать, что любая деятельность реального мира тем или иным образом связана со временем. Темпоральность подразумевает временную сущность фактов, определение времени их совершения или динамики их движения и изменения по временной шкале.

Время является неотъемлемой частью многих областей деятельности человека. Для решения задач, связанных со временем, эффективным аналитическим показателем [1] является коэффициент неточности – интервальная метрика, позволяющая обобщать и гранулировать периоды, использующая неопределенность в интерпретации времени, достаточная для соблюдения семантических особенностей интерпретации фактов.

Таким образом, неточность во времени образует новое направление для исследований – так называемую *нечеткую темпоральность*, которая подразумевает наличие неопределенности, неоднозначности или отсутствия информации о времени как таковом. Если говорить, что темпоральная БД – это организованная структура, построенная на основе темпоральной модели данных и содержа-

щая темпоральные данные, то *нечетко-темпоральная БД* – это структура, построенная на совокупности алгоритмов и унифицированных процедур хранения, извлечения и обработки нечетких темпоральных данных.

Теоретические достижения и значимость исследований в области обработки нечетко-темпоральных данных отражены в научно-технических публикациях [2–4]. Прикладные системы обработки нечетко-темпоральных данных востребованы в различных прикладных областях [5–8] и являются одной из индустриальных задач информационных технологий.

Проведение исследования вызвано отсутствием необходимых средств в системах различного назначения для использования и обработки нечетко-темпоральной информации с целью ее дальнейшей обработки и интерпретации. Приведем примеры таких систем.

1. Системы планирования и прогнозирования. Фиксирование нечеткой временной информации в календаре или в планировщике задач, в частности, фиксация/планирование/смещение задач с учетом погрешности во времени.

2. Фиксирование исторических неточных фактов или данных.

3. Последовательная цепочка выполнения задач, где каждый последующий процесс зависит от факта и времени выполнения предыдущего.



Потребность в использовании темпоральной модели хранения данных актуализировалась с появлением темпоральной поддержки [9] в стандарте SQL:2011. С появлением стандарта разработчики аналитических систем стали активно добавлять в свои продукты темпоральную поддержку, реализация которой способствовала активному внедрению на предприятиях для решения задач, связанных с учетом времени. Темпоральная поддержка в стандарте SQL:2011 ограничивается в решении вопросов обработки нечетких темпоральных данных. Необходимость использования нечеткой логики в обработке темпоральной информации вызвана возможным наличием неточности в исходных данных, где периоды начала/завершения действия интервалов или событий могут быть неточно определены или не определены вообще. Расширенными возможностями, прорабатываемыми в новой модели, являются следующие:

- рассмотрение и проработка не только исторических темпоральных фактов, но и фактов, которые могут наступить в предположительное время в будущем, от времени выполнения которых могут зависеть другие задачи;
- изучение не только неточных интервалов, но и временных точек, принимающих множество значений;
- рассмотрение погрешности во временных точках, а также в диапазонах значений интервалов;
- изучение и проработка нетемпоральных периодов (например integer и string);
- изучение применения неточности в нескольких временных измерениях;
- рассмотрение методов конвертации исторических данных, хранящихся в действующих реляционных структурах, в возможно-темпоральные.

#### Построение модели обработки неточной временной информации

В темпоральной модели данных выделяются следующие типы нечеткости [4].

**Неопределенность.** Какая-либо часть темпоральной информации неизвестна или определена неточно. Например, дата рождения Сократа – «около 469 год д.н.э.».

**Субъективность, двусмысленность.** Темпоральные события или периоды могут быть субъективно или неоднозначно определены. Например, «эпоха нового времени» или в сфере грузоперевозки «даты начала или завершения транспортировки груза».

**Неясность, расплывчатость.** События могут быть определены с различной гранулярностью или нечеткостью. Например, «доставка груза в течение дня».

С целью реализации модели, рассматривающей и обрабатывающей темпоральные данные, в кото-

рых присутствуют неопределенность, субъективность или неясность, необходимо ввести понятие возможностной временной точки.

**Возможностная временная точка ( $pt$ )** – предполагаемая временная отметка (или момент времени) на временной оси, значение которой неизвестно либо неточно определено.

Введем следующие значения возможностных временных точек.

1. **Значение определено с возможным отклонением.** Предполагаемое значение временной точки принимает степень возможности 1, значения возможностного отклонения от времени наступления находятся в диапазоне  $[0, 1]$ . Вид:  $[pt, pt - \alpha, pt + \beta]$ , где  $pt \in T$  – возможностная временная точка, предполагаемый момент времени на временном отрезке  $T$ ;  $\alpha$  и  $\beta$  – значения возможностного отклонения от наступления (рис. 1).

Значения отклонения должны быть заданы с соответствующей гранулярностью  $\alpha$  и  $\beta$ , то есть в единицах времени из существующих форматов: ss, mi, hh24, dd, mm, уууу.

2. **Значение в диапазоне не определено.** Значения в диапазоне имеют одинаковую степень возможности наступления  $PExtent$ .

Вид:  $[\{pt_x, pt_y\}, PExtent]$ , где  $\{pt_x, pt_y\} \in T$  – диапазон значений возможностных временных точек предполагаемого момента времени на временном отрезке  $T$ ;  $pt_x$  и  $pt_y$  могут принимать пустое значение. В таком случае (рис. 2) предполагается диапазон значений в виде:  $[\{pt_x, \infty\}, PExtent] \vee [\{\infty, pt_y\}, PExtent]$ .

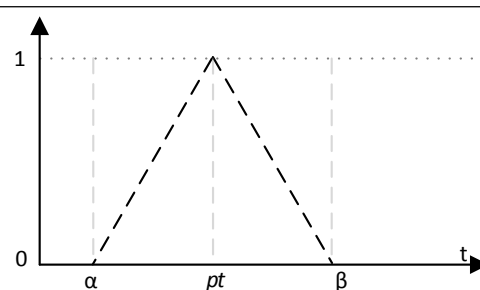


Рис. 1. Значение определено с возможным отклонением

Fig. 1. The value is determined with possible deviation

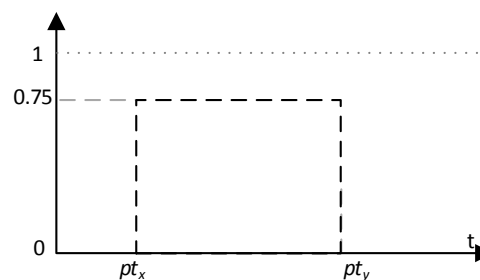


Рис. 2. Значение в диапазоне не определено

Fig. 2. A value in a range is not determined

3. Значение не определено. Любые значения имеют одинаковую степень возможности наступления = 0. Вид: [undefined, 0].

### Возможностные отношения временных точек и интервалов

Одной из наиболее распространенных моделей учета временной неопределенности является темпоральная алгебра Аллена. Дж.Ф. Аллен определил набор из 13 качественных отношений, которые могут существовать между двумя временными интервалами [10]. Алгебра Аллена ограничивается использованием отношений между периодами (отрезками, интервалами) времени и характеризуется мерой длительности временных отрезков.

Однако практика временных отношений базируется не только на временных интервалах, но и на понятиях временных точек. Временная точка и временной интервал являются неотъемлемой составляющей реализации темпорального анализа данных.

Для построения модели, рассматривающей возможностные отношения между точками и/или интервалами, темпоральная алгебра Аллена была расширена путем декомпозиции интервальных отношений до отношений между двумя точками, а также точками и интервалами. Введенные отношения отражены в таблице 1.

На основании темпоральных отношений между точками и/или интервалами, а также введенных значений возможностных временных точек была разработана алгебра возможностных отношений точек и/или интервалов.

Рассмотрим фрагмент алгебры на примере расширения одного отношения «Перед» (Before) (табл. 1) до трех типов отношений (табл. 2).

Для учета эффекта неопределенности применяется функция принадлежности теории нечетких множеств, которая позволяет определить степень принадлежности элементов фундаментального множества пространства рассуждения нечеткому множеству. За основу меры степени принадлежности взяты введенные Л. Заде [1] вещественные числа в интервале от 0 до 1: 0 означает, что элемент не включен в нечеткое множество, 1 – описывает полностью включенный элемент. Значения между 0 и 1 характеризуют нечетко включенные элементы.

Данное отношение предназначено для определения признака наступления возможностной временной точки  $a$  до  $b$  и принимает следующий вид:  $PosBefore(a, b[, PExtent\_IN])$ , где  $a$  и  $b$  – обязательные сравниваемые значения временных точек;  $PExtent\_IN \in [0, 1]$  (опционально) – степень, требуемая для выполнения условия.

Отношение возвращает  $(PExtent\_OUT[, boolean])$ , где  $PExtent\_OUT \in [0, 1]$  – результирующая

Таблица 1

Введенные отношения временных точек и интервалов

Table 1

Input dependencies between time points and intervals

Отношение	Точки A, B и интервал J		Визуализация
Перед (Before)	A перед J A перед B	$A_t < S_j$ $A_t < B_t$	
После (After)	A после J A после B	$A_t > E_j$ $A_t > B_t$	
Равен (Equal)	A равен B	$A_t = B_t$	
Содержит (Contains)	J содержит A	$S_j < A_t < E_j$	
Начинает (Starts)	A начинает J	$S_j = A_t < E_j$	
Завершает (Finishes)	A завершает J	$S_j < A_t = E_j$	

Таблица 2

Отношения возможностных точек и интервалов «Before»

Table 2

Dependencies between probabilistic points and intervals “Before”

Отношение	Точки a, b, интервалы IPeriod, JPeriod	Вид
Точки	a перед b	$PosBefore(a, b[, PExtent\_IN])$
Интервалы	IPeriod перед JPeriod	$PosBefore(IPeriod, JPeriod[, PExtent\_IN])$
Точки-Интервалы	a перед IPeriod	$PosBefore(a, IPeriod.startdate)$

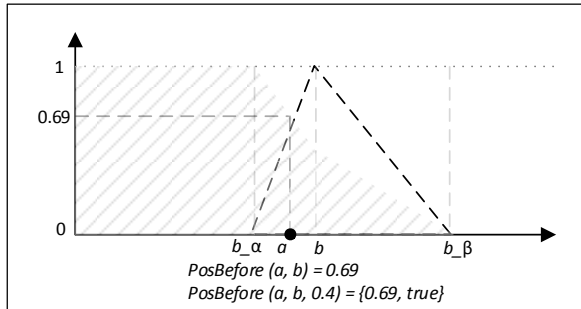


Рис. 3. Отношение «Перед»: значение точки  $b$  определено с возможным отклонением»

Fig. 3. Relation "Before:  $b$  point value is determined with possible deviation"

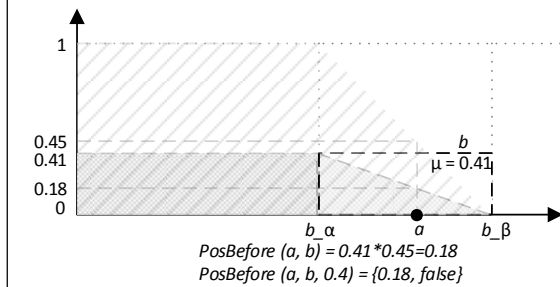


Рис. 4. Отношение «Перед»: значение точки  $b$  в диапазоне не определено»

Fig. 4. Relation "Before:  $b$  point value in a range is not determined"

возможная степень выполнения; *boolean* – признак выполнения: *true*, если  $PExtent\_OUT \geq PExtent\_IN$ , или *false*, если  $PExtent\_OUT < PExtent\_IN$ .

Варианты использования отношения «Перед» для двух введенных значений возможных временных точек.

#### 1. Точное значение точки $a$ .

##### 1.1. Точное значение точки $b$ :

$$PosBefore = \begin{cases} 0, & b \leq a, \\ 1, & b > a. \end{cases}$$

##### 1.2. Значение точки $b$ определено с возможным отклонением:

$$PosBefore = \begin{cases} 0, & a \geq b_{\beta}, \\ 1 - \frac{a - b_{\alpha}}{b - b_{\alpha}} / 2, & b_{\alpha} \leq a < b, \\ \left(1 - \frac{a - b}{b_{\beta} - b}\right) / 2, & b \leq a < b_{\beta}, \\ 1, & a < b_{\alpha}. \end{cases}$$

Возможная точка  $b$  с периодами возможного отклонения наступления  $b_{\alpha}$  и  $b_{\beta}$  (рис. 3).

##### 1.3. Значение точки $b$ в диапазоне не определено.

Возможная точка  $b$  с возможным диапазоном времени наступления  $b_{\alpha}$  и  $b_{\beta}$  и возможной степенью  $b\_PExtent\_IN$  (рис. 4):

$$PosBefore = \begin{cases} 0 * b\_PExtent\_IN, & b_{\beta} \leq a, \\ \left(1 - \frac{a - b_{\alpha}}{b_{\beta} - b_{\alpha}}\right) * b\_PExtent\_IN, & b_{\alpha} \leq a < b_{\beta}, \\ 1 * b\_PExtent\_IN, & a < b_{\alpha}. \end{cases}$$

В случаях, когда  $b_{\alpha}$  и/или  $b_{\beta}$  принимают пустое значение, для проведения корректных расчетов подставляются даты начала или завершения значениями по умолчанию ( $d\_sd$  и  $d\_ed$ ), которые должны быть определены в соответствии с предметной областью.

При  $b_{\alpha} = \text{null}$  и  $b_{\beta} \neq \text{null}$ :

$$PosBefore = \begin{cases} 0, & b_{\beta} \leq a, \\ \left(1 - \frac{a - d\_sd}{b_{\beta} - d\_sd}\right) * b\_PExtent\_IN, & d\_sd \leq a < b_{\beta}, \\ 1 * b\_PExtent\_IN, & a < d\_sd. \end{cases}$$

При  $b_{\alpha} \neq \text{null}$  и  $b_{\beta} = \text{null}$ :

$$PosBefore = \begin{cases} 1 * b\_PExtent\_IN, & a \leq b_{\alpha}, \\ \left(1 - \frac{a - b_{\alpha}}{d\_ed - b_{\alpha}}\right) * b\_PExtent\_IN, & b_{\alpha} \leq a < d\_ed, \\ 0 * b\_PExtent\_IN, & d\_ed < a. \end{cases}$$

При  $b_{\alpha} = \text{null}$  и  $b_{\beta} = \text{null}$ :

$$PosBefore = \begin{cases} 0, & d\_ed \leq a, \\ \left(1 - \frac{a - d\_sd}{d\_ed - d\_sd}\right) * b\_PExtent\_IN, & d\_sd \leq a < d\_ed, \\ 1 * b\_PExtent\_IN, & a < d\_sd. \end{cases}$$

#### 1.4. Значение точки $b$ не определено.

Возможная точка  $b$  с неизвестным временем наступления.

Отношение возвращает исходное значение  $PExtent\_IN$ , которое принимает значение 0.

#### 2. Точка $a$ – значение определено с возможным отклонением.

##### 2.1. Точное значение точки $b$ :

$$PosBefore = \begin{cases} 0, & b \leq a_{\alpha}, \\ 1 - \frac{a_{\beta} - b}{a_{\beta} - a_{\alpha}} / 2, & a_{\alpha} < b \leq a_{\beta}, \\ \left(1 - \frac{a - b}{a - a_{\alpha}}\right) / 2, & a_{\alpha} < b \leq a, \\ 1, & b > a_{\beta}. \end{cases}$$

Возможная точка  $a$  с периодами возможного отклонения наступления  $a_{\alpha}$  и  $a_{\beta}$  (рис. 5).

##### 2.2. Значение точки $b$ определено с возможным отклонением.

Возможная точка  $a$  с периодами возможного отклонения наступления  $a_{\alpha}$  и  $a_{\beta}$ . Возможная точка  $b$  с периодами возможного отклонения наступления  $b_{\alpha}$  и  $b_{\beta}$  (рис. 6).

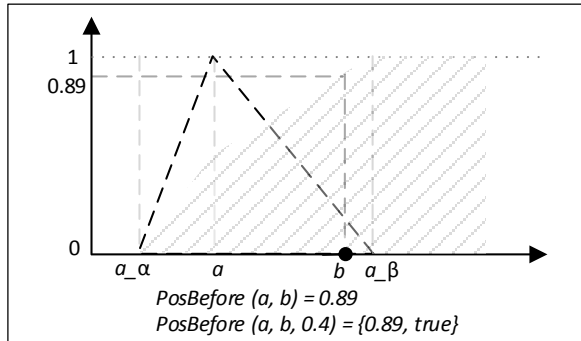


Рис. 5. Отношение «Перед»: значение точки  $b$  не определено

Fig. 5. Relation "Before:  $b$  point value is not determined"

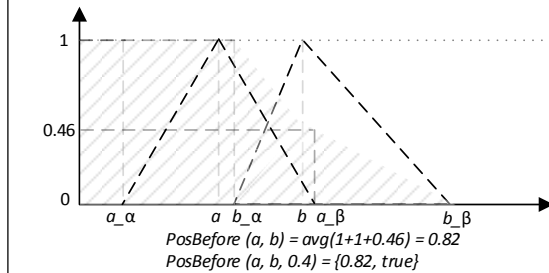


Рис. 6. Отношение «Перед»: значение точки  $b$  определено с возможным отклонением

Fig. 6. Relation "Before:  $b$  point value is determined with possible deviation"

Для каждого  $a_x \in A$ , где  $A = \{a, a_\alpha, a_\beta\}$ , выполняется следующий расчет степени  $pos\_extent$ :

$$pos\_extent = \begin{cases} 0, & a_x > b_\beta, \\ 1 - \frac{a_x - b_\alpha}{b - b_\alpha} / 2, & b_\alpha < a_x \leq b, \\ \left(1 - \frac{a_x - b}{b_\beta - b}\right) / 2, & b < a_x \leq b_\beta, \\ 1, & a_x \leq b_\alpha. \end{cases}$$

Степень возможности наступления  $a$  до  $b$  вычисляется при помощи среднего значения полученных  $pos\_extent$ :  $PosBefore = AVG(pos\_extent(a), pos\_extent(a_\alpha), pos\_extent(a_\beta))$ .

2.3. Значение точки  $b$  в диапазоне не определено.

Возможностная точка  $a$  с периодами возможного отклонения наступления  $a_\alpha$  и  $a_\beta$ . Возможностная точка  $b$  с возможностным диапазоном времени наступления  $b_\alpha$  и  $b_\beta$  и возможностной степенью  $b\_PEntent\_IN$  (рис. 7).

Для каждого  $a_x \in A$ , где  $A = \{a, a_\alpha, a_\beta\}$ , выполняется следующий расчет степени  $pos\_extent$ :

$$pos\_extent = \begin{cases} 0, & a_x > b_\beta, \\ \left(1 - \frac{a_x - b_\alpha}{b_\beta - b_\alpha}\right) * b\_PEntent\_IN, & b_\alpha < a_x \leq b_\beta, \\ 1 * b\_PEntent\_IN, & a_x \leq b_\alpha. \end{cases}$$

Степень возможности наступления  $a$  до  $b$  вычисляется при помощи среднего значения полученных  $a\_extent$ :  $PosBefore = AVG(pos\_extent(a), pos\_extent(a_\alpha), pos\_extent(a_\beta))$ .

2.4. Значение точки  $b$  не определено.

Возможностная точка  $b$  с неизвестным временем наступления.

Отношение возвращает исходное значение  $PEntent\_IN$ , которое принимает значение 0.

3. Значение точки  $a$  в диапазоне не определено.

3.1. Точное значение точки  $b$ .

Возможностная точка  $a$  с возможностным диапазоном времени наступления  $a_\alpha$  и  $a_\beta$  и возможностной степенью  $a\_PEntent\_IN$  (рис. 8):

$$PosBefore = \begin{cases} 0, & a_\alpha > b, \\ \left(\frac{b - a_\alpha}{a_\beta - a_\alpha}\right) * a\_PEntent\_IN, & a_\alpha \leq b < a_\beta, \\ 1 * a\_PEntent\_IN, & a_\beta \leq b. \end{cases}$$

3.2. Значение точки  $b$  определено с возможным отклонением.

Возможностная точка  $a$  с возможностным диапазоном времени наступления  $a_\alpha$  и  $a_\beta$  и возможностной степенью  $a\_PEntent\_IN$  (рис. 9).

Возможностная точка  $b$  с периодами возможного отклонения наступления  $b_\alpha$  и  $b_\beta$ .

Для каждого  $a_x \in A$ , где  $A = \{a_\alpha, a_\beta\}$ , выполняется следующий расчет степени  $pos\_extent$ :

$$pos\_extent = \begin{cases} 0, & a_x \geq b_\beta, \\ 1 - \frac{a_x - b_\alpha}{b - b_\alpha} / 2 * a\_PEntent\_IN, & b_\alpha \leq a_x < b, \\ \left(1 - \frac{a_x - b}{b_\beta - b}\right) / 2 * a\_PEntent\_IN, & b \leq a_x < b_\beta, \\ 1 * a\_PEntent\_IN, & a_x < b_\alpha. \end{cases}$$

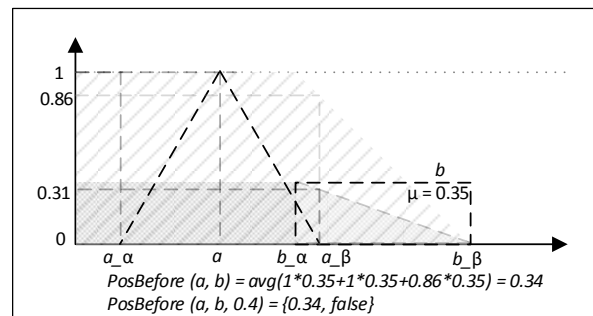
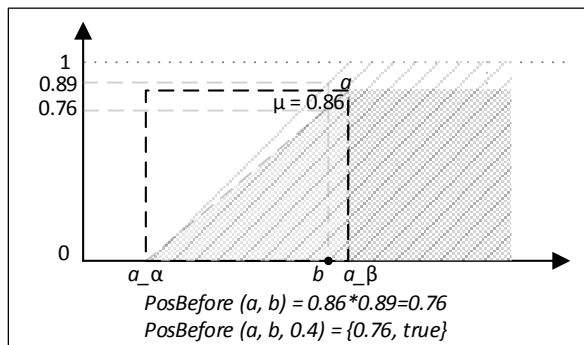
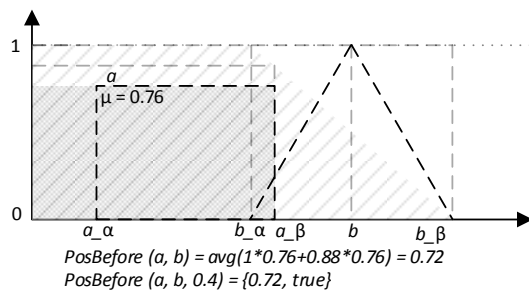
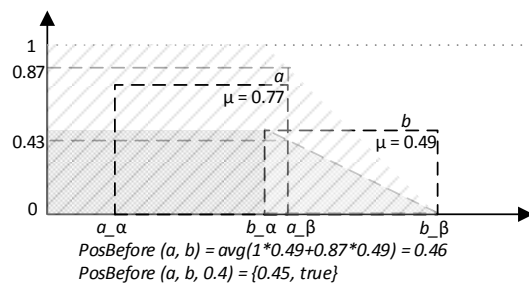


Рис. 7. Отношение «Перед»: значение точки  $b$  в диапазоне не определено

Fig. 7. Relation "Before:  $b$  point value in a range is not determined"

Рис. 8. Отношение «Перед: точное значение точки  $b$ »Fig. 8. Relation "Before:  $b$  point precise value"Рис. 9. Отношение «Перед: значение точки  $b$  определено с возможным отклонением»Fig. 9. Relation "Before:  $b$  point value is determined with possible deviation"Рис. 10. Отношение «Перед: значение точки  $b$  в диапазоне не определено»Fig. 10. Relation "Before:  $b$  point value in a range is not determined"

Степень возможности наступления  $a$  до  $b$  вычисляется при помощи среднего значения полученных  $pos\_extent$ :

$$PosBefore = \text{AVG}(pos\_extent(a_\alpha), pos\_extent(a_\beta)).$$

3.3. Значение точки  $b$  в диапазоне не определено.

Возможностная точка  $a$  с возможностным диапазоном времени наступления  $a_\alpha$  и  $a_\beta$  и возможностной степенью  $a\_PExtent\_IN$ . Возможностная точка  $b$  с возможностным диапазоном времени наступления  $b_\alpha$  и  $b_\beta$  и возможностной степенью  $b\_PExtent\_IN$  (рис. 10).

Для  $a_x \in A$ , где  $A = \{a_\alpha, a_\beta\}$ , выполняется следующий расчет степени  $pos\_extent$ :

$$pos\_extent = \begin{cases} 1 * \min(a\_PExtent\_IN, b\_PExtent\_IN), & a_x < b_\alpha, \\ \left(1 - \frac{a_x - b_\alpha}{b_\beta - b_\alpha}\right) \times \\ \times \min(a\_PExtent\_IN, b\_PExtent\_IN), & b_\alpha \leq a_x < b_\beta, \\ 0, & b_\beta \leq a_x. \end{cases}$$

Степень возможности наступления  $a$  до  $b$  вычисляется при помощи среднего значения полученных  $pos\_extent$ :

$$PosBefore = \text{AVG}(pos\_extent(a_\alpha), pos\_extent(a_\beta)).$$

3.4. Значение точки  $b$  не определено.

Возможностная точка  $b$  с неизвестным временем наступления.

Отношение возвращает исходное значение  $PExtent\_IN$ , которое принимает значение 0.

По аналогии с представленным отношением «Перед» формализованы все заданные отношения интервалов Аллена [10] и расширенные отношения в таблице 1.

### Библиотека алгоритмов

Алгебра отношений возможностей временных точек и/или интервалов была реализована в виде библиотеки алгоритмов на основе средств языка pl/sql. Разработанная библиотека алгоритмов дополняет и расширяет возможности языка SQL для обработки информации на базе реляционных моделей данных в развитие теории темпоральных моделей данных. В перспективе разработка отдельного инструментария систем управления БД для работы с темпоральными БД.

Апробация библиотеки алгоритмов возможностей временных точек и/или интервалов проводилась в телекоммуникационной компании. Библиотека была интегрирована в модуль анализа работы биллинговых систем оператора связи и их взаимодействия со смежными системами. Модуль позволяет выявлять процессы, в которых присутствуют неопределенность и неточность в информации, связанной с временными показателями.

Разработанная библиотека алгоритмов возможностей временных точек и/или интервалов была применена в системе мониторинга работы биллинговой системы при решении задачи управления программами лояльности, в том числе:

- прогнозирование и фиксирование предполагаемых дат оплаты счетов абонентами на основании показателей прошлых периодов;
- анализ временных периодов внесения платежей абонентами и их сопоставление с возможностными датами оплаты непогашенных счетов;
- прогнозирование предполагаемой даты завершения биллингового расчетного периода – сбора счетов и платежей абонентов за предыдущий месяц;

– расчет и хранение возможных дат предоставления премиальных пакетов абонентам за пользование услугами связи.

На основании изложенного сделаем следующие выводы. Построение и реализация модели обработки темпоральных данных, в которых присутствует неточность или неопределенность, а также разработка процедур на основе подхода возможных отношений позволяют решить ряд практических задач анализа и прогнозирования временных данных.

В дальнейшем исследовании планируется продолжить развитие нечетко-темпоральной модели:

– расширить класс возможных отношений между возможными временными точками и интервалами;

– интегрировать в модель возможные не-темпоральные периоды и применение нескольких возможных временных осей в одном отношении;

– реализовать процедурную модель для языков определения и манипулирования данными.

В целом исследования направлены на изучение и построение моделей извлечения информации в задачах с темпоральной неопределенностью, с опорой на развитый и массово используемый аппарат реляционных моделей данных. Унификация такого подхода, разработка методов и алгоритмов на базе теории реляционной алгебры и их программная реализация в отдельных продуктах (библиотеки,

классы, языки, инструментарии) могут быть эффективно применены в различных прикладных областях и технологиях обработки данных.

### Литература

1. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976. 162 с.
2. Burney A., Mahmood N., Jilani T., Saleem H. Conceptual fuzzy temporal relational model (ftm) for patient data. *WSEAS Transactions on Information Science and Applications*, 2010, vol. 7, no. 5, pp. 725–734.
3. Garrido C., Marin N., Pons O. Fuzzy intervals to represent fuzzy valid time in a temporal relational database. *Int. Journ. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2009, vol. 17 (suppl. 1), pp. 173–192.
4. Nagyp'al G., Motik B. A fuzzy model for representing uncertain, subjective, and vague temporal knowledge in ontologies. *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, OTM Confed. Int. Conf.*, 2003, pp. 906–923.
5. Pons J., Billiet C., Pons O., De Tré G. Aspects of dealing with imperfect data in temporal databases. *Flexible approaches in data, information and knowledge management. Springer Int. Publ., Swetzerland*, 2014, vol. 497, part. 3, pp. 189–220.
6. Galindo J., Urrutia A., Piattini M. Fuzzy databases: modeling, design and implementation. Idea Group Publ., Hershey, USA, 2006, pp. 150–151.
7. Pons J.E. Fuzzy temporal information treatment in relational databases. Thesis, Granada, Univ. Granada, 2014, 257 p.
8. Snodgrass R.T. Developing time-oriented database applications in SQL. Morgan Kaufmann Publ., Inc., SF, 1999, 504 p.
9. Kulkarni K.G., Michels J.-E. Temporal features in SQL: 2011. *ACM SIGMOD Record*, 2012, vol. 41, no. 3, pp. 34–43.
10. Allen J.F. Maintaining Knowledge about Temporal Intervals. *Commun. ACM*, 1983, vol. 26, no. 11, pp. 832–843.

DOI: 10.15827/0236-235X.115.029-035

Received 01.07.16

### IMPRECISE TEMPORAL INFORMATION PROCESSING

<sup>1</sup>Fomin V.V., Dr.Sc. (Engineering), Professor, v\_v\_fomin@mail.ru; <sup>1</sup>Maltsev S.V., Postgraduate Student, maltsevsergey@live.com

<sup>1</sup> Herzen State Pedagogical University of Russia, Reki Moyki Quay 48, St. Petersburg, 191186, Russian Federation

**Abstract.** The article presents the results of extensive research resulting from a practical need in improving the effectiveness of representation methods, extraction algorithms, processing fuzzy temporal data. The authors focus on ontological models of prediction and interpretation during interaction with imprecise temporal information. They propose a data storage methodology as well as models and unified procedures for extracting fuzzy information. The pragmatic of temporal data representation algorithms is considered using a new so-called possibilistic temporal information in analytical systems, where the parameters of inaccuracy and efficiency are crucial for making business decisions. The article proposes solutions for the problems of imprecise temporal information storage and retrieval by constructing a model of processing fuzzy temporal relations between time points and/or intervals. The concept of fuzzy temporal information analysis and presentation is described within development of temporal relational databases framework taking into account the criteria of inaccuracies, ambiguity and uncertainty. The authors developed the algebra of possibilistic relations between points and/or intervals. It defines the concept of possibilistic time point, possibilistic interval and corresponding values of possibilistic time points and intervals. The example of extension of one of the relations describes the fragment of a formalized algorithm of algebra implementation. The authors propose an approach to design and modeling a fuzzy temporal domain as well as the relations between possibilistic intervals and time points. The article describes an implemented library of algorithms for processing relations between possibilistic time points and/or intervals, as well as corresponding procedures and functions on SQL platform. At the end, the article describes the results of integration of the library of algorithms in a commercial project of analysis system for a telecommunication company. It also describes areas for further development of application tools based on the theory of a fuzzy-temporal model.

**Keywords:** temporal databases, temporal uncertainty, programs and algorithms.

### References

1. Zadeh L.A. *Ponyatie lingvisticheskoy peremennoy i ego primeneniye k prinyatiyu priblizhennykh resheniy* [The Concept of a Linguistic Variable and its Application to Approximate Reasoning]. Moscow, Mir Publ., 1976, 162 p.
2. Burney A., Mahmood N., Jilani T., Saleem H. Conceptual fuzzy temporal relational model (ftm) for patient data. *WSEAS Transactions on Information Science and Applications*, 2010, vol. 7, no. 5, pp. 725–734.
3. Garrido C., Marin N., Pons O. Fuzzy intervals to represent fuzzy valid time in a temporal relational database. *Int. Journ. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2009, vol. 17 (suppl. 1), pp. 173–192.
4. Nagyp'al G., Motik B. A fuzzy model for representing uncertain, subjective, and vague temporal knowledge in ontologies. *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, OTM Confed. Int. Conf.*, 2003, pp. 906–923.
5. Pons J., Billiet C., Pons O., De Tré G. Aspects of dealing with imperfect data in temporal databases. *Flexible Approaches in Data, Information and Knowledge Management. Springer Int. Publ., Swetzerland*, 2014, pp. 189–220.
6. Galindo J., Urrutia A., Piattini M. *Fuzzy Databases: Modeling, Design and Implementation*. Idea Group Publ., Hershey, USA, 2006, pp. 150–151.
7. Pons J.E. *Fuzzy Temporal Information Treatment in Relational Databases*. Thesis, Granada, Univ. Granada, 2014, 257 p.
8. Snodgrass R.T. *Developing Time-oriented Database Applications in SQL*. Morgan Kaufmann Publ., Inc., SF, 1999, 504 p.
9. Kulkarni K.G., Michels J.-E. Temporal features in SQL: 2011. *ACM SIGMOD Record*, 2012, vol. 41, no. 3, pp. 34–43.
10. Allen J.F. Maintaining Knowledge about Temporal Intervals. *Commun. ACM*, 1983, vol. 26, no. 11, pp. 832–843.

УДК 004.056

Дата подачи статьи: 21.03.16

DOI: 10.15827/0236-235X.115.036-041

**БЕЗОПАСНОСТЬ БАЗ ДАННЫХ: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ**

*М.А. Полтавцева, к.т.н., доцент, maria.poltavtseva@ibks.icc.spbstu.ru  
(Санкт-Петербургский политехнический университет Петра Великого,  
ул. Политехническая, 29, г. Санкт-Петербург, 195251, Россия);*

*А.Р. Хабаров, к.т.н., профессор, al\_xabarov@mail.ru  
(Тверской государственный технический университет,  
наб. Аф. Никитина, 22, г. Тверь, 170026, Россия)*

Вопросы защиты данных являются критическими при обеспечении безопасности современных корпоративных систем. Статья посвящена сложившейся ситуации в области безопасности данных, находящихся под управлением систем БД. Сделан обзор существующих работ на указанную тему как отечественных, так и зарубежных авторов. Рассмотрены этапы эволюции систем защиты информации в СУБД. Выделены современные комплексные проблемы в области защиты данных теоретического и практического характера. При анализе проблем безопасности учтены новые направления развития СУБД, такие как in memory и NoSQL-решения. В статье выделены особенности БД с точки зрения информационной безопасности по отношению к другим видам ПО. Предложено базовое деление уязвимостей СУБД на основании выделенных особенностей. Сформулированы перспективные пути исследований и развития систем защиты для построения надежных и защищенных серверов по обработке информации.

**Ключевые слова:** БД, СУБД, информационная безопасность, NoSQL, уязвимости, защита данных, безопасность данных, ПО.

Атаки на хранилища и БД являются одними из самых опасных для предприятий и организаций. Согласно статистике компании infowatch [1], в последние годы количество утечек данных в мире неуклонно растет, при этом на 2015 год более тридцати процентов из них приходится на внешних нарушителей и более шестидесяти выполнено с участием сотрудников организации. Даже если предположить, что в ряде случаев утечка включала данные, к которым сотрудник имеет легальный доступ, каждый третий случай приходился на внешнюю атаку. Также нужно отметить, что, согласно приведенным в [1] данным, на внешние атаки приходится семь из восьми утечек объемом более десяти миллионов записей.

Злоумышленников интересуют такие виды информации, как внутренняя операционная информация, персональные данные сотрудников, финансовая информация, информация о заказчиках/клиентах, интеллектуальная собственность, исследования рынка/анализ деятельности конкурентов, платежная информация [2]. Эти сведения в итоге хранятся в корпоративных хранилищах и БД различного объема.

Все это приводит к необходимости обеспечения защиты не только коммуникаций, операционных систем и других элементов инфраструктуры, но и хранилищ данных как еще одного барьера на пути злоумышленника. Однако на сегодняшний день работы в области обеспечения безопасности БД направлены в основном на преодоление существующих и уже известных уязвимостей, реализацию основных моделей доступа и рассмотрение вопросов, специфичных для конкретной СУБД. Целью данной работы являются комплексное рассмотрение и систематизация вопросов безопасности различных БД в свете новых угроз, общих тенденций

развития информационной безопасности и возрастающей роли и разнообразия хранилищ данных.

Вопросы комплексной безопасности БД привлекают внимание исследователей, им ежегодно посвящается ряд работ как в России, так и за рубежом. Можно отметить такие исследования, как классическая работа [3]. В ней рассматриваются подходы к обеспечению конфиденциальности, целостности и доступности СУБД, предотвращение, определение и игнорирование атак. Предлагаются подходы к обеспечению мандатного и ролевого дискреционного доступа к реляционному серверу. Данную тему развивает работа [4], затрагивающая те же вопросы обеспечения разделения доступа, привилегий, аудита и шифрования данных, а также вопросы применения для обеспечения защищенного доступа встроенных механизмов, таких как триггеры, представления и хранимые процедуры. Резюмирующая их работа [5] обобщает развитие подходов к безопасности в историческом аспекте.

Среди зарубежных работ, освещающих современные направления исследований, можно отметить [6, 7]. Работы российских исследователей в основном посвящены узким вопросам безопасности СУБД, например [8–12].

Нельзя не отметить также монографии [13, 14]. Однако эти работы, как и известные учебные пособия и материалы, в частности [15, 16], также не выходят за рамки вышеозначенных тем или же, как например [14], отражают специфику конкретной СУБД.

Аналогичную ситуацию можно наблюдать и в работе [17]. STIG включает известные вопросы безопасности и критерии уровневой сертификации программных средств СУБД, оценивая безопасность ПО на основании известных угроз, без учета специфики хранимых данных.



Таким образом, сегодняшние исследования в области безопасности СУБД ограничиваются развитием концепции конфиденциальности, целостности и доступности данных, что не соответствует современным требованиям к системам защиты и информационной безопасности программных решений (например [18]), причем в контексте конкретных методов защиты, а не целостного рассмотрения проблемы. При этом они зачастую посвящены конкретным программным продуктам, а не всему классу соответствующего ПО.

### **Эволюция систем безопасности БД**

Исторически развитие систем безопасности БД происходило как реакция на действия злоумышленников в соответствии с этапами эволюции самих хранилищ (БД) и изменениями типа и вида возрастающих угроз. Эти изменения были обусловлены общим развитием БД от решений на мейнфреймах до облачных хранилищ.

В архитектурном плане [5] можно выделить следующие подходы:

- полный доступ всех пользователей к серверу БД;
- разделение пользователей на доверенных и частично доверенных средствами СУБД (системы управления БД);
- введение системы аудита (логов действий пользователей) средствами СУБД;
- введение шифрования данных; вынос средств аутентификации за пределы СУБД в операционные системы и промежуточное ПО; отказ от полностью доверенного администратора данных.

Тем не менее, введение средств защиты как реакции на угрозы не обеспечивает защиту от новых способов атак и формирует разрозненное представление о самой проблеме обеспечения безопасности. С одной стороны, крупные компании могут выделить достаточное количество средств обеспечения безопасности для своих продуктов, с другой стороны, именно по этой причине имеется большое количество разнородных решений, отсутствует понимание комплексной безопасности данных (и ее компоненты разнятся от производителя к производителю), нет общего, единого подхода к безопасности хранилищ данных и, как следствие, возможности. Усложняются прогнозирование будущих атак и перспективная разработка защитных механизмов, для многих систем сохраняется актуальность уже давно известных атак, усложняется подготовка специалистов по безопасности.

Именно разработка программных средств перспективной защиты (на опережение злоумышленника), обеспечение возможности внедрения такой технологии представляются авторам статьи наиболее актуальными задачами на текущем этапе.

### **Современные проблемы обеспечения безопасности БД**

Список основных уязвимостей хранилищ данных, актуальный на сегодняшний день [19], не претерпел существенных изменений за последние более чем пять лет. Проанализировав средства обеспечения безопасности СУБД, архитектуру БД, интерфейсы, известные уязвимости и инциденты безопасности, можно выделить следующие причины возникновения такой ситуации:

- проблемами безопасности серьезно занимаются только крупные производители прежде всего в ведущих продуктах линейки для хранения данных;
- программисты БД, прикладные программисты и администраторы БД не уделяют должного внимания вопросам безопасности;
- разные масштабы и виды хранимых данных требуют разных подходов к безопасности;
- различные СУБД используют разные языковые диалекты для доступа к данным, организованным на основе одной и той же модели;
- появляются новые виды и модели хранения данных.

Рассмотрим эти положения более подробно на примере линейки продуктов от Oracle. СУБД Oracle Database Server имеет достаточно развитую систему безопасности, включающую основные и дополнительные модули и содержащую средства гранулирования доступа до уровня записи и маскирования данных. Отметим, что продукт компании СУБД MySQL не может похвастаться таким уровнем защищенности. Это достаточно серьезная проблема, так как MySQL – широко применяемая СУБД как в электронной коммерции, так и в БД государственных структур.

Многие уязвимости, обозначенные в исследованиях (например в [19]), сохраняют актуальность за счет невнимания или незнания администраторами систем БД вопросов безопасности. Например, известные техники простой SQL-инъекции широко эксплуатируются сегодня в отношении различных web и иных приложений, в которых не уделяется внимание контролю входных данных запроса. Причинами этого являются как недостаточная информированность или внимание администраторов СУБД и прикладных программистов, так и отсутствие встроенных средств контроля известных уязвимостей в большинстве СУБД. Хорошим решением были бы автоматизация и перенос контроля таких угроз на уровень сервера, однако многообразие языковых диалектов не позволяет это сделать.

Также нужно отметить, что применение различных средств обеспечения информационной безопасности является для организации компромиссом в финансовом плане: внедрение более защищенных продуктов и подбор более квалифицированного персонала требуют больших затрат. К тому же компоненты безопасности могут отрицательно влиять

на производительность систем управления БД, например уровни согласованности транзакций. Полное соответствие модели ACID – самый медленный способ обеспечения целостности при многопользовательской работе. Такие подходы, как маскирование данных или введение проверок безопасности доступа, также замедляют работу [13, 20].

Еще одна причина такой ситуации – разрозненность диалектов языка запросов к СУБД. Если рассматривать только известные реляционные СУБД, несмотря на наличие развивающегося стандарта SQL (SQL-92, SQL-99, SQL-2003, SQL-2008), даже крупные производители не только используют собственные расширения языка, но и не поддерживают до конца операции принятой версии стандарта. Этот факт осложняет разработку единых механизмов защиты БД уровня сервера.

Приведенные выше проблемы усугубляются с появлением и широким распространением нереляционных хранилищ данных и СУБД, оперирующих другой моделью данных, однако построенных по тем же принципам и имеющих аналогичное назначение, что и традиционные, реляционные серверы. По сути многообразие современных так называемых NoSQL (нереляционных) решений приводит к разнообразию применяемых моделей данных и размывает границу понятия БД.

Следствием этих проблем и отсутствия единых методик является нынешняя ситуация с безопасностью NoSQL-систем. Эти решения появились на рынке недавно и еще не успели пройти «путь ошибок и уязвимостей», характерный для их более зрелых реляционных аналогов. В большинстве NoSQL-систем отсутствуют не только общепринятые механизмы безопасности вроде шифрования, поддержки целостности и аудита данных, но даже развитые средства аутентификации пользователей.

#### **Особенности систем БД как объекта защиты**

В связи с появлением новых решений в области нереляционных хранилищ, размывающих границу традиционного представления о СУБД (например, система кэширования данных в памяти MemcacheDB или Hadoop HDFS), определим функции, отличающие СУБД от файлового хранилища и других типов программных продуктов. В этом ключе в [20] выделено несколько признаков. Переформулировав первый признак – «поддержание логически согласованного набора файлов», в силу активного развития in memory СУБД, осуществляющих хранение и все операции над данными в оперативной памяти, приведем эти критерии в следующей редакции:

- поддержание логически согласованного набора данных;
- обеспечение языка манипулирования данными;

– восстановление информации после разного рода сбоев;

– реальная параллельная работа нескольких пользователей (процессов).

Используя такой подход, можно отделить именно СУБД от файловых систем и ПО другого вида.

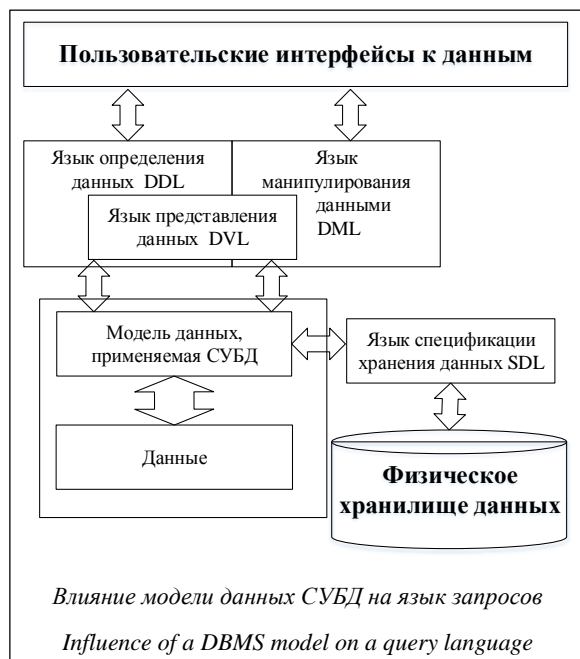
Отличительной особенностью систем БД от остальных видов прикладного ПО является (относительно информационной безопасности и не только) их двойственная природа. С этой точки зрения СУБД включает в себя два компонента: хранимые данные (собственно БД) и программы управления (СУБД).

Обеспечение безопасности хранимой информации, в частности, невозможно без обеспечения безопасного управления данными. Исходя из этой концепции, все уязвимости и вопросы безопасности СУБД можно разделить на две категории: зависящие от данных и независимые от данных.

Отметим, что уязвимости, независимые от данных (их структуры, организации и т.д.), являются характерными для всех прочих видов ПО. К этой группе можно отнести несвоевременное обновление ПО или наличие неиспользуемых функций.

Зависимыми от данных (в той или иной степени) является большое число аспектов безопасности. В частности, зависимыми напрямую можно назвать механизмы логического вывода и агрегирования данных, называемые специфичными проблемами СУБД. В то же время многие уязвимости являются косвенно зависимыми от данных. Например, современные СУБД (считая и реляционные, и нереляционные решения) поддерживают запросы к данным с использованием некоторого языка запросов. В свою очередь, в этом качестве используются специализированные языки запросов (SQL, CQL, OQL и других), наборы доступных пользователю функций (которые, в свою очередь, тоже можно считать операторами запросного языка) или произвольные функции на языке программирования (чаще всего Java). Обобщенные интерфейсы работы с данными представлены на рисунке.

Архитектура применяемых языков, по крайней мере, то, что касается специализированных языков (запросов) и наборов функций, напрямую связана с моделью данных, применяемой для хранения информации. Таким образом, модель определяет особенности языка, а особенности языка – наличие в нем тех или иных уязвимостей. Причем уязвимости общего типа, например инъекция (под инъекцией будем понимать атаку на БД путем модификации входных запросов, заставляющую сервер СУБД выполнить нелегитимный набор действий), выполняются по-разному (SQL-инъекция, JAVA-инъекция) в зависимости от синтаксиса и семантики языка, которые, как уже сказано выше, отчасти определяются моделью данных и, следовательно, являются зависимым от данных компонентом.



### Требования к безопасности БД

Таким образом, на основании разделения уязвимостей можно выделить зависимые и независимые от данных меры обеспечения безопасности хранилищ информации. Независимыми от данных можно назвать следующие требования к безопасной системе БД.

- Функционирование в доверенной среде.

Под доверенной понимается информационная среда, интегрирующая совокупность защитных механизмов, которые обеспечивают обработку информации без нарушения политики безопасности [21]. В данном случае СУБД должна функционировать в доверенной информационной системе с соответствующими методами обмена данными.

- Организация физической безопасности файлов данных.

Данный вопрос требует более детального изучения, так как применяемые структуры данных в различных моделях данных СУБД могут иметь значение при шифровании и защите файлов данных. Однако в первом приближении вопрос физической безопасности файлов данных сходен с вопросом физической безопасности любых других файлов пользователей и приложений.

- Организация безопасной и актуальной настройки СУБД.

К данному аспекту относятся такие общие вопросы обеспечения безопасности, как своевременная установка обновлений, отключение неиспользуемых модулей или применение эффективной политики паролей.

Следующие требования можно назвать зависимыми от данных.

- Безопасность пользовательского слоя ПО.

К этой категории относятся задачи построения безопасных интерфейсов и вызовов (в том числе с учетом интерфейса СУБД и механизма доступа к данным).

- Безопасная организация данных и манипулирование ими.

Вопрос организации данных и управления ими является ключевым в системах хранения информации. Несмотря на то, что в приведенном перечне он указан последним, именно в эту область входят задачи организации данных с контролем целостности, обеспечение защиты от логического вывода и другие, специфичные для СУБД проблемы безопасности. Фактически эта задача включает в себя основной пул зависимых от данных уязвимостей и защиты от них.

### Пути создания защищенных БД

Для преодоления названных проблем обеспечения информационной безопасности СУБД необходимо перейти от метода закрытия уязвимостей к комплексному подходу обеспечения безопасности хранилищ информации. Основными этапами этого перехода, по мнению авторов, должны стать следующие положения.

1. Разработка комплексных методик обеспечения безопасности хранилищ данных на текущем этапе.

Создание комплексных методик позволит применять их (или их соответствующие версии) при разработке хранилищ данных и пользовательского ПО. Основой для создания таких документов могут стать обобщающие проблематику работы, например [16] или [22]. Следование комплексной методике позволит избежать многих ошибок управления СУБД и защититься от наиболее распространенных на сегодняшний день уязвимостей.

2. Оценка и классификация угроз и уязвимостей СУБД.

Специализированная классификация угроз и уязвимостей СУБД позволит упорядочить их для последующего анализа и защиты, даст возможность установить зависимость между уязвимостями и причинами (источниками) их возникновения. В результате при введении конкретного механизма в СУБД появится возможность установить и спрогнозировать связанные с ним угрозы и заранее подготовить соответствующие средства обеспечения безопасности.

3. Разработка стандартных (применимых к различным СУБД без внесения изменений или с минимальными изменениями) механизмов обеспечения безопасности.

Стандартизация подходов и языков работы с данными позволит создать мультиплатформенные средства обеспечения безопасности, применимые к разным СУБД. С одной стороны, это методические и теоретические подходы, применимые в рамках

модели данных. На сегодняшний день есть разработки таких механизмов по реляционной модели, однако они не решают всех насущных вопросов безопасности. С другой – это разработка теоретического базиса для новых СУБД, в частности, конкретизация и формализация агрегатных моделей данных. Появление готовых программных средств во многом зависит от производителей и разработчиков СУБД и их следования стандартам, а также достаточности определенных в стандарте средств для построения развитых механизмов безопасности.

4. Разработка теоретической базы информационной защиты систем хранения и манипулирования данными.

Выше отмечен ряд характерных особенностей, специфичных для хранилищ данных: двойственная природа СУБД, зависимость уязвимостей и механизмов управления от данных и описывающей их организацию модели, угрозы логического вывода, различная значимость сочетаний данных. Все это определяет специфический характер безопасности СУБД и требует новых теоретических подходов к обеспечению защиты данных в программных системах такого рода. Отдельный большой вопрос – развитие теоретического базиса в контексте формализации модели данных, а также разработка подходов обеспечения целостности информации для новых NoSQL-хранилищ.

На основании изложенного сделаем следующие выводы. В результате рассмотрения российских и зарубежных работ, а также ситуации на рынке СУБД в статье выделены текущий подход к обеспечению безопасности, принцип и основные эволюционные этапы систем БД. Сформулированы проблемы информационной безопасности современных СУБД: разнообразие языковых средств, появление новых моделей данных, не подкрепленных теоретическим базисом, необходимость поиска баланса между безопасностью и ее стоимостью, развитие систем защиты как реакции на потерю средств и престижа, а также общее невнимание к вопросам безопасности.

Сформулированы критерии, выделяющие СУБД из сходных программных продуктов, с учетом новых кластерных и in memoгу решений, особенности этого класса ПО с точки зрения информационной безопасности и предложено базовое деление уязвимостей на зависимые и независимые от данных и их организации.

В результате сформулированы общие требования к безопасности БД, перспективные пути исследования и развития систем защиты для построения надежных и защищенных серверов по обработке информации. В них вошли как систематизация и развитие существующих подходов в виде разработки методик и стандартизации механизмов защиты, так и направления новых исследований, например, классификация уязвимостей СУБД и

формализация новых моделей данных, построение прогнозирующих средств защиты.

### Литература

1. Исследование утечек информации за первое полугодие 2015 года. URL: <http://www.infowatch.ru/analytics/reports/16340> (дата обращения: 26.02.2016).
2. Информационная безопасность бизнеса. Исследование текущих тенденций в области информационной безопасности бизнеса. 2014. URL: [http://media.kaspersky.com/pdf/IT\\_risk\\_report\\_Russia\\_2014.pdf](http://media.kaspersky.com/pdf/IT_risk_report_Russia_2014.pdf) (дата обращения: 26.02.2016).
3. Sandhu Ravi S., Jajodia Sushil. Data and database security and controls. Handbook of Information Security Management, Auerbach Publishers, 1993, pp. 181–199.
4. Qiu M., Davis S. Database security mechanisms and implementation. IACIS, Issues in Inform. Syst. 2002, vol. 03, pp. 529–534.
5. Lesov P. Database security: a historical perspective. 2010. URL: <http://arxiv.org/ftp/arxiv/papers/1004/1004.4022.pdf> (дата обращения: 26.02.2016).
6. Burtescu E. Database security – attacks and control methods. Journ. of Applied Quantitative Methods, 2009, vol. 4, no. 4, pp. 449–454.
7. Rohilla S., Mittal P.K. Database Security: Threads and Challenges. Intern. Journ. of Advanced Research in Computer Science and Software Engineering, 2013, vol. 3, iss. 5, pp. 810–813.
8. Потапов А.Е., Манухина Д.В., Соломатина А.С., Бадмаев А.И., Яковлев А.В., Нилова А.С. Безопасность локальных баз данных на примере SQL Server Compact // Вестн. Тамбов. ун-та. Серия: Естественные и технические науки. 2014. № 3. С. 915–917.
9. Бортовчук Ю.В., Крылова К.А., Ермолаева Л.В. Информационная безопасность в современных системах управления базами данных // Современные проблемы экономического и социального развития. 2010. № 6. С. 224–225.
10. Горбачевская Е.Н., Катянов А.Ю., Краснов С.С. Информационная безопасность средствами СУБД Oracle // Вестн. ВУиТ. 2015. № 2 (24). С. 72–85.
11. Ткаченко Н.О. Реализация монитора безопасности СУБД MySQL в dbf/dam системах // ПДМ. Приложение. 2014. № 7. С. 99–101.
12. Полтавцева М.А. Задача хранения прав доступа к данным в РСУБД на примере Microsoft SQL Server // Актуальные направления фундаментальных и прикладных исследований: матер. V Междунар. науч.-практ. конф. 2015. С. 118–120.
13. Баранчиков А.И., Баранчиков П.А., Пылкин А.Н. Алгоритмы и модели доступа к записям БД. М.: Горячая линия–Телеком, 2011. 182 с.
14. Поляков А.М. Безопасность Oracle глазами аудитора: нападение и защита. М.: ДМК Пресс, 2014. 336 с.
15. Смирнов С.Н. Безопасность систем баз данных. М.: Гелиос АРВ, 2007. 352 с.
16. Murray M.C. Database security: what students need to know. JITE:IP, vol. 9, 2010, pp. 61–77.
17. Database Security Technical Implementation Guide (STIG). US Department of Defense. Vers. 7. Release 1. 2004. URL: [https://www.computer.org/cms/s2esc/s2esc\\_excom/Minutes/2005-03/DISA%20STIGs/DATABASE-STIG-V7R1.pdf](https://www.computer.org/cms/s2esc/s2esc_excom/Minutes/2005-03/DISA%20STIGs/DATABASE-STIG-V7R1.pdf) (дата обращения: 26.02.2016).
18. Зегжда П.Д. Обеспечение безопасности информации в условиях создания единого информационного пространства // Защита информации. Инсайд. 2007. № 4 (16). С. 28–33.
19. Top Ten Database Security Threats. URL: [http://www.imperiva.com/docs/wp\\_topten\\_database\\_threats.pdf](http://www.imperiva.com/docs/wp_topten_database_threats.pdf) IMPREVA 2015 (дата обращения: 26.02.2016).
20. Кузнецов С.Д. Базы данных: учебник для студ. М.: Академия, 2012. 496 с.
21. Зегжда Д.П., Калинин М.О. Обеспечение доверенности информационной среды на основе расширения понятия «целостность» и управления безопасностью // Проблемы информ. безопасности. Компьютерные системы. 2009. № 4. С. 7–16.
22. Полтавцева М.А., Зегжда Д.П., Супрун А.Ф. Безопасность баз данных: учеб. пособие. СПб: Изд-во СПбПУ, 2015. 125 с.

DOI: 10.15827/0236-235X.115.036-041

Received 21.03.16

## DATABASE SECURITY: PROBLEMS AND PROSPECTS

<sup>1</sup>*Poltavtseva M.A., Ph.D. (Engineering), Associate Professor, maria.poltavtseva@ibks.icc.spbstu.ru*<sup>2</sup>*Khabarov A.R., Ph.D. (Engineering), Professor, al\_xabarov@mail.ru*<sup>1</sup>*Peter the Great Saint-Petersburg Polytechnic University,  
Politechnicheskaya St. 29, St. Petersburg, 195251, Russian Federation*<sup>2</sup>*Tver State Technical University, Nikitin Quay 22, Tver, 170026, Russian Federation*

**Abstract.** Data protection issues are critical when ensuring modern enterprise systems security. The paper is devoted to the current situation in database systems security. The paper provides an overview of existing publications, both Russian and foreign authors. It reviews the stages of a database information security systems evolution. The paper describes modern complex problems in DBMS data protection, both theoretical and practical. The analysis of security issues takes into account new directions of the database development, such as “in memory” and NoSQL DBMS. The article describes database features in terms of information security compared with other types of software. It includes the basic database vulnerabilities division based on selected features. The authors formulate promising ways of research and development of security systems for building reliable and secure information management systems.

**Keywords:** databases, DBMS, information security, NoSQL, vulnerabilities, data protection, data security, software.

## References

1. *Issledovanie utechek informatsii za pervoe polugodie 2015 goda* [A Research on Information Leakage in the First Half of 2015]. Available at: <http://www.infowatch.ru/analytics/reports/16340> (accessed February 26, 2016).
2. *Buziness Data Security. A Research on current trends in Buziness Data Security. Informatsionnaya bezopasnost biznesa* [Business Data Security]. 2014. Available at: [http://media.kaspersky.com/pdf/IT\\_risk\\_report\\_Russia\\_2014.pdf](http://media.kaspersky.com/pdf/IT_risk_report_Russia_2014.pdf) (accessed February 26, 2016).
3. Sandhu Ravi S., Jajodia Sushil. *Data and Database Security and Controls. Handbook of Information Security Management*. Auerbach Publ., 1993.
4. Qiu M., Davis S. Database Security Mechanisms and Implementation. *Issues in Information Systems*. 2002, vol. 3, pp. 529–534.
5. Lesov P. *Database security: a historical perspectiv*. CoRR, 2010. Available at: <http://arxiv.org/ftp/arxiv/papers/1004/1004.4022.pdf> (accessed February 26, 2016).
6. Burtescu E. Database security – attacks and control methods. *Journ. of Applied Quantitative Methods*. 2009, vol. 4, no. 4, pp. 449–454.
7. Rohilla S., Mittal P.K. Database Security: Threads and Challenges. *Int. Journ. of Advanced Research in Computer Science and Software Engineering*. 2013, vol. 3, iss. 5, pp. 810–813.
8. Potapov A.E., Manukhina D.V., Solomatina A.S., Badmaev A.I., Yakovlev A.V., Nilova A.S. Lokal databases' security on the example of SQL Server Compact. *Vestnik Tambovskogo universiteta. Ser. Estestvennye i tekhnicheskie nauki* [The Bulletin of Tambov Univ. Series Natural and Technical Sciences]. 2014, no. 3, pp. 915–917 (in Russ.).
9. Bortovchuk Yu.V., Krylova K.A., Ermolaeva L.V. Information security in modern data base management systems. *Sovremennye problemy ekonomicheskogo i sotsialnogo razvitiya* [Modern Problems of Economic and Social Development]. 2010, no. 6, pp. 224–225 (in Russ.).
10. Gorbachevskaya E.N., Katyanov A.Yu., Krasnov S.S. Information security using Oracle DBMS. *Vestnik VUit* [Vestnik of Volzhsky University after V.N. Tatischev]. 2015, no. 2 (24), pp. 72–85 (in Russ.).
11. Tkachenko N.O. Implementation of RDBMS MySQL security monitor in DBF/DAM systems. *PDM* [Applied discrete mathematics]. Supplement. 2014, no. 7, pp. 99–101 (in Russ.).
12. Poltavtseva M.A. *Aktualnye napravleniya fundamentalnykh i prikladnykh issledovaniy* [Proc. 5th Int. Science and Practice Conf. “Important Trends of Fundamental and Applied Research”]. 2015, pp. 118–120 (in Russ.).
13. Baranchikov A.I., Baranchikov P.A., Pylin A.N. *Algorithms and database record access models*. Moscow, Goryachaya liniya–Telekom Publ., 2011, 182 p. (in Russ.).
14. Polyakov A.M. *Oracle security through the eyes of an auditor: attack and defense*. Moscow, DMK Press, 2014, 336 p. (in Russ.).
15. Smirnov S.N. Database Systems' Security. Moscow, Gelios ARV Publ., 2007, 352 p. (in Russ.).
16. Murray M.C. Database Security: What Students Need To Know. *Journ. of Information Technology Education: Innovations in Practice (JITE:IIP)*. 2010, vol. 9, pp. 61–77.
17. *Database Security Technical Implementation Guide (STIG)* US Department of Defense. Ver. 7, Release 1, 2004. available at: [https://www.computer.org/cms/s2esc/s2esc\\_excom/Minutes/2005-03/DISA%20STIGs/DATABASE-STIG-V7R1.pdf](https://www.computer.org/cms/s2esc/s2esc_excom/Minutes/2005-03/DISA%20STIGs/DATABASE-STIG-V7R1.pdf) (accessed February 26, 2016).
18. Zegzhda P.D. Information security ensuring in the context of creating single information space. *Zashchita informatsii. Insayd* [Information Security. Inside]. 2007, no. 4 (16), pp. 28–33 (in Russ.).
19. *Top ten database security threats*. Available at: [http://www.imperva.com/docs/wp\\_topen\\_database\\_threats.pdf](http://www.imperva.com/docs/wp_topen_database_threats.pdf) IMPREVA 2015 (accessed February 26, 2016).
20. Kuznetsov S.D. *Databases*. Textbook for students. Moscow, Akademiya Publ., 2012, 496 p. (in Russ.).
21. Zegzhda D.P., Kalinin M.O. Ensuring information environment confidence based on na extending the concept of “completeness” and security control. *Problemy informatsionnoy bezopasnosti. Kompyuternye sistemy* [Information Security Problems. Computer Systems]. 2009, no. 4, pp. 7–16 (in Russ.).
22. Poltavtseva M.A., Zegzhda D.P., Suprun A.F. *Database security*. Study guide. St. Petersburg, Politekh. un-t Publ., 2015, 125 p. (in Russ.).

УДК 004.056

Дата подачи статьи: 16.03.16

DOI: 10.15827/0236-235X.115.042-050

**ПОСТРОЕНИЕ МОДЕЛИ УГРОЗ БЕЗОПАСНОСТИ ИНФОРМАЦИИ  
В АВТОМАТИЗИРОВАННОЙ СИСТЕМЕ УПРАВЛЕНИЯ  
КРИТИЧЕСКИ ВАЖНЫМИ ОБЪЕКТАМИ  
НА ОСНОВЕ СЦЕНАРИЕВ ДЕЙСТВИЙ НАРУШИТЕЛЯ**

*Е.Б. Дроботун, к.т.н., докторант, drobotun@xakep.ru;*

*О.В. Цветков, начальник службы защиты государственной тайны*

*(Военная академия воздушно-космической обороны им. Маршала Советского Союза Г.К. Жукова,  
ул. Жигарева, 50, г. Тверь, 170100, Россия)*

Разработка и построение модели угроз безопасности информации в различных автоматизированных и информационно-вычислительных системах является одним из ключевых этапов деятельности по обеспечению безопасности информации в этих системах и их безопасного и правильного функционирования. На основе модели угроз безопасности информации определяются требования к системе защиты информации в автоматизированных и информационно-вычислительных системах, а также может быть определен комплекс организационных, технических и организационно-технических мер по обеспечению безопасности информации в автоматизированных и вычислительных системах различного назначения.

В статье предлагается методика построения модели угроз для автоматизированных систем управления критически важными и потенциально опасными объектами, основанная на моделировании возможных сценариев действий нарушителя безопасности информации.

Предложенная методика учитывает многоуровневость построения современных автоматизированных систем управления критически важными объектами и возможность реализации одной и той же угрозы безопасности информации с помощью различных сценариев нарушениями различных категорий.

Применение предложенной методики позволит сократить время, необходимое для построения модели угроз безопасности информации в автоматизированных системах управления критически важными и потенциально опасными объектами.

**Ключевые слова:** критическая информационная система, угроза безопасности информации, нарушитель безопасности информации, модель угроз, модель нарушителя.

Одной из основных задач обеспечения безопасности информации в АСУ *критически важными объектами* (КВО) является определение перечня угроз безопасности информации с оценкой рисков воздействия актуальных угроз на защищаемую АСУ КВО. Решение данных задач позволяет определить требования к системе защиты АСУ КВО, сформировать перечень защитных мер и в итоге определить оптимальную структуру системы защиты АСУ КВО от различного рода информационно-технических воздействий.

Согласно ГОСТ Р 50922-2006, под угрозой безопасности информации понимается совокупность условий и факторов, создающих потенциальную или реально существующую опасность нарушения конфиденциальности, доступности и (или) целостности информации.

Поскольку при обеспечении безопасности АСУ КВО, помимо конфиденциальности, целостности и доступности информации, необходимо, в первую очередь, обеспечить безопасность производственного или технологического процесса, реализуемого АСУ КВО [1], при определении угроз безопасности информации первоочередное внимание необходимо уделять оценке угроз, связанных с целенаправленными воздействиями на АСУ КВО, реализуемыми посредством программных (программно-технических) средств и предпринимаемыми в

целях нарушения или прекращения их функционирования (компьютерные атаки).

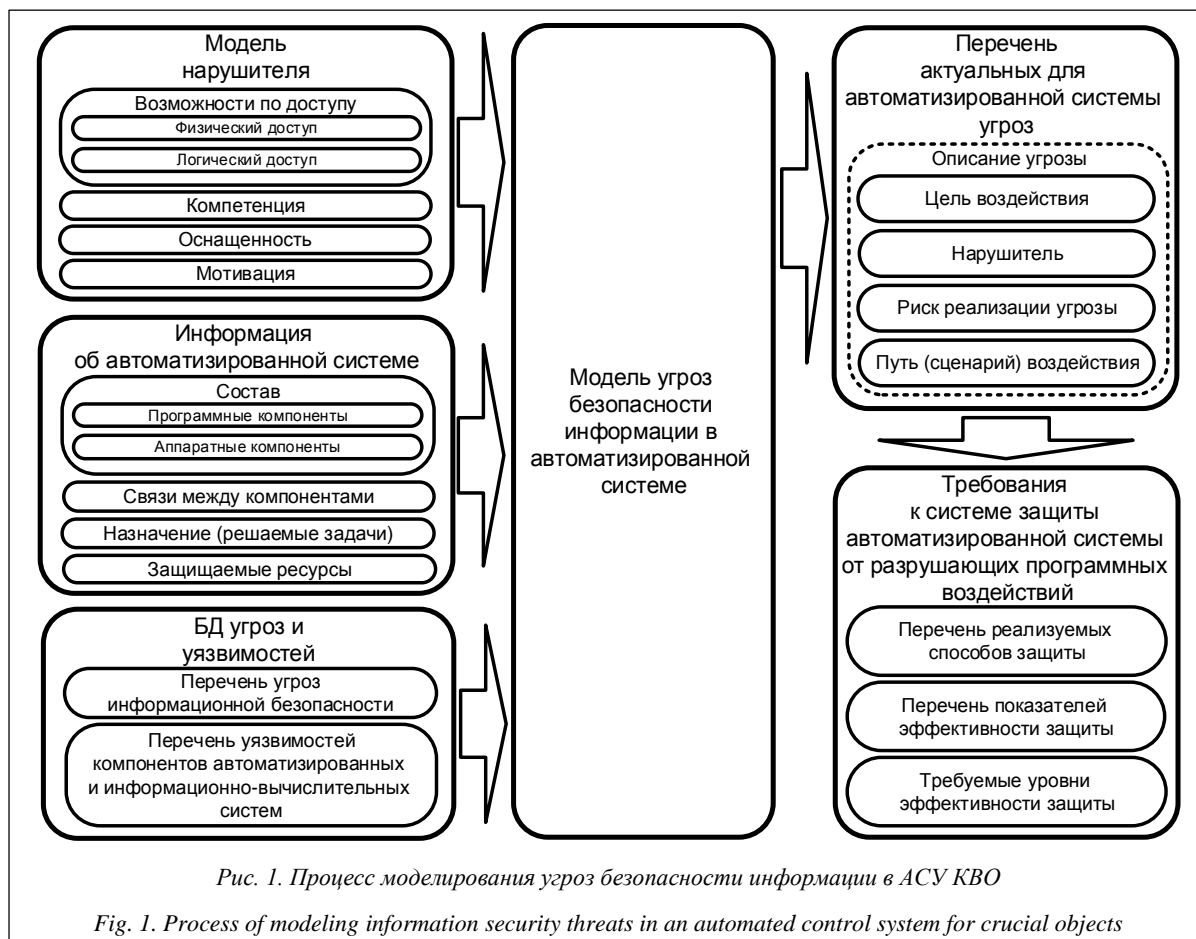
Процесс оценки угроз безопасности информации и формирования перечня угроз, актуальных для защищаемой АСУ КВО, называется моделированием угроз безопасности информации, а результатом этого процесса является модель угроз, которая представляет собой физическое, математическое или описательное представление свойств или характеристик угроз безопасности информации.

Целью моделирования угроз безопасности информации является определение системы конкретных требований к защите информации в АСУ КВО, обеспечивающих ее безопасное функционирование, и создание на основе этих требований адекватной системы защиты информации в АСУ КВО (рис. 1).

Исходные данные для моделирования угроз безопасности информации в АСУ КВО: модель нарушителя безопасности информации, информация об АСУ КВО, БД угроз безопасности информации и уязвимостей компонентов информационно-вычислительных систем (в том числе и АСУ КВО).

**Модель нарушителя безопасности информации.** В качестве нарушителей информационной безопасности АСУ КВО могут выступать лица

– осуществляющие преднамеренные действия с целью доступа к информации (воздействия на ин-



формацию), содержащейся в АСУ КВО, или нарушения функционирования АСУ КВО или обслуживающей ее инфраструктуры (преднамеренные угрозы безопасности информации);

- имеющие доступ к АСУ КВО, непреднамеренные действия которых могут привести к нарушению безопасности информации (непреднамеренные угрозы безопасности информации).

Целью моделирования нарушителя безопасности информации является формирование предположений о потенциальных возможностях реальных нарушителей при реализации ими угроз безопасности информации в АСУ КВО.

Результатом моделирования нарушителя безопасности информации является его модель, которая входит в модель угроз безопасности АСУ КВО в качестве одной из составных частей (рис. 2) и содержит:

- уровень физического доступа (*PhL*) к информации и (или) к компонентам АСУ КВО;
- уровень логического доступа (*LogL*) к информации и (или) к компонентам АСУ КВО;
- уровень компетенции (*C*) нарушителя безопасности информации;
- уровень оснащенности (*A*) нарушителя безопасности информации;
- мотивацию (*M*) нарушителя безопасности информации.

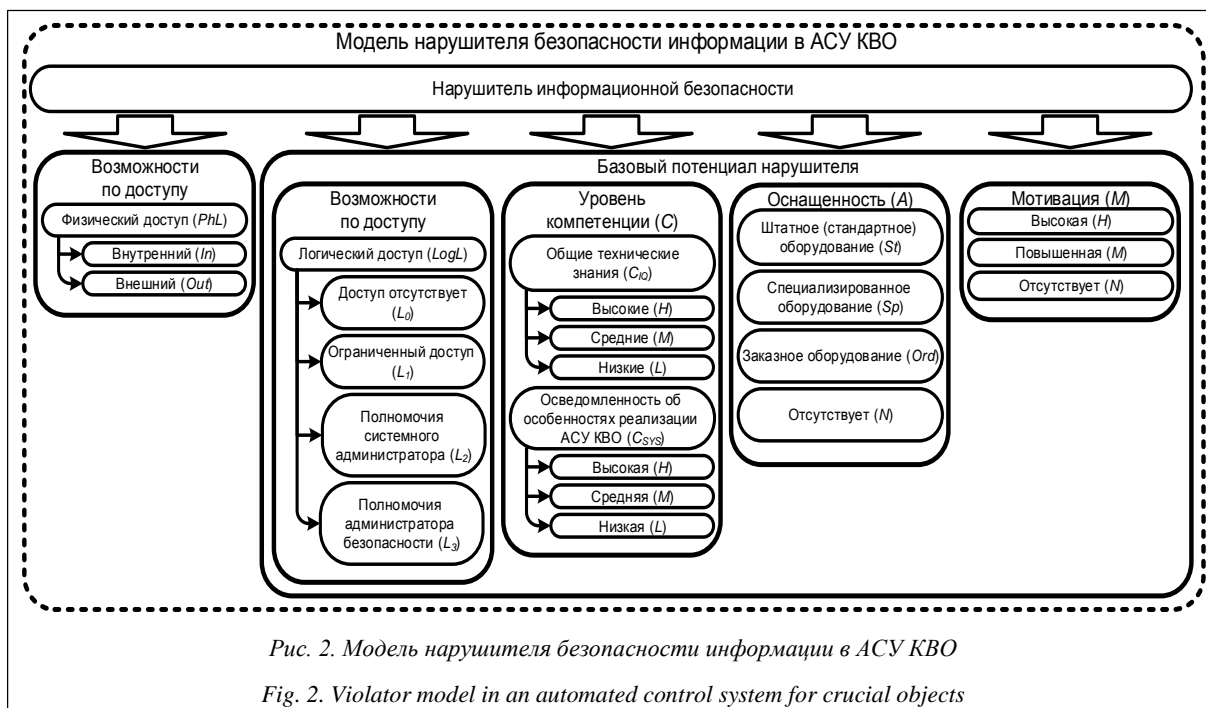
Общая запись модели нарушителя производится в виде строки  $PhL:[x]/LogL:[x]/C:[x]/A:[x]/M:[x]$ , где вместо  $[x]$  записывается значение соответствующего параметра.

Уровень физического доступа определяется исходя из наличия возможностей физического доступа к информации и (или) компонентам АСУ КВО. С учетом этого можно выделить два уровня нарушителей безопасности информации в АСУ КВО [2–4]:

- внутренние нарушители (*In*) – лица, имеющие возможность постоянного или разового доступа к информационной системе или ее отдельным компонентам;
- внешние нарушители (*Out*) – лица, не имеющие возможности доступа к информационной системе или к ее отдельным компонентам и реализующие угрозы безопасности информации дистанционно (удаленно), находясь вне информационной системы.

Уровень логического доступа определяется с учетом прав доступа субъектов (физических лиц) к объектам АСУ КВО [2, 4]. Исходя из этого можно выделить следующие уровни логического доступа:

- лица, не имеющие доступа к объектам АСУ КВО (*L<sub>0</sub>*);
- зарегистрированные пользователи АСУ КВО, имеющие ограниченный доступ к объектам



АСУ КВО с автоматизированного рабочего места ( $L_1$ );

- зарегистрированные пользователи АСУ КВО, имеющие доступ к объектам АСУ КВО с полномочиями системного администратора АСУ КВО ( $L_2$ );

- зарегистрированные пользователи АСУ КВО, имеющие доступ к объектам АСУ КВО с полномочиями администратора безопасности АСУ КВО ( $L_3$ ).

Уровень компетенции нарушителя безопасности информации определяется исходя из предположений об общих технических знаниях нарушителя и его осведомленности об особенностях построения и эксплуатации атакуемой АСУ КВО, а также из специфики ее функционирования.

Общие технические знания нарушителя ( $C_{IQ}$ ) (по ГОСТ Р ИСО/МЭК 18045-2008) оцениваются тремя уровнями:

- высокий – нарушитель имеет высокую осведомленность о способах и средствах защиты информации, применяемых в различных информационных системах, а также обладает знаниями о методах выявления уязвимостей и реализации угроз безопасности в различных автоматизированных системах;

- средний – нарушитель имеет осведомленность о способах и средствах защиты информации, применяемых в различных автоматизированных системах;

- низкий – нарушитель имеет слабую осведомленность о способах и средствах защиты информации, применяемых в различных информационных системах, и не обладает знаниями о реализации угроз безопасности.

Каждому из вышеуказанных уровней общих технических знаний присваивается численное значение. Так, численные значения характеристик уровней общих технических знаний нарушителя безопасности информации АСУ КВО следующие: высокий уровень – 5, средний – 2, низкий – 0.

Осведомленность нарушителя об особенностях построения и эксплуатации атакуемой АСУ КВО ( $C_{SYS}$ ) также оценивается тремя уровнями:

- высокий – нарушитель имеет возможность получения доступа к сведениям о структуре и функциональных связях в АСУ КВО, системе защиты информации в АСУ КВО, а также к сведениям из проектной, конструкторской и эксплуатационной документации;

- средний – нарушитель имеет возможность получения доступа к сведениям о целях и задачах, решаемых АСУ КВО, а также к сведениям из эксплуатационной документации;

- низкий – нарушитель не имеет возможности получения доступа к сведениям о структуре и функциональных связях в АСУ КВО, системе защиты информации в АСУ КВО, а также к сведениям из проектной, конструкторской и эксплуатационной документации.

Каждому из вышеуказанных уровней осведомленности нарушителя об особенностях реализации атакуемой АСУ КВО присваивается численное значение. Так, численные значения характеристик уровней осведомленности нарушителя об особенностях реализации атакуемой АСУ КВО следующие: высокий уровень – 5, средний – 2, низкий – 0.

Общий уровень компетенции (C) нарушителя безопасности информации в АСУ КВО оценивается исходя из суммы численных значений  $C_{IQ}$  и



$C_{SYS}$  следующим образом:

$$C = \begin{cases} \text{высокий (H)} & \text{при } (C_{IQ} + C_{SYS}) > 5, \\ \text{средний (M)} & \text{при } 2 < (C_{IQ} + C_{SYS}) \leq 5, \\ \text{низкий (L)} & \text{при } (C_{IQ} + C_{SYS}) \leq 2. \end{cases}$$

Мотивация – действенный фактор потенциала нарушителя, который может быть использован для описания различных аспектов, относящихся к нарушителю, а также к ресурсам АСУ КВО, интересующим нарушителя. Мотивация может подразумевать определенную вероятность нападения: из угрозы, оцененной как высокомотивированная, можно предположить, что атака на АСУ КВО неизбежна или что вследствие немотивированной угрозы нападения не ожидается.

В целом мотивация нарушителя безопасности информации в АСУ КВО может быть оценена тремя уровнями:

- высокая (H) – имеются сведения об устремлениях нарушителей к конкретной информационной системе, отдельным ее компонентам или имеются сведения, полученные от органов исполнительной власти, о намерениях нарушителя осуществить неправомерные действия в отношении АСУ КВО или информации, содержащейся в ней;
- повышенная (M) – нарушитель действует из каких-либо личных побуждений (любопытство, желание самоутвердиться и т.п.);
- отсутствует (N) – нарушения безопасности информации происходят вследствие непреднамеренных, неосторожных (ошибочных) или некачественных действий.

Оснащенность нарушителя характеризует возможность его доступа к программным и (или) программно-аппаратным средствам, которыми он может воспользоваться при реализации угроз безопасности в АСУ КВО:

- стандартное (штатное) оборудование (St) – программные и (или) программно-аппаратные средства, легкодоступные для нарушителя (эти средства могут входить в состав самой АСУ КВО, а также могут быть легко получены, как, например, известные и доступные программные средства, которые можно загрузить из Интернета);
- специализированное оборудование (Sp) – программные и (или) программно-аппаратные средства, ограниченно доступные нарушителю, но которые он может приобрести без значительных усилий;
- заказное оборудование (Ord) – программные и (или) программно-аппаратные средства, не доступные широкому кругу, поскольку либо может потребоваться их специальная разработка, либо распространение данных средств является контролируемым и ограниченным, либо данные средства имеют очень высокую стоимость;
- отсутствует (N) – доступ к каким-либо программным и (или) программно-аппаратным средствам у нарушителя отсутствует.

Возможности нарушителя по логическому доступу к компонентам системы, уровень его компетенции и оснащенности, а также его мотивация определяют базовый потенциал ( $Pt_{BASE}$ ), которым обладает нарушитель безопасности информации в АСУ КВО и который используется на этапе моделирования угроз при оценке вероятностей их реализации. Оценка предварительного базового потенциала нарушителя производится на основе числовых значений уровней компетенции и оснащенности в соответствии с таблицей 1.

Таблица 1

**Численные значения характеристик уровней компетенции нарушителя и его оснащенности**

Table 1

**Numerical values of violator's competence level characteristics and his equipment status**

Уровень	Показатель возможностей нарушителя	Численное значение характеристики
Уровень компетенции (C)	низкий	0
	средний	2
	высокий	5
Уровень оснащенности (A)	отсутствует	0
	стандартное оборудование	1
	специализированное оборудование	3
	заказное оборудование	5

Полученные из таблицы 1 значения суммируются, и определяется предварительный базовый потенциал нарушителя:

$$Pt_{BASE} = \begin{cases} \text{высокий (H)} & \text{при } (C + A) \geq 8, \\ \text{средний (M)} & \text{при } 5 \leq (C + A) < 8, \\ \text{низкий (L)} & \text{при } 3 \leq (C + A) < 5, \\ \text{отсутствует (N)} & \text{при } (C + A) < 3. \end{cases}$$

Далее в соответствии с таблицей 2 определяется базовый потенциал нарушителя с учетом уровней его логического доступа к объектам АСУ КВО.

В случае высокой мотивации потенциал нарушителя переходит на следующий уровень (от низкого к среднему или от среднего к высокому), в случае отсутствия мотивации потенциал от среднего или высокого переходит к низкому. В случае отсутствия у нарушителя возможностей для реализации угроз (базовый потенциал имеет уровень «отсутствует») независимо от мотивации нарушителя базовый потенциал остается на уровне «отсутствует».

Примеры описания различных категорий нарушителей в соответствии с предлагаемой моделью нарушителя безопасности информации в АСУ КВО приведены в таблице 3.

**Информация об АСУ КВО.** При моделировании угроз безопасности информации в АСУ КВО

Таблица 2

**Определение базового потенциала нарушителя безопасности информации с учетом прав  
логического доступа к объектам АСУ КВО**

Table 2

**Determination of information security violator's basic potential taking into account logical access rights  
to objects of an automated control system for crucial objects**

Уровень логического доступа нарушителя к объектам АСУ КВО	Предварительный базовый потенциал нарушителя безопасности информации в АСУ КВО			
	отсутствует	низкий	средний	высокий
$L_0$	отсутствует	низкий	средний	высокий
$L_1$	отсутствует	низкий	средний	высокий
$L_2$	низкий	средний	высокий	высокий
$L_3$	низкий	средний	высокий	высокий

Таблица 3

**Примеры описания различных категорий  
нарушителей безопасности информации в АСУ КВО**

Table 3

**Case descriptions of various categories  
of information security violators in an automated control system for crucial objects**

Категория нарушителя	Мотив действий нарушителя	Общая запись модели нарушителя	Базовый потенциал
Специальные службы иностраных государств	Нанесение ущерба государству, отдельным сферам его деятельности или секторам экономики	$PhL:Out/LogL:L_0/C:H/A:Ord/M:H$	высокий
Разработчики, произво- дители и поставщики компонентов АСУ КВО	Причинение имущественного ущерба путем обмана или злоупотребления доверием	$PhL:Out/LogL:L_3/C:H/A:Sp/M:M$	высокий
Оператор АСУ КВО	Непреднамеренные, неосторожные или неквалифицированные действия	$PhL:In/LogL:L_1/C:/A:St/M:N$	низкий
Администратор АСУ КВО	Любопытство или желание самореализации	$PhL:In/LogL:L_2/C:H/A:St/M:M$	высокий

необходимо учитывать ее структуру, включая наличие уровней (сегментов), решаемые задачи, состав, физические, логические, функциональные и технологические взаимосвязи, взаимодействие с другими автоматизированными (информационными) системами и информационно-телекоммуникационными сетями, режимы функционирования АСУ КВО, а также критичность ресурсов (информации и компонентов АСУ КВО), подлежащих защите.

В АСУ КВО объектами защиты (см. <http://www.fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/110-prikazy/864-prikaz-fstek-rossii-ot-14-marta-2014-g-n-31>), как правило, являются следующие:

- информация (данные) о параметрах (состоянии) управляемого (контролируемого) объекта или процесса: входная (выходная), управляющая (командная), контрольно-измерительная, иная критически важная (технологическая) информация;
- программно-технический комплекс, включающий технические средства (в том числе автоматизированные рабочие места, промышленные серверы, телекоммуникационное оборудование, каналы связи, программируемые логические контроллеры, исполнительные устройства), ПО

(в том числе микропрограммное, общесистемное, прикладное), а также средства защиты информации.

В [1] показано, что критически важными в АСУ КВО являются закреплённая в документации технологическая информация, воздействия на которую могут привести к нарушению функционирования автоматизированной системы, а также сведения непосредственно о самой АСУ КВО, которые (в случае их хищения) могут быть использованы для деструктивных воздействий на автоматизированную систему. При этом под технологической понимается

- оперативная (динамическая) информация (телеметрия, телеизмерения, телеуправление) о протекании управляемого технологического (производственного) процесса;
- архивная (статическая) информация (нормативно-техническая, проектная и эксплуатационная документация, параметры технологических (производственных) процессов и другая архивная информация).

Критичность ресурсов АСУ КВО определяется степенью возможного ущерба или степенью негативных последствий от воздействия на каждый из ресурсов АСУ КВО:

- высокая: воздействие на ресурс АСУ КВО повлекло нарушение штатного режима функционирования АСУ КВО, при этом возможны существенные негативные последствия (АСУ КВО не может выполнять функции контроля и управления технологическими (производственными) процессами);
- средняя: воздействие на ресурс АСУ КВО повлекло нарушение штатного режима функционирования АСУ КВО, при этом возможны умеренные негативные последствия (АСУ КВО не может выполнять хотя бы одну из функций контроля и управления технологическими (производственными) процессами);
- низкая: воздействие на ресурс АСУ КВО повлекло нарушение штатного режима функционирования АСУ КВО, при этом возможны незначительные негативные последствия (АСУ КВО может выполнять функции контроля и управления технологическими (производственными) процессами с потерей эффективности или выполнение функций возможно с привлечением дополнительных сил и средств);
- отсутствует: воздействие на ресурс АСУ КВО не повлекло никаких негативных последствий.

**Моделирование угроз безопасности информации в АСУ КВО.** Результатом процесса моделирования угроз безопасности информации в АСУ КВО является перечень актуальных для конкретной системы угроз в виде их формального описания. Угрозы безопасности информации определяются по результатам

- оценки возможностей нарушителя безопасности информации (в соответствии с моделью нарушителя безопасности информации в АСУ КВО);
- анализа возможных уязвимостей АСУ КВО (с использованием различных источников и БД об уязвимостях [5–7]);
- анализа возможных способов (сценариев) реализации угроз безопасности информации в АСУ КВО (с использованием банка данных угроз информационной безопасности [8]);
- анализа последствий реализации угроз безопасности информации.

Каждая угроза безопасности информации в АСУ КВО описывается следующим образом:

$T_{биi} = [\text{нарушитель, уязвимости, сценарий реализации угрозы, объекты воздействия, последствия реализации угрозы}]$ .

Наиболее простым и наглядным способом моделирования и описания возможных сценариев реализации угрозы безопасности информации является применение аппарата деревьев атак.

Деревья атак – достаточно мощный и гибкий инструмент для решения множества проблем, связанных с безопасностью, включая корреляцию и предотвращение воздействий на информацию, а также реакцию на них [9, 10].

В данном случае под атакой следует понимать действия, направленные на реализацию угроз несанкционированного доступа к информации, путем воздействия на нее или на ресурсы автоматизированной системы с применением программных и (или) программно-технических средств (ГОСТ Р 50.1.053-2005).

Деревья атак в связи с их графовой структурой наглядно представляют варианты информационно-технического воздействия на автоматизированную систему и по аналогии с деревом неисправностей (по ГОСТ Р 51901.13-2005) являются организованным графическим представлением условий или других факторов, вызывающих нежелательное событие, называемое вершиной событий.

По каждому возможному для конкретной АСУ КВО угрозе и нарушителю строится свое дерево атак. Вершиной этого дерева является реализация угрозы безопасности информации. Для достижения вершины дерева нарушителю безопасности информации необходимо выполнить некоторое количество условий на различных уровнях АСУ КВО, которые будут являться промежуточными вершинами дерева атак. На рисунке 3 показан пример двухуровневого дерева атак для угрозы подмены информации от датчиков полевого уровня (исходя из многоуровневой структуры АСУ КВО), из которого видно, что угроза будет реализована, если нарушителю удастся выполнить хотя бы одно из условий второго уровня («ИЛИ-декомпозиция» угрозы безопасности информации [9]).

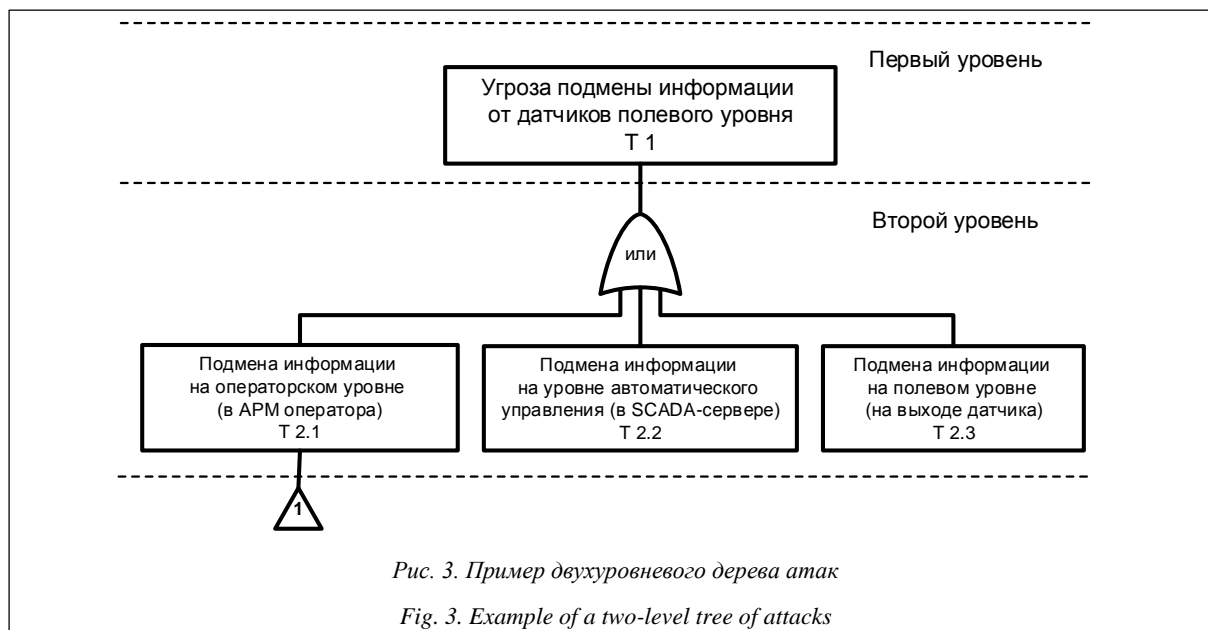
Аналогично может быть изображена «И-декомпозиция», обеспечивающая возможность достижения цели при одновременном выполнении всех условий на низшем уровне дерева, и «ИЛИ-НЕ-декомпозиция», при которой достижение цели возможно при выполнении только одного из условий на низшем уровне.

Количество уровней при построении дерева атак определяется исходя из требуемого уровня детализации сценария реализации угрозы безопасности в АСУ КВО. На рисунке 4 показан пример детализации на более низком уровне угрозы подмены информации от датчиков полевого уровня при реализации сценария подмены информации на операторском уровне АСУ КВО (в АРМ оператора).

Сценариями реализации угрозы безопасности являются все возможные пути от элементов самого нижнего уровня до конечной цели на первом уровне в построенном дереве атак.

Например, для дерева атак, изображенного на рисунках 3 и 4, одним из сценариев реализации угрозы подмены информации от датчиков полевого уровня будет такой путь:  $T\ 3.1 \rightarrow T\ 2.1 \rightarrow T\ 1$ .

Основной проблемой построения деревьев атак является масштабируемость [11], то есть возможность построения деревьев атак для информационных или автоматизированных систем, имеющих сложную структуру с большим количеством ком-



понентов и связей между ними с высокой степенью детализации (то есть на самом нижнем уровне дерева атаки будут отражены условия, не подлежащие дальнейшей детализации).

Для снижения сложности анализа и времени построения деревьев атак для АСУ КВО, имеющих сложную структуру и большое количество компонентов, детализацию на более низком уровне необходимо производить только для актуальных (применительно к данной АСУ КВО) сценариев (путей) реализации угрозы.

Степень актуальности сценария (пути) реализации угрозы зависит от вероятности того, что нарушитель будет использовать именно этот путь, и от степени возможного ущерба, вызванного последствиями реализации угрозы.

Сценарий (W) реализации угрозы является актуальным, если для автоматизированной системы (с заданными структурой, связями между компонентами и особенностями функционирования) существует ненулевая вероятность (P) реализации угрозы рассматриваемым сценарием нарушителем с соответствующим потенциалом, а реализация самой угрозы приведет к неприемлемым негативным последствиям (ущербу) (X):  $W_{ij} = [P_{ij}, X_i]$ , где  $i$  – номер рассматриваемой угрозы безопасности информации в АСУ КВО;  $j$  – номер возможного сценария (пути) реализации рассматриваемой угрозы безопасности информации в АСУ КВО.

Вероятность реализации  $i$ -й угрозы  $j$ -м сценарием ( $P_{ij}$ ) определяется исходя из базового потенциала нарушителя, его уровня физического доступа к компонентам АСУ КВО и необходимостью физического доступа к компонентам АСУ КВО при реализации угрозы рассматриваемым путем, а также наличия и уровня опасности уязвимостей, способствующих реализации угрозы рассматриваемым сценарием (табл. 4, 5). Уровень опасности

уязвимостей оценивается исходя из методики оценки CVSS [12].

Оценка актуальности  $j$ -го сценария (пути) при реализации  $i$ -й угрозы безопасности информации определяется исходя из вероятности реализации  $i$ -й угрозы  $j$ -м сценарием (путем) и степени возможного ущерба ( $X_i$ ) от воздействий этой угрозы на ресурсы АСУ КВО (табл. 6).

Если по результатам оценки какой-либо сценарий реализации угрозы оказывается неактуальным, его дальнейшая детализация на низших уровнях не производится. После оценки актуальности всех возможных сценариев реализации уязвимости вплоть до самого низшего уровня детализации производится оценка актуальности (для рассматриваемой АСУ КВО) угрозы в целом. Если хотя бы один из всех возможных сценариев реализации уязвимо-



Рис. 4. Пример детализации угрозы подмены информации от датчиков полевого уровня на один уровень ниже

Fig. 4. A detailed example of information substitution threat from field level sensors at one level lower

Таблица 4

**Предварительная оценка вероятности реализации угрозы безопасности информации в АСУ КВО  
(без учета наличия и уровня опасности уязвимостей)**

Table 4

**Preliminary assessment of information security threat probability in an automated control system for crucial objects (regardless of existence and a hazard level of vulnerabilities)**

Базовый потенциал нарушителя ( $P_{BASE}$ )	Уровень физического доступа нарушителя ( $LogL$ )	Необходимость физического доступа к компонентам АСУ КВО при реализации $i$ -й угрозы $j$ -м сценарием (путем)	
		да	нет
высокий	<i>Out</i>	низкая	высокая
высокий	<i>In</i>	высокая	высокая
средний	<i>Out</i>	отсутствует	средняя
средний	<i>In</i>	средняя	средняя
низкий	<i>Out</i>	отсутствует	низкая
низкий	<i>In</i>	низкая	низкая
отсутствует	<i>Out</i>	отсутствует	отсутствует
отсутствует	<i>In</i>	отсутствует	отсутствует

Таблица 5

**Оценка вероятности реализации угрозы безопасности информации в АСУ КВО  
с учетом наличия уязвимостей и уровня их опасности**

Table 5

**Assessment of information security threat probability in an automated control system for crucial objects taking into account existence and a hazard level of vulnerabilities**

Вероятность реализации угрозы безопасности информации без учета уровня опасности уязвимости	Наличие и уровень опасности уязвимостей (по CVSS)				
	отсутствует	низкий	средний	высокий	критический
отсутствует	отсутствует	отсутствует	отсутствует	отсутствует	отсутствует
низкая	отсутствует	отсутствует	низкая	средняя	средняя
средняя	низкая	низкая	средняя	высокая	высокая
высокая	средняя	средняя	высокая	высокая	высокая

Таблица 6

**Оценка актуальности сценария (пути) реализации угрозы безопасности информации в АСУ КВО**

Table 6

**Relevance assessment of information security threat scenarios (paths) in an automated control system for crucial objects**

Вероятность реализации $i$ -й угрозы $j$ -м сценарием ( $P_{ij}$ )	Степень возможного ущерба ( $X_i$ ) от реализации $i$ -й угрозы безопасности информации в АСУ КВО (в соответствии с таблицей 4)			
	отсутствует	низкая	средняя	высокая
отсутствует	неактуальный	неактуальный	неактуальный	неактуальный
низкая	неактуальный	неактуальный	неактуальный	актуальный
средняя	неактуальный	неактуальный	актуальный	актуальный
высокая	неактуальный	актуальный	актуальный	актуальный

сти является актуальным, эта угроза считается актуальной для рассматриваемой АСУ КВО и включается в результирующий перечень актуальных угроз.

Таким образом, общая последовательность моделирования угроз представляется в следующем виде:

- моделирование нарушителя безопасности информации в рассматриваемой АСУ КВО;
- построение множества всех возможных угроз для рассматриваемой АСУ КВО;
- построение дерева атак для каждой угрозы с выявлением актуальных сценариев реализации

угрозы для нарушителей с различным базовым потенциалом;

- формирование множества (перечня) актуальных угроз для рассматриваемой АСУ КВО.

Предлагаемая методика моделирования угроз для АСУ КВО позволит, во-первых, за счет исключения из анализа неактуальных сценариев реализации угрозы сократить время построения перечня актуальных (для рассматриваемой АСУ КВО) угроз, а во-вторых, учесть возможность реализации одной и той же угрозы в АСУ КВО несколькими сценариями нарушителями с разным базовым потенциалом.

Методика может использоваться для определения требований к системе защиты информации АСУ КВО или формирования комплекса технических и организационных мероприятий по обеспечению безопасности информации в АСУ КВО.

### Литература

1. Горбачев И.Е., Глухов А.П. Моделирование процессов нарушения информационной безопасности критической инфраструктуры // Тр. СПИИ РАН. 2015. Вып. 1 (38). С. 112–135.
2. Жуков В.Г., Жукова М.Н., Стефаров А.П. Модель нарушителя прав доступа в автоматизированной системе // Программные продукты и системы. 2012. № 2 (98). С. 75–78.
3. Гришина Н.В. Модель потенциального нарушителя объекта информатизации // Изв. ЮФУ. Технич. науки. 2003. № 4. Т. 33. С. 356–358.
4. Климов С.М. Методы и модели противодействия компьютерным атакам. Люберцы: КАТАЛИТ, 2008. 316 с.
5. Банк данных угроз информационной безопасности.

Список уязвимостей. URL: <http://www.bdu.fstec.ru/vul> (дата обращения: 25.02.2016).

6. National Vulnerability Database. URL: <http://www.nvd.nist.gov> (дата обращения: 25.02.2016).

7. Open Source Vulnerability Database. URL: <http://www.osvdb.org> (дата обращения: 25.02.2016).

8. Банк данных угроз информационной безопасности. Список угроз. URL: <http://www.bdu.fstec.ru/threat> (дата обращения: 25.02.2016).

9. Полаженко С. Деревья атак и их применение при анализе проблемы безопасности и защищенности программных продуктов. URL: <http://www.software-testing.ru/library/testing/security/140-attack-trees> (дата обращения: 25.02.2016).

10. Чечулин А.А. Методика оперативного построения, модификации и анализа деревьев атак // Тр. СПИИ РАН. 2013. Вып. 3 (26). С. 40–53.

11. Колегов Д.Н. Проблемы синтеза и анализа графов атак. URL: <http://www.securitylab.ru/contest/299868.php> (дата обращения: 25.02.2016).

12. Калькулятор CVSS. URL: <http://www.bdu.fstec.ru/calc> (дата обращения: 25.02.2016).

DOI: 10.15827/0236-235X.115.042-050

Received 16.03.16

## MODELING INFORMATION SECURITY THREATS IN THE AUTOMATED CONTROL SYSTEM FOR CRUCIAL OBJECTS ON THE BASIS OF ATTACK SCENARIOS

<sup>1</sup>*Drobotun E.B., Ph.D. (Engineering), Doctoral Student, drobotun@xakep.ru*

<sup>1</sup>*Tsvetkov O.V., Chief of service of protection state secrets*

<sup>1</sup>*Military Academy of the Aerospace Defense, Zhigareva St. 50, Tver, 170100, Russian Federation*

**Abstract.** Development and construction of information security threat models in various automated and data-processing systems is one of key steps in ensuring information security in these systems and their safe and proper operation. Based on the information security threat model we can define the requirements to information security in automated data-processing systems. We can also define a set of organizational, technical, organizational and technical measures to ensure information security in automated and computer systems for various purposes.

The article proposes a method of constructing a threat model for automated control systems of crucial and potentially dangerous objects. The method is based on modeling information security violator's possible action scenarios. The proposed method takes into account a multilevel structure modern automated control systems for crucial objects and the possibility of implementing the same information security threats using different scenarios by different violators.

The proposed methodology would reduce the time for building a model of information security threats in automated control systems for crucial and potentially dangerous objects.

**Keywords:** critical information system, information security threat, security information violator, threat model, violator model.

### References

1. Gorbachev I.E., Glukhov A.P. Modeling of processes of information security violations of critical infrastructure. *Trudy SPIIRAN* [SPIIRAN Proc.]. 2015, no. 1 (38), pp. 112–135 (in Russ.).
2. Zhukov V.G., Zhukova M.N., Stefarov A.P. Model of an access rights violator in an automated system. *Programmnye produkty i sistemy* [Software & Systems]. 2012, no. 2 (98), pp. 75–78 (in Russ.).
3. Grishina N.V. A model of a potential violator of informatization object. *Izvestiya YUFU. Tekhnicheskie nauki* [News of SFedU. Technical Sciences]. 2003, no. 4, vol. 33, pp. 356–358 (in Russ.).
4. Klimov S.M. *Metody i modeli protivodeystviya kompyuternym atakam* [Methods and Models of Counteractions to Computer Attacks]. Lubertsy, KATALIT Publ., 2008, 316 p.
5. *Bank dannykh ugroz informatsionnoy bezopasnosti. Spisok uязvimostey* [information Security Threats Databank. List of Vulnerabilities]. FSTEC Russia. Available at: <http://www.bdu.fstec.ru/vul> (accessed February 25, 2016).
6. *National vulnerability database*. Available at: <http://www.nvd.nist.gov> (accessed February 25, 2016).
7. *Open source vulnerability database*. Available at: <http://www.osvdb.org> (accessed February 25, 2016).
8. *Bank dannykh ugroz informatsionnoy bezopasnosti. Spisok ugroz* [Information Security Threats Databank. List of Threats]. FSTEC Russia. Available at: <http://www.bdu.fstec.ru/threat> (accessed February 25, 2016).
9. Polazhenko S. *Derevya atak i ikh primeneniye pri analize problemy bezopasnosti i zashchishchennosti programmnykh produktov* [Trees of Attacks and Their Application When Analyzing a Problem of Safety and Security of Software Products]. Available at: <http://www.software-testing.ru/library/testing/security/140-attack-trees> (accessed February 25, 2016).
10. Chechulin A.A. Methods of operational formation, modification and analysis of attack trees. *Trudy SPIIRAN* [SPIIRAN Proc.]. 2013, no. 3 (26), pp. 40–53 (in Russ.).
11. Kolegov D.N. *Problemy sinteza i analiza grafov atak* [Problems of Synthesis and Analysis of Attack Graphs]. Available at: <http://www.securitylab.ru/contest/299868.php> (accessed February 25, 2016).
12. *Kalkulyator CVSS* [CVSS Calculator]. FSTEC Russia. Available at: <http://www.bdu.fstec.ru/calc> (accessed February 25, 2016).

УДК 004.056

Дата подачи статьи: 04.05.16

DOI: 10.15827/0236-235X.115.051-059

## **СИНТЕЗ СИСТЕМ ЗАЩИТЫ АВТОМАТИЗИРОВАННЫХ СИСТЕМ УПРАВЛЕНИЯ ОТ РАЗРУШАЮЩИХ ПРОГРАММНЫХ ВОЗДЕЙСТВИЙ**

*Е.Б. Дроботун, к.т.н., докторант, drobotun@xakep.ru*

*(Военная академия воздушно-космической обороны им. Маршала Советского Союза  
Г.К. Жукова, ул. Жигарева, 50, г. Тверь, 170022, Россия)*

Информатизация и автоматизация различных объектов производства, в том числе создающих потенциальную угрозу обеспечению безопасности жизнедеятельности, сделала обеспечение безопасности функционирования таких объектов одной из наиболее острых проблем современности. Процессы информатизации и автоматизации обуславливают появление новых видов угроз информационной безопасности, направленных на нарушение функционирования производственных объектов, которые могут быть подвергнуты деструктивным информационно-техническим воздействиям, и создающих потенциальную угрозу обеспечению безопасности жизнедеятельности. Данные факторы в сочетании с возрастающим уровнем зависимости общества от устойчивости функционирования объектов, создающих потенциальную угрозу обеспечению безопасности жизнедеятельности, требуют принятия комплекса мер (в том числе и технического характера), направленных на снижение риска нарушения или полного прекращения функционирования таких объектов в условиях информационно-технических воздействий.

В статье представлен один из возможных подходов к построению рациональной (с точки зрения обеспечения требуемого уровня защищенности и минимального влияния системы защиты на защищаемую автоматизированную систему) системы защиты от разрушающих программных воздействий как основного вида информационно-технических воздействий, которым могут быть подвергнуты автоматизированные системы.

**Ключевые слова:** *автоматизированная система управления технологическими процессами, потенциально опасный производственный объект, информационная безопасность, разрушающее программное воздействие, защита от разрушающих программных воздействий.*

Внедрение АСУ технологическими и производственными процессами на объектах промышленности обусловило то, что обеспечение безопасности их функционирования стало одной из наиболее острых проблем современности. В настоящее время необходимым условием эффективного и безопасного функционирования таких объектов является обеспечение защищенности АСУ этими объектами от различного рода информационно-технических воздействий, в том числе с помощью вредоносных программ и *разрушающих программных воздействий* (РПВ).

Наиболее известным случаем информационно-технического воздействия на потенциально опасный объект может служить информационно-техническая атака, проведенная, предположительно, спецслужбами США и Израиля на объекты Иранской атомной инфраструктуры с помощью вредоносной программы Stuxnet. Используя уязвимости операционной системы, эта вредоносная программа, по некоторым данным, вывела из строя 1 368 из 5 000 центрифуг на заводе по обогащению урана в Натанзе, а также сорвала сроки запуска ядерной АЭС в Бушере [1, 2].

АСУ технологическими процессами (ТП) имеют более высокие степени риска по сравнению с системами, предназначенными только для хранения, обработки и предоставления информации, вплоть до рисков нарушения ТП и работоспособности системы, выброса вредных веществ, техногенных катастроф и человеческих жертв.

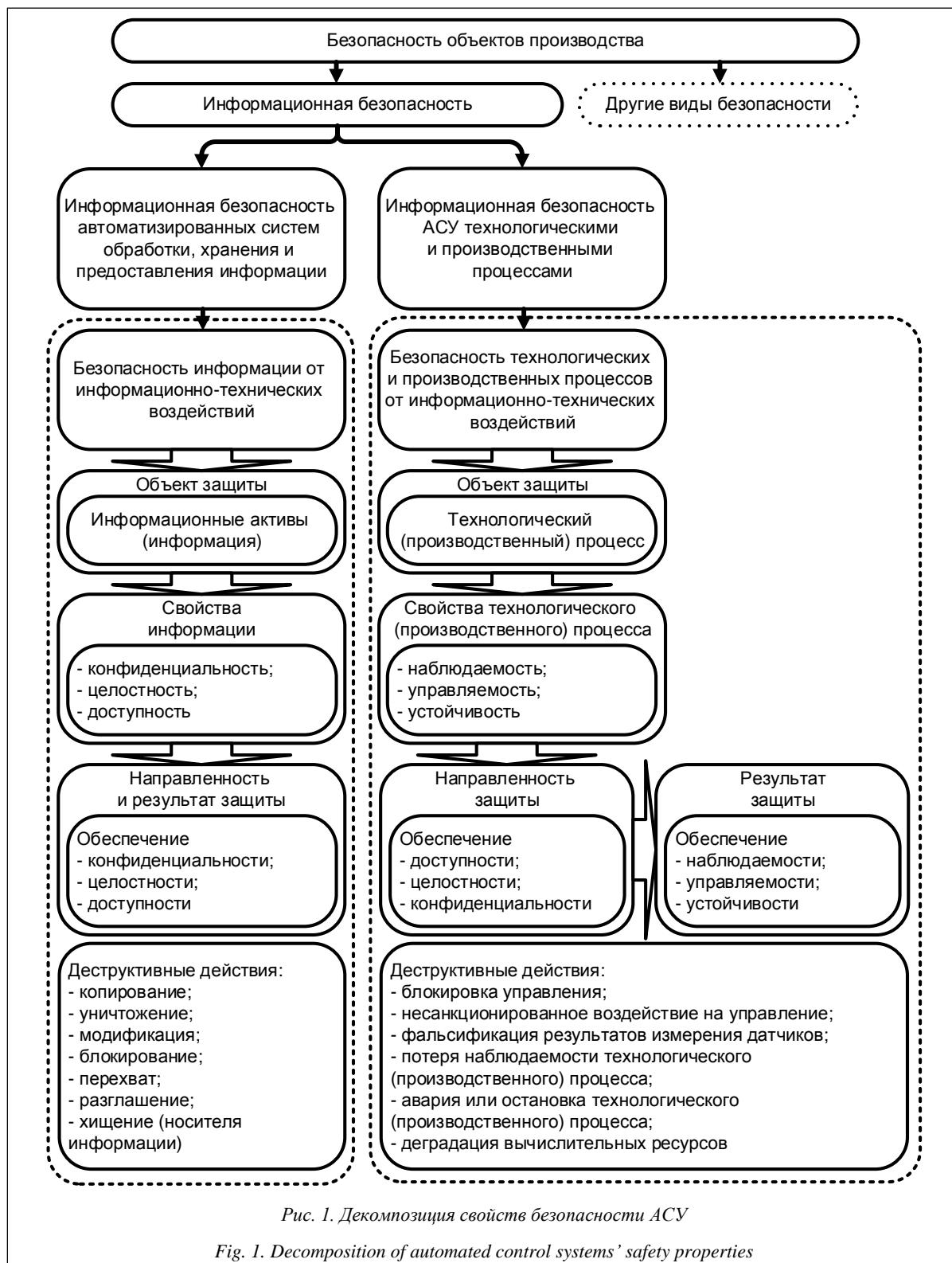
В связи с этим при обеспечении защиты АСУ технологическими и производственными процессами от информационно-технических воздействий,

помимо конфиденциальности, целостности и доступности информации (как указано в ГОСТ Р 50922–2006), прежде всего необходимо обеспечить безопасность самого технологического или производственного процесса.

Целью защиты технологических или производственных процессов, а также систем управления ими является обеспечение требований управляемости, наблюдаемости и устойчивости этих процессов и систем. При невыполнении этих требований возможны деструктивные действия на АСУ ТП, связанные с потерями управляемости технологического или производственного процесса, включающими блокировку управления и (или) несанкционированное управление; наблюдаемости технологического или производственного процесса (модификация параметров процессов и (или) фальсификация измерений датчиков); работоспособности системы (авария или остановка технологического или производственного процесса и (или) деградация вычислительных ресурсов) (рис. 1) [3–5].

Все деструктивные воздействия на АСУ ТП являются производными трех причин: нарушение доступности (отказ в обслуживании) критически важной информации, нарушение ее целостности (модификация) и нарушение конфиденциальности (утечка).

Для технологических или производственных процессов основным направлением защиты является обеспечение доступности и целостности информации, поскольку нарушения конфиденциальности (утечки) информации непосредственно угрозу безопасности этим процессам не создают. Однако при этом следует учитывать, что утечка так



называемой технологической информации об АСУ (о составе, характеристиках управляемого процесса, программного и программно-аппаратного обеспечения, о размещении, коммуникациях и т.п.) может в случае ее хищения использоваться для деструктивных информационно-технических воздействий [5].

Одним из основных способов деструктивных информационно-технических воздействий на автоматизированные системы являются воздействия с помощью различных программ, в том числе и специально созданных для таких воздействий (вредоносных программ). Такой вид воздействий относится к РПВ на автоматизированную систему.



К программам, с помощью которых осуществляются РПВ, следует отнести программы, способные выполнять любое непустое подмножество из множества следующих функций [6]:

- скрывать признаки своего присутствия в системе;
- обладать способностью к самодублированию, в том числе к созданию своих модифицированных копий;
- обладать способностью к ассоциированию себя с другими программами;
- обладать способностью к переносу своих фрагментов в иные области оперативной или внешней памяти, в том числе находящиеся на удаленном компьютере;
- получать несанкционированный доступ к компонентам или ресурсам системы;
- разрушать или искажать код программ в оперативной памяти;
- наблюдать за процессами обработки информации и принципами функционирования средств защиты;
- сохранять фрагменты информации из оперативной памяти в некоторой области внешней памяти;
- искажать, блокировать или подменять выводимый во внешнюю память или канал связи информационный массив, образовавшийся в результате работы прикладных программ;
- искажать находящиеся во внешней памяти массивы данных;
- подавлять информационный обмен в компьютерных сетях, фальсифицировать информацию в каналах связи;
- нейтрализовывать работу тестовых программ и средств защиты информационных ресурсов системы;
- постоянно или кратковременно изменять степень защищенности конфиденциальных данных;
- приводить в неработоспособное состояние или разрушать компоненты системы;
- создавать скрытые каналы передачи данных;
- инициировать ранее внедренные РПВ.

При этом стоит отметить, что РПВ возможны с помощью не только специально созданных для этого вредоносных программ [7], но и многих других программ, изначально не создававшихся с целью этих воздействий (например, с помощью утилит анализа сетевого трафика возможны перехват и искажение информации, циркулирующей в информационно-вычислительной сети, с помощью различных системных утилит возможно изменение конфигурации и настроек операционной системы с целью нарушения корректности ее функционирования и т.п.).

Согласно [8], технические меры защиты информации (с точки зрения противодействия РПВ), реа-

лизуемые в АСУ в рамках ее системы защиты, должны обеспечивать (рис. 2):

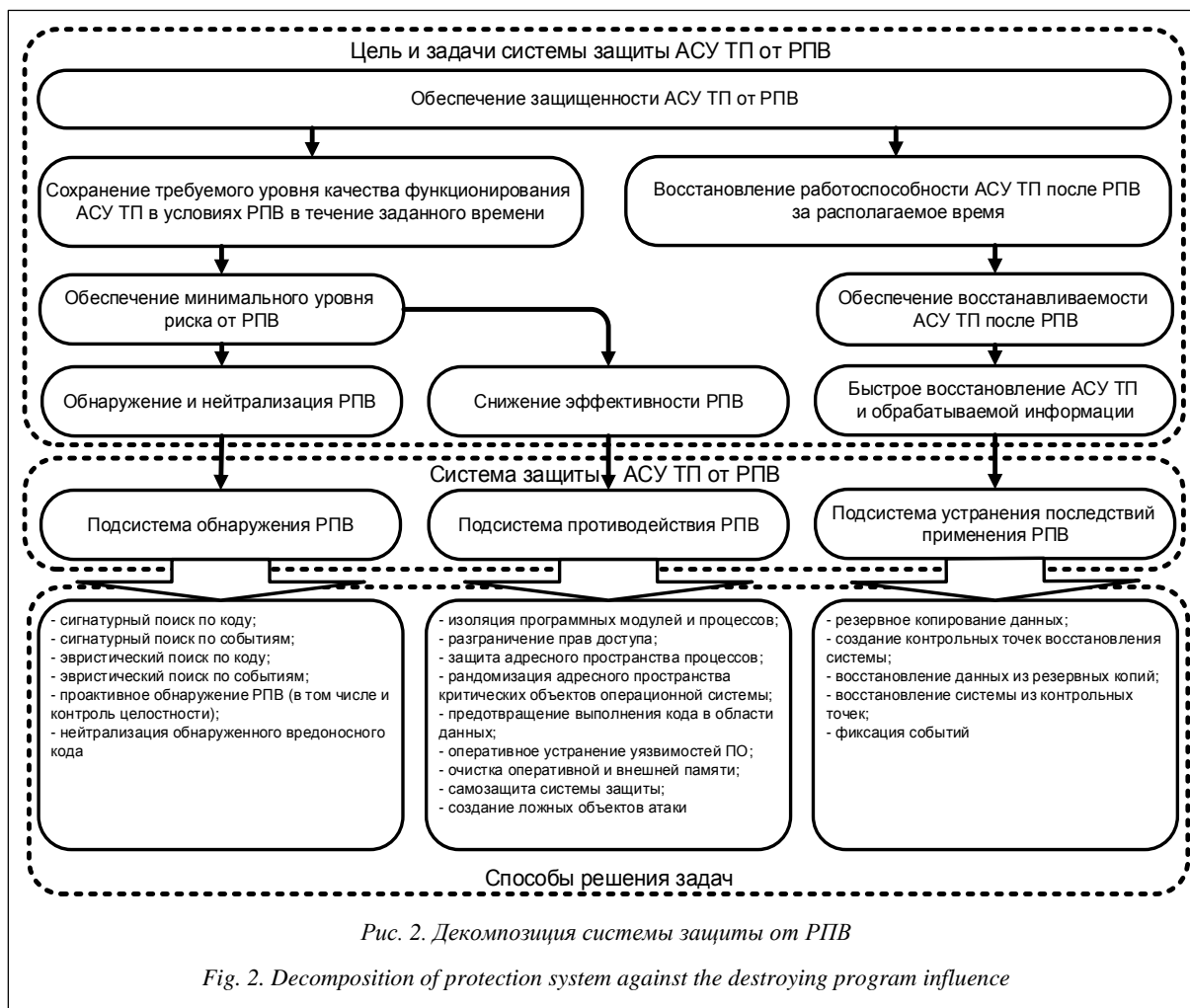
- идентификацию и аутентификацию субъектов и объектов доступа;
- управление доступом субъектов доступа к объектам доступа;
- ограничение программной среды;
- защиту машинных носителей информации;
- регистрацию событий безопасности;
- антивирусную защиту;
- обнаружение (предотвращение) вторжений;
- контроль (анализ) защищенности информации;
- целостность АСУ и информации;
- доступность технических средств и информации.

Помимо этого, реализация технических мер в рамках системы защиты от РПВ должна обеспечивать доступность обрабатываемой в АСУ информации (исключение неправомерного блокирования информации), ее целостность (исключение неправомерного уничтожения, модифицирования информации), а также при необходимости конфиденциальность (исключение неправомерного доступа, копирования, предоставления или распространения информации); должна соотноситься с мерами по промышленной, физической, пожарной, экологической, радиационной безопасности, иными мерами по обеспечению безопасности АСУ и управляемого (контролируемого) объекта и (или) процесса и не должна отрицательно влиять на штатный режим функционирования АСУ [8].

На основании перечисленных требований задачу синтеза системы защиты от РПВ можно представить как задачу выбора таких рациональных структуры и параметров системы защиты  $A^*$  от РПВ для АСУ технологическими и производственными процессами, чтобы система защиты, отвечающая выбранным структуре и параметрам, позволяла обеспечить минимальное снижение качества функционирования  $K$  этой автоматизированной системы при сохранении требуемого уровня защиты от РПВ  $Q$ , при ограничениях на ресурсы  $R$  и стоимость создания системы защиты  $C$ :

$$A^* = \arg \min_{A_i \in V^{**}} K(A_i) \quad \text{при} \quad V^{**} = \{A_i \mid (Q(A_i) \geq Q_{\text{ТРЕБ}}) \wedge (C(A_i) \leq C_{\text{ДОП}}) \wedge (R(A_i) \leq R_{\text{ДОП}})\}, \quad V^{**} \subseteq V,$$

где  $K$  – множество показателей, характеризующих степень снижения качества функционирования АСУ при вводе в ее состав системы защиты от РПВ;  $A_i$  – потенциально возможный вариант системы защиты;  $V^{**}$  – множество вариантов реализации системы защиты, удовлетворяющих ограничениям по стоимости и потребляемым ресурсам, а также требованиям по уровню защиты;  $V$  – множество всех возможных вариантов реализации системы защиты от РПВ;  $Q$  – эффективность системы



защиты;  $C$  – стоимость построения системы защиты;  $R$  – потребность в ресурсах;  $Q_{\text{треб}}$  – требуемый уровень защиты от РПВ;  $C_{\text{доп}}$  – максимально допустимая стоимость построения системы защиты;  $R_{\text{доп}}$  – максимально допустимая потребность в ресурсах.

Исходя из того, что, по ГОСТ 34.003–90, АСУ представляет собой систему, состоящую из персонала и комплекса средств автоматизации его деятельности, реализующей технологию выполнения установленных функций, степень снижения качества функционирования АСУ при вводе в ее состав системы защиты от РПВ  $K$  можно представить двумя следующими совокупностями показателей:  $K_C$  – множество показателей снижения производительности функционирования комплекса средств автоматизации, входящего в состав АСУ;  $K_{\text{П}}$  – множество показателей снижения производительности обслуживающего персонала.

Используя широко применяемые показатели производительности информационных систем [9, 10], снижение производительности комплекса средств автоматизации можно оценить совокупностью показателей:  $K_C = \{K_{\text{ПС}}; K_{\text{ПАР}}; K_{\text{РЕС}}; K_{\text{ТОТКЛ}}; K_{\text{ОШ}}\}$ , где  $K_{\text{ПС}}$  – коэффициент снижения пропуск-

ной способности (количества операций, выполняемых системой за определенный период времени);  $K_{\text{ПАР}}$  – коэффициент снижения параллелизма (количества операций, выполняемых системой одновременно);  $K_{\text{РЕС}}$  – коэффициент увеличения запаса ресурса (количества ресурсов, необходимых для обеспечения роста нагрузки);  $K_{\text{ТОТКЛ}}$  – коэффициент увеличения времени отклика (времени выполнения одной операции);  $K_{\text{ОШ}}$  – коэффициент увеличения частоты ошибок (частоты генерации системой исключений типа «отказ в обслуживании»).

Снижение производительности обслуживающего персонала, в свою очередь, можно оценить двумя показателями:  $K_{\text{П}} = \{K_{\text{Твп}}; K_{\text{Д}}\}$ , где  $K_{\text{Твп}}$  – коэффициент увеличения времени выполнения операций обслуживающим персоналом при выполнении им функциональных обязанностей;  $K_{\text{Д}}$  – коэффициент повышения дискомфорта при выполнении функциональных обязанностей.

Максимально допустимая потребность в ресурсах может быть представлена как  $R_{\text{доп}} = \{R_{\text{Вдоп}}; R_{\text{Лдоп}}\}$ , где  $R_{\text{Вдоп}}$  – максимально допустимая потребность в программных и аппаратных (вычислительных) ресурсах;  $R_{\text{Лдоп}}$  – максимально допустимая потребность в людских ресурсах.

Варьируемыми параметрами в потенциально возможных вариантах построения системы защиты  $A_i$  от РПВ являются следующие:

- структура системы защиты, которую можно описать как  $S = \{L; B\}$ , где  $L$  – множество составных элементов системы защиты;  $B$  – функциональные связи между ними;
- множество параметров системы защиты  $F$ .

Изменение структуры и параметров системы защиты от РПВ ведет к изменению всех вышеописанных показателей.

Поскольку при построении системы защиты от РПВ возникает необходимость учета большого количества ограничений и частных показателей, задача построения системы защиты от РПВ с требуемыми характеристиками не может быть решена методом прямого синтеза (синтез оптимальной структуры с заданными характеристиками).

Исходя из этого, решение данной задачи заключается в выборе наилучших вариантов при проектировании и построении системы защиты в условиях ограничений. При этом с целью уменьшения количества вариантов, упрощения их формирования и проведения расчетов показателей, характеризующих эти варианты, необходима декомпозиция задачи выбора рационального варианта системы защиты.

В данном случае декомпозицию можно осуществить как по последовательности выполнения процедур формирования и выбора вариантов, так и по подсистемам системы защиты АСУ от РПВ (функциональная декомпозиция).

Декомпозиция по подсистемам основывается на том, что систему защиты от РПВ можно представить в виде множества подсистем, реализующих определенный перечень функций. На основании этого в качестве подсистем, обеспечивающих защищенность автоматизированной системы от РПВ, в составе системы защиты можно выделить три подсистемы: обнаружения РПВ, противодействия РПВ и устранения последствий применения РПВ (рис. 2).

Соотношение технических мер защиты (согласно [8]) и способов решения задач подсистемами системы защиты от РПВ показано на рисунке 3.

Последовательность выполнения процедур формирования и выбора вариантов будет следующей:

- формирование альтернативных вариантов для каждой подсистемы;
- оценка сформированных вариантов и выбор из них рациональных;
- формирование на основе выбранных вариантов подсистем общего рационального варианта системы защиты АСУ от РПВ.

Исходя из проведенной декомпозиции системы защиты от РПВ по подсистемам, вариант построения системы защиты  $A$  можно представить в виде

множества  $A = \{A_{\text{обн}}; A_{\text{пр}}; A_{\text{устр}}\}$ , где  $A_{\text{обн}}$  – вариант реализации подсистемы обнаружения РПВ;  $A_{\text{пр}}$  – вариант реализации подсистемы противодействия РПВ;  $A_{\text{устр}}$  – вариант реализации подсистемы устранения последствий применения РПВ.

Соответственно, структура системы защиты может быть представлена как  $S = \{\{L_{\text{обн}}; B_{\text{обн}}\}; \{L_{\text{пр}}; B_{\text{пр}}\}; \{L_{\text{устр}}; B_{\text{устр}}\}; B_{\text{сз}}\}$ , где  $L_{\text{обн}}$  – множество элементов подсистемы обнаружения РПВ;  $B_{\text{обн}}$  – функциональные связи между элементами обнаружения РПВ;  $L_{\text{пр}}$  – множество элементов подсистемы;  $B_{\text{пр}}$  – функциональные связи между элементами противодействия РПВ;  $L_{\text{устр}}$  – множество элементов подсистемы устранения последствий применения РПВ;  $B_{\text{устр}}$  – функциональные связи между элементами подсистемы устранения последствий применения РПВ;  $B_{\text{сз}}$  – функциональные связи между подсистемами, входящими в состав системы защиты.

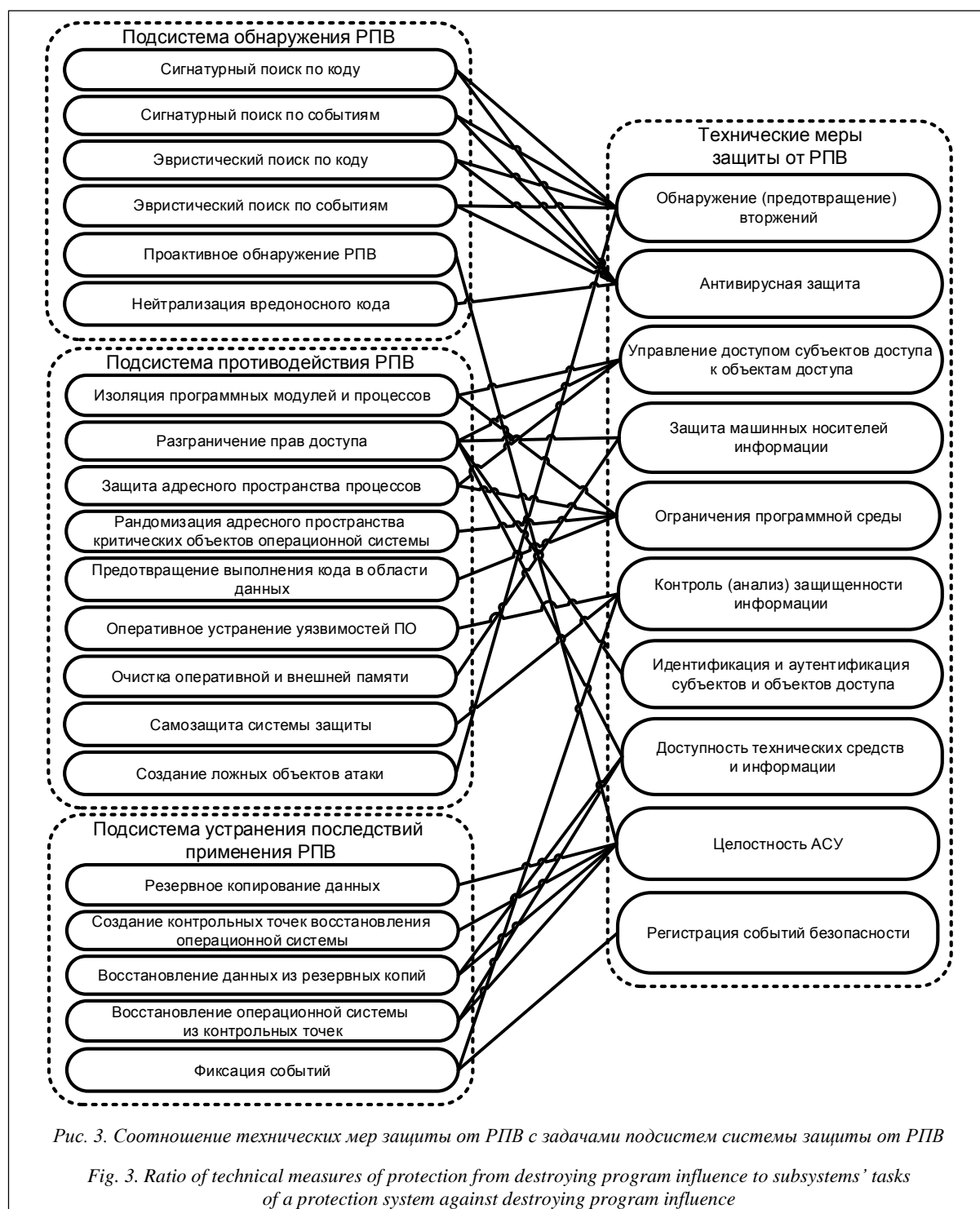
В свою очередь, параметры системы защиты могут быть представлены в виде совокупности трех множеств:  $F = \{F_{\text{обн}}; F_{\text{пр}}; F_{\text{устр}}\}$ , где  $F_{\text{обн}}$  – параметры подсистемы обнаружения РПВ;  $F_{\text{пр}}$  – параметры подсистемы противодействия РПВ;  $F_{\text{устр}}$  – параметры подсистемы устранения последствий применения РПВ.

Задачи выбора рациональных вариантов подсистем решаются последовательно, то есть сначала формируются все возможные варианты подсистем, далее из этих вариантов выбираются рациональные (с точки зрения соответствия требованиям по эффективности выполнения задач, возложенных на подсистему). Затем из этих рациональных вариантов построения подсистем формируется множество вариантов построения системы защиты, из которых сначала отбираются варианты, приемлемые по стоимости и по потребляемым ресурсам, далее из них отбираются варианты, удовлетворяющие требованиям по обеспечению защищенности автоматизированной системы. После чего из этих отобранных вариантов выбирается рациональный вариант построения системы защиты, оказывающий минимальное влияние на качество функционирования защищаемой автоматизированной системы.

Общая схема решения задачи синтеза системы защиты АСУ ТП от РПВ показана на рисунке 4.

Из рисунка видно, что исходными данными для построения системы защиты являются следующие:

- данные о структуре защищаемой автоматизированной системы  $Y = \{Y_{\text{п}}; Y_{\text{а}}; Y_{\text{с}}\}$ , где  $Y_{\text{п}}$  – составные компоненты ПО автоматизированной системы;  $Y_{\text{а}}$  – составные аппаратные компоненты АСУ;  $Y_{\text{с}}$  – функциональные связи между компонентами автоматизированной системы;
- максимально допустимая стоимость создания системы защиты  $C_{\text{доп}}$ ;
- максимально допустимые ресурсы, необходимые для функционирования системы  $R_{\text{доп}}$ ;



– данные об угрозах и уязвимостях: БД актуальных угроз, БД актуальных уязвимостей;

– данные о потенциальном нарушителе (профиль нарушителя)  $X = \{X_K; X_O; X_M; X_C\}$ , где  $X_K$  – уровень компетенции нарушителя;  $X_O$  – уровень оснащённости нарушителя;  $X_M$  – уровень мотивации нарушителя;  $X_C$  – класс нарушителя (внутренний или внешний).

Формирование множества всех возможных вариантов построения подсистем (блок 4) осуществ-

ляется методом морфологического анализа [11, 12], при этом

– для подсистемы обнаружения РПВ исходными данными является совокупность множеств  $\{L_{обн}; V_{обн}; F_{обн}\}$ , а формируемыми данными – множество вариантов построения подсистемы обнаружения РПВ  $V_{обн}$ ;

– для подсистемы противодействия РПВ исходными данными является совокупность множеств  $\{L_{пр}; B_{пр}; F_{пр}\}$ , а формируемыми данными –

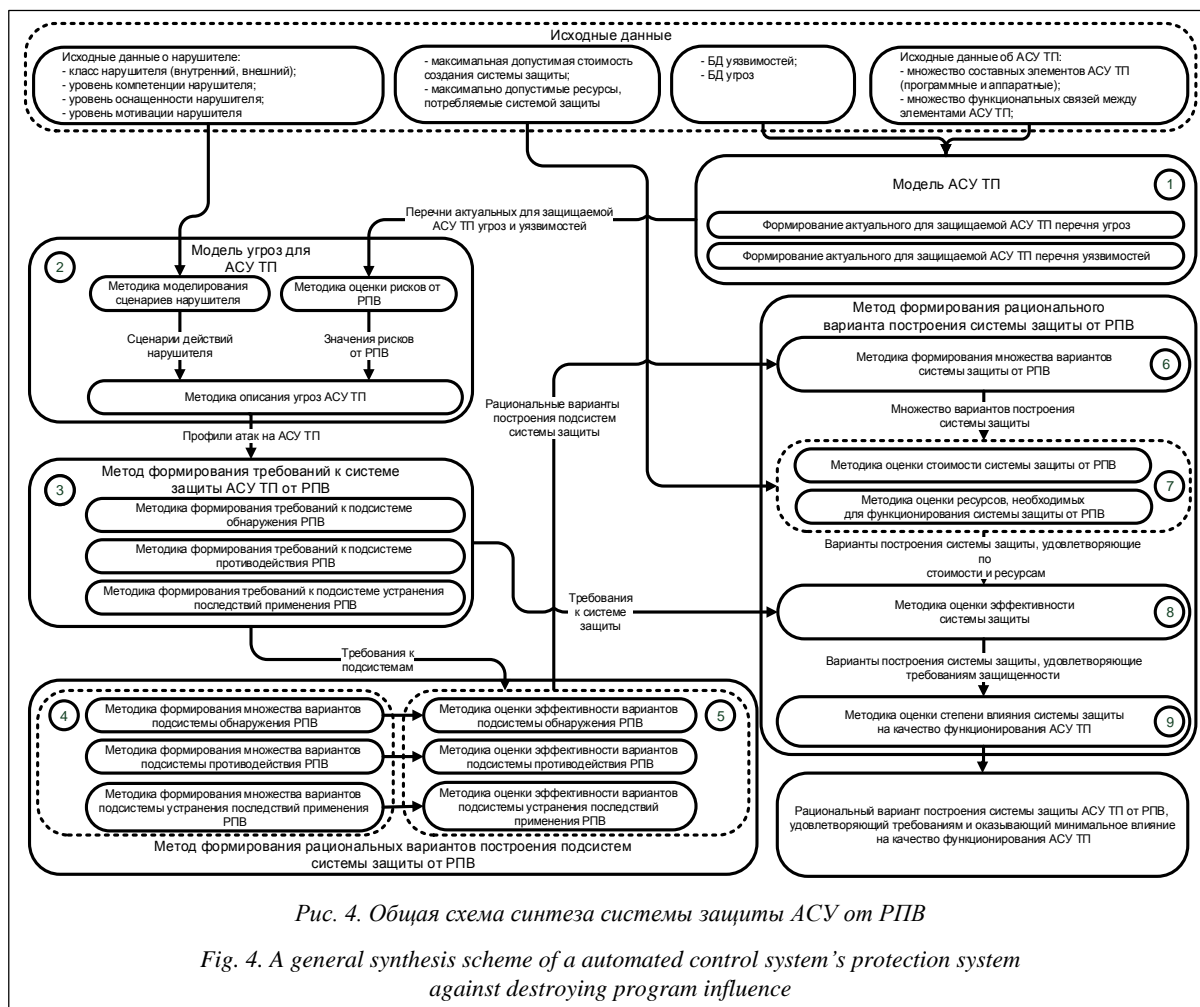


Рис. 4. Общая схема синтеза системы защиты АСУ от РПВ

Fig. 4. A general synthesis scheme of a automated control system's protection system against destroying program influence

множество вариантов построения подсистемы противодействия РПВ  $V_{\text{ПР}}$ ;

– для подсистемы устранения последствий применения РПВ исходными данными является совокупность множеств  $\{L_{\text{УСТР}}, V_{\text{УСТР}}, F_{\text{УСТР}}\}$ , а формируемыми данными – множество вариантов построения подсистемы противодействия РПВ  $V_{\text{УСТР}}$ .

Формирование множеств рациональных вариантов построения подсистем ( $V_{\text{ОБН}}^*$ ,  $V_{\text{ПР}}^*$ ,  $V_{\text{УСТР}}^*$ ) осуществляется отбором из полученных на первом этапе множеств ( $V_{\text{ОБН}}$ ,  $V_{\text{ПР}}$ ,  $V_{\text{УСТР}}$ ) (блок 5) по условиям

– соответствия эффективности функционирования подсистемы обнаружения РПВ требуемой эффективности:

$V_{\text{ОБН}}^* = \{A_{\text{ОБН}i} \mid Q_{\text{ОБН}}(A_{\text{ОБН}i}) \geq Q_{\text{ОБН}}^{\text{ТРЕБ}}\}$ , где эффективность  $Q_{\text{ОБН}}$  выражается вероятностью обнаружения РПВ и вероятностью ложной тревоги –  $Q_{\text{ОБН}} = \{P_{\text{ОБН}}; P_{\text{ЛТ ОБН}}\}$ ;

– соответствия эффективности функционирования подсистемы противодействия РПВ требуемой эффективности:

$V_{\text{ПР}}^* = \{A_{\text{ПР}i} \mid Q_{\text{ПР}}(A_{\text{ПР}i}) \geq Q_{\text{ПР}}^{\text{ТРЕБ}}\}$ , где эффективность  $Q_{\text{ПР}}$  выражается вероятностью блокирования

РПВ и вероятностью ложного блокирования –  $Q_{\text{ПР}} = \{P_{\text{БЛ}}; P_{\text{ЛТ БЛ}}\}$ ;

– соответствия эффективности функционирования подсистемы устранения последствий применения РПВ требуемой эффективности:

$V_{\text{УСТР}}^* = \{A_{\text{УСТР}i} \mid Q_{\text{УСТР}}(A_{\text{УСТР}i}) \geq Q_{\text{УСТР}}^{\text{ТРЕБ}}\}$ , где эффективность  $Q_{\text{УСТР}}$  выражается вероятностью восстановления системы за требуемое время и вероятностью восстановления данных за требуемое время –  $Q_{\text{УСТР}} = \{P_{\text{СИСТ}}; P_{\text{ДАН}}\}$ .

Для формирования требований по эффективности функционирования подсистем используются модель АСУ и модель угроз для нее. В модели АСУ на основе данных о структуре системы (блок 1)  $\{Y_{\text{П}}; Y_{\text{Л}}; Y_{\text{С}}\}$  формируются перечни актуальных для защищаемой автоматизированной системы угроз и уязвимостей. Исходные данные берутся из соответствующих БД [13, 14].

Модель угроз для АСУ (блок 2) формирует множество профилей возможных атак на АСУ с помощью РПВ. Профиль атаки описывается следующим образом:  $\{H_{\text{АТ}}, Z_{\text{АТ}}, P_{\text{АТ}}, M_{\text{АТ}}\}$ , где  $H_{\text{АТ}}$  – атакуемый элемент АСУ (цель атаки);  $Z_{\text{АТ}}$  – средство проведения атаки;  $P_{\text{АТ}}$  – вероятность успешного заверше-

ния атаки;  $M_{AT}$  – возможный ущерб от атаки на автоматизированную систему.

Исходными данными для модели угроз являются перечни актуальных для защищаемой автоматизированной системы угроз и уязвимостей, полученные в модели автоматизированной системы, а также данные о нарушителе (множество профилей нарушителя). Профили нарушителей, а также сценарии их действий формируются исходя из уровня опасности и критичности защищаемого объекта с помощью методов экспертных оценок [15].

На основе полученного множества профилей атак формируются требования к подсистемам (блок 3) в виде требуемых значений показателей эффективности функционирования подсистем:

- для подсистемы обнаружения РПВ – требуемые вероятность обнаружения и вероятность ложной тревоги:  $Q_{OBN\ TREB} = \{P_{OBN\ TREB}; P_{LT\ OBN\ TREB}\}$ ;

- для подсистемы противодействия РПВ – требуемые вероятность блокирования и вероятность ложного блокирования:  $Q_{PR\ TREB} = \{P_{BL\ TREB}; P_{LT\ BL\ TREB}\}$ ;

- для подсистемы устранения последствий применения РПВ – требуемые вероятность восстановления системы и вероятность восстановления данных:  $Q_{USTR\ TREB} = \{P_{CIST\ TREB}; P_{DAN\ TREB}\}$ .

Для выбора рационального варианта построения системы защиты  $A^*$  с использованием морфологического анализа формируется множество возможных вариантов построения системы защиты  $V$  (блок 6). Исходными данными при этом являются множества рациональных вариантов построения подсистем –  $V^*_{OBN}$ ,  $V^*_{PR}$ ,  $V^*_{USTR}$ .

Из полученного множества отбираются варианты  $V^*$ , удовлетворяющие по стоимости:  $V^* = \{A_i \mid C(A_i) \leq C_{доп}\}$  и по потребляемым ресурсам:  $V^* = \{A_i \mid R(A_i) \leq R_{доп}\}$  (блок 7).

Далее из множества  $V^*$  отбираются варианты построения системы защиты  $V^{**}$ , удовлетворяющие требуемому уровню защищенности автоматизированной системы:  $V^{**} = \{A_i \mid Q(A_i) \geq Q_{TREB}\}$  (блок 8). Уровень защищенности автоматизированной системы можно определить обобщенным коэффициентом защищенности по методике, изложенной в [16].

Рациональный вариант построения системы защиты от РПВ  $A^*$  отбирается из множества вариантов  $V^{**}$  по условию минимального снижения качества функционирования защищаемой АСУ:  $A^* = \arg \min_{A_i \in V^{**}} K(A_i)$  (блок 9), где  $K$  – обобщенный

коэффициент снижения качества функционирования защищаемой автоматизированной системы, который определяется следующим образом:  $K = r_c K_C + r_{пк} K_{пк} + r_{пс} K_{пс} + r_{пар} K_{пар} + r_{рес} K_{рес} + r_{Тоткл} K_{Тоткл} + r_{ош} K_{ош}$  – коэффициент снижения производительности автоматизированной системы;

$K_{пк} = r_{Твып} K_{Твып} + r_d K_d$  – коэффициент снижения производительности обслуживающего персонала;  $r_c, r_{пк}, r_{пс}, r_{пар}, r_{рес}, r_{Тоткл}, r_{ош}, r_{Твып}, r_d$  – весовые коэффициенты значимости при соответствующих коэффициентах снижения производительности, которые определяются экспертным путем исходя из структуры защищаемой автоматизированной системы и решаемых ею задач.

Предлагаемый подход к построению систем защиты АСУ технологическими и производственными процессами может служить основой для разработки комплекса моделей и методов решения задачи структурно-параметрического синтеза систем защиты, позволяющих обеспечить, с одной стороны, требуемый уровень защищенности АСУ от РПВ, а с другой – минимальное влияние системы защиты на процессы функционирования защищаемой АСУ.

### Литература

1. Маринин С. Анализ причинно-следственных факторов применения компьютерного вируса «Стакснет» против ядерных объектов Ирана // Зарубежное военное обозрение. 2011. № 8. С. 34–39; № 9. С. 30–33.
2. Дроботун Е.Б. Малварь для промышленной автоматики // Хакер. 2012. № 12 (167). С. 86–90.
3. Воронов А.А. Введение в динамику сложных управляемых систем. М.: Наука. Глав. ред. Физматлит, 1985. 351 с.
4. Горбачев И.Е., Глухов А.П. Моделирование процессов нарушения информационной безопасности критической инфраструктуры // Тр. СПИИ РАН. 2015. Вып. 1 (38). С. 112–135.
5. Мальнев А. Противодействие реальным угрозам АСУ ТП // Информационная безопасность. 2015. № 4. С. 26–29.
6. Разрушающие программные воздействия: учеб.-метод. пособие; [под ред. М.А. Иванова]. М.: Изд-во НИЯУ МИФИ, 2011. 328 с.
7. Дроботун Е.Б. Об определении компьютерного вируса и вредоносной программы // Новые информационные технологии и системы 2010: сб. матер. IX Междунар. науч.-технич. конф. Пенза, 2010. С. 104–107.
8. Требования к обеспечению защиты информации ... 2014. URL: <http://www.fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/110-prikazy/864-prikaz-fstek-rossii-ot-14-marta-2014-g-n-31> (дата обращения: 01.12.2015).
9. Narumoto M. Assessing system performance against key performance metrics. October 2015. URL: <https://www.github.com/msnpn/performance-optimization/blob/master/Assessing-System-Performance-Against-KPI.md#insertlink#> (дата обращения: 01.12.2015).
10. Анализ ключевых показателей производительности. 2015. Ч. 1. URL: <http://www.habrahabr.ru/company/microsoft/blog/271547> (дата обращения: 01.12.2015).
11. Альтшуллер Г. Найти идею: Введение в ТРИЗ – теорию решения изобретательских задач. М.: Алпина Паблишерз, 2011. 400 с.
12. Одрин В.М. Метод морфологического анализа технических систем. М.: Изд-во ВНИИПИ, 1989. 205 с.
13. Банк данных угроз информационной безопасности. Список угроз. URL: <http://www.bdu.fstec.ru/threat> (дата обращения: 01.12.2015).
14. Банк данных угроз информационной безопасности. Список уязвимостей. URL: <http://www.bdu.fstec.ru/vul> (дата обращения: 01.12.2015).
15. Орлов А.И. Экспертные оценки: учеб. пособие. М.: Изд-во ИВСТЭ, 2002. 131 с.
16. Мукминов В.А., Хуцишвили В.М., Лобузько А.В. Методика оценки реального уровня защищенности автоматизированных систем // Программные продукты и системы. 2012. № 1 (97). С. 39–42.

DOI: 10.15827/0236-235X.115.051-059

Received 04.05.16

# SYNTHESIS OF PROTECTION SYSTEMS OF AUTOMATED CONTROL SYSTEMS AGAINST DESTROYING PROGRAM INFLUENCE

<sup>1</sup>*Drobotun E.B., Ph.D. (Engineering), Doctoral Student, drobotun@xakep.ru*<sup>1</sup>*Military Academy of the Aerospace Defence, Zhigarev St. 50, Tver, 170022, Russian Federation*

**Abstract.** Informatization and automation of various industrial facilities (including facilities that create potential safety threats) has made functioning safety of such objects one of the most important issues of modern age. Informatization and automation cause emergence of new types of threats for information security directed to violation of functioning of production objects that create a potential threat of safety, which can be exposed to destructive information and technical influences. These factors in combination with the increasing level of society dependence on functioning stability of the objects, which create potential safety threat, demand a package of measures (including technical measures) to decrease the risk of violation or complete cessation of such objects' functioning under information and technical influences.

The article presents one of possible approaches to constructing rational (in the context of ensuring the demanded security level and minimum influence of protection system on a protected automated system) system of protection against destroying program influences as the main type of possible information and technical influences on automated control systems for crucial and potentially dangerous objects.

**Keywords:** automated control system of technological processes, potentially dangerous industrial object, information security, destroying program influence, protection against destroying program influences.

## References

1. Marinin S. The analysis of cause and effect factors of the "Stuxnet" computer virus application against Iran's nuclear objects. *Zarubezhnoe voennoe obozrenie* [Foreign Military Review]. 2011, no. 8, pp. 34–39; no. 9, pp. 30–33 (in Russ.).
2. Drobotun E.B. Malware for industrial automatic equipment. *Haker* [Hacker]. 2012, no. 12 (167), pp. 86–90 (in Russ.).
3. Voronov A.A. *Vvedenie v dinamiku slozhnykh upravlyaemykh sistem* [Introduction to Dynamics of Difficult Operated Systems]. Moscow, Nauka, Fizmatlit Publ., 1985, 351 p. (in Russ.).
4. Gorbachev I.E., Glukhov A.P. Modeling of processes of information security violations of critical infrastructure. *Trudy SPIIRAN* [SPIIRAS Proc.]. 2015, no. 1 (38), pp. 112–135 (in Russ.).
5. Malnev A. Counteraction to real threats of an industrial control system. *Informatsionnaya bezopasnost* [Information Security]. 2015, no. 4, pp. 26–29 (in Russ.).
6. *Razrushayushchie programmnye vozdeystviya* [Destroying Program Influences]. Study guide. M.A. Ivanov (Ed.). Moscow, National Research Nuclear Univ. MEPhI Publ., 2011, 328 p.
7. Drobotun E.B. On the definition of a computer virus and the malicious program. *Sb. dokl. IX Mezhdunar. nauch.-tekhnich. konf. "Novye informatsionnye tekhnologii i sistemy 2010"* [Proc. 9th Int. Science and Technical Conf. "New Information Technologies and Systems 2010"]. Penza, 2010, pp. 104–107 (in Russ.).
8. *Trebovaniya k obespecheniyu zashchity informatsii ... 2014* [Requirements to Ensuring Information Security in Automated Control Systems for Production and Technological Processes on Crucial Objects, Potentially Dangerous Objects, and also the Objects Posing the Increased Hazard to Life and Human Health and for Surrounding Environment]. 2014, FSTEC Russia. Available at: <http://www.fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/110-prikazy/864-prikaz-fstek-ros-sii-ot-14-marta-2014-g-n-31> (accessed December 1, 2015).
9. Narumoto M. *Assessing system performance against key performance metrics*. 2015. Available at: <https://www.github.com/mspnp/performance-optimization/blob/master/Assessing-System-Performance-Against-KPI.md#insertlink#> (accessed December 1, 2015).
10. *Analiz klyuchevykh pokazateley proizvoditelnosti* [Analysis of Key Performance Indicators]. 2015, part 1. Available at: <http://www.habrahabr.ru/company/microsoft/blog/271547> (accessed December 1, 2015).
11. Altshuller G. *Nayti ideyu: Vvedenie v TRIZ – teoriyu resheniya izobretatelskikh zadach* [To Find an Idea: Introduction to TSIT – the Theory of Solving Inventive Tasks]. Moscow, Alpina Publ., 2011, 400 p.
12. Odrin V.M. *Metod morfologicheskogo analiza tekhnicheskikh sistem* [A Method of Morphological Analysis of Technical Systems]. Moscow, VNIPI Publ., 1989, 205 p.
13. *Bank dannykh ugroz informatsionnoy bezopasnosti. Spisok ugroz* [Information Security Threats Databank. List of Threats]. FSTEC Russia. Available at: <http://www.bdu.fstec.ru/threat> (accessed December 1, 2015).
14. *Bank dannykh ugroz informatsionnoy bezopasnosti. Spisok uyazvimostey* [Information Security Threats Databank. List of Vulnerabilities]. FSTEC Russia. Available at: <http://www.bdu.fstec.ru/vul> (accessed December 1, 2015).
15. Orlov A.I. *Ekspertnye otsenki* [Expert Estimates]. Study guide. Moscow, IVSTE Publ., 2002, 131 p.
16. Mukminov V.A., Khutishvili V.M., Lobuzko A.V. The assessment technique of real security level of an automated system. *Programmnye produkty i sistemy* [Software & Systems]. 2012, no. 1 (97), pp. 39–42 (in Russ.).

УДК 519.1

Дата подачи статьи: 08.07.16

DOI: 10.15827/0236-235X.115.060-066

## **АЛГОРИТМ ПОИСКА ОПТИМАЛЬНОГО РАСПОЛОЖЕНИЯ СЕНСОРОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ МОНИТОРИНГА ПРОСТРАНСТВА**

(Работа выполнена при поддержке РФФИ, гранты №№ 16-01-00342, 16-29-04268,  
и гранта Президента РФ № НШ-6831.2016.8)

**А.А. Кочкаров**, к.ф.-м.н., зам. директора Научно-тематического центра, [AKochkarov@paorti.ru](mailto:AKochkarov@paorti.ru)  
(Компания «РТИ», ул. 8 Марта, 10, стр. 1, г. Москва, 127083, Россия;  
Финансовый университет при Правительстве Российской Федерации,  
Ленинградский просп., 49, г. Москва, 125993, Россия);

**Д.В. Яцкин**, аспирант, [Danil@frtk.ru](mailto:Danil@frtk.ru)  
(Московский физико-технический институт (государственный университет),  
Институтский пер., 9, г. Долгопрудный, 141700, Россия)

Рассмотрены задача мониторинга пространства, переход к задаче обнаружения, а затем к задаче геометрического расположения сенсоров. Для решения поставленной задачи предлагается использовать децентрализованную сеть сенсоров. Устанавливаются ограничения и допущения, приводящие к задаче покрытия пространства. Проводится дискретизация задачи, обосновывается ее необходимость. Задача подробно рассматривается с математической точки зрения, разрабатывается алгоритм ее решения, оценивается его сложность. Проводится математическое и программное моделирование задачи. Разрабатывается программный комплекс, реализующий указанный алгоритм. По заданной карте местности и параметрам среды, а также с использованием ряда начальных условий, определяющихся физическими характеристиками сенсоров, программный комплекс позволяет рассчитать точки пространства, при размещении в которых устройств-обнаружителей задача мониторинга будет считаться решенной. На данном программном комплексе осуществлен ряд испытаний, показавших работоспособность, корректность и оптимальность разработанных подходов и алгоритма. Благодаря невысокой полиномиальной вычислительной сложности алгоритма, с помощью программного комплекса можно решать задачи мониторинга в конкретном случае и для больших зон мониторинга, и для сотен устройств-обнаружителей.

**Ключевые слова:** мониторинг, обнаружение, сенсор, покрытие пространства, теория множеств, алгоритмы покрытия, мобильный сенсор, сеть сенсоров, групповое управление, программный комплекс.

Для решения разнообразных задач все чаще требуется мониторинг пространства. Под мониторингом в общем случае понимается систематическое наблюдение за каким-либо процессом с целью фиксации соответствия (или несоответствия) результатов этого процесса первоначальным предположениям. В данной работе рассмотрим мониторинг пространства, то есть поставим задачу оперативного обнаружения заранее заданного объекта на определенной территории (при его появлении). Среди таких задач наиболее интересной кажется задача обнаружения заранее заданных объектов на определенной ограниченной территории. Следует отметить, что целевой объект может иметь разную природу: являться не только физическим объектом (предметом), но и, например, сигналом. Такие задачи часто возникают при проведении метеорологических наблюдений [1], экологического мониторинга [2], мониторинга сложных технических систем [3] и т.д.

Задача делится на две составные части: задачу распознавания и задачу геометрического расположения. Действительно, необходимо, с одной стороны, определить механизм обнаружения и распознавания целевого объекта неким обнаружителем, а с другой, расположить этот обнаружитель таким образом, чтобы обнаружение осуществлялось наиболее качественно.

Первая задача носит частный характер для каждого конкретного случая и зависит от свойств и параметров объекта, обнаружителя и среды. Предлагается решать ее децентрализованной сетью мобильных сенсоров. Такой подход имеет ряд преимуществ, в числе которых масштабируемость, гибкость, оперативность развертывания и свертывания системы (следовательно, перебазирования), самоорганизационная устойчивость к изменению факторов воздействия внешней среды (вплоть до факторов, способствующих выходу из строя устройств), а также возможность создания за счет изменения количества сенсоров широкой области обзора. При этом под областью обзора понимается область, при появлении целевого объекта в которой обнаружение может быть осуществлено без изменения геометрического положения устройств (алгоритмов поиска цели). Сами устройства (сенсоры) будем также называть абонентами.

Задача оптимального геометрического расположения сенсоров сводится к поиску такого набора геометрических положений сенсоров, который гарантирует решение задачи мониторинга. Понятие эффективности решения на данном этапе не определяется, так как существенным образом зависит от специфики задачи, о которой будет упомянуто далее. Дело в том, что от начальных условий зависят сама постановка задачи, подходы к ее решению,



а также понятия об эффективности решения задачи. В зависимости от геометрических параметров зоны мониторинга, а также количества сенсоров и геометрических параметров их зон принципиально выделяются два случая [4].

1. Существует такое множество точек в зоне мониторинга (за вычетом препятствий), что при размещении указанного (при задании начальных условий) числа сенсоров в них зона мониторинга полностью содержится в множестве, представляющем собой объединение зон видимости всех сенсоров. Иными словами, возможно покрытие зоны мониторинга указанным числом сенсоров с соответствующими характеристиками.

2. Зону мониторинга невозможно покрыть указанным (при задании начальных условий) числом зон видимости, соответствующих сенсорам, размещенным в зоне мониторинга (в любых точках, кроме препятствий).

Первый случай называется *задачей покрытия*, второй – *задачей патрулирования*. В данной статье рассматривается только первая задача.

#### Математическое описание задачи

Наиболее общая формулировка задачи покрытия такова:

- задана некая рабочая зона (зона мониторинга) – связная область пространства;
- даны  $N$  одинаковых устройств, каждому из которых соответствует своя зона покрытия (определяется аналогично для всех устройств);
- в рабочей зоне задано распределение конечного числа препятствий – связных множеств точек рабочей зоны, в которых устройства не могут находиться;
- каждое препятствие каким-либо образом (известным для каждой частной задачи) изменяет зону покрытия каждого устройства (в зависимости от взаимного расположения препятствия и устройства); после учета влияния всех препятствий образуется реальная зона видимости.

Требуется найти такие  $k \leq N$  наборов координат, что, если каждое устройство займет позицию, соответствующую одному из этих наборов (так, чтобы два устройства не использовали один набор координат), в рабочей зоне не будет точек, не входящих в реальную зону видимости какого-либо из устройств. При этом значение  $k$  должно быть минимальным из всех возможных.

Следует отметить следующий принципиальный момент. При подобной постановке задачи возможна ситуация, при которой препятствия будут изменять зону покрытия таким образом, что мониторинг в некоторых точках будет невозможен при любом взаимном расположении устройств. В таком случае будет сделан вывод о невозможном полном решении задачи с помощью сети таких устройств.

Однако в какой-то мере задача может быть решена, поэтому при реальной постановке требуется указать допустимые размеры слепых зон (зон, которые не могут входить в зону видимости ни одного из устройств независимо от расположения), при которых решение задачи считается допустимым.

Соответственно, математически представим задачу следующим образом.

- Дано связное ограниченное множество  $A$ , на котором задано  $K$  связных подмножеств  $B_i$ . При этом  $\bigcup_{i=1}^K B_i = B$ . Считаем, что множества задаются

на некоей евклидовой плоскости, а значит, можно вычислять расстояния между элементами этих множеств.

- Задан закон, ставящий в соответствие некоей точке рабочей зоны  $X=(x_1, x_2, x_3)$  ограниченное связное множество  $C(X)$ , которое содержит в себе точку  $X$ .

- Для каждого  $B_i$  задана функция  $f_i$ , преобразующая  $C(X)$  в каждой точке  $X$  рабочей зоны. Таким образом, при заданном распределении  $B_i$  каждой точке  $X$  можно поставить в соответствие множество  $D(X) = \bigcap_{i=1}^K f_i(C(X))$ .

Требуется найти  $k \leq N$  точек  $X_i$ , таких, что  $\forall i \in \{1, 2, \dots, k\} \rightarrow X_i \notin B; A \subset \bigcup_{i=1}^k D(X_i); k \rightarrow \min$ .

Соответствие введенных обозначений реальным физическим характеристикам приводится в таблице 1.

Пример взаимного расположения рабочей зоны, препятствий и зон видимости показан на рисунке 1.

#### Дискретизация задачи и алгоритм ее решения

Сведем задачу к дискретной.

Рабочая зона  $A$  представляется в виде сетки (некоего аналога расчетной сетки [5, 6])  $L=\{l_i\}$  с шагом  $\Delta$ .

Определяется сетка следующим образом:

- элементом сетки являются точки в рассматриваемом пространстве;
- для каждого элемента  $l_i$  на расстоянии  $\Delta$  находится как минимум один элемент  $l_j$ ; если таких элементов несколько, расстояние между ними равно или  $2\Delta$ , или  $\sqrt{2}\Delta$  соответственно; на евклидовой плоскости это соответствует некоей структуре, состоящей из квадратов, вершинами которых являются элементы  $L$ ;
- нет такой точки пространства  $A$ , в  $\Delta$ -окрестности которой (на расстоянии, не превышающем  $\Delta$ ) не находился бы хотя бы один элемент сетки  $L$ .

С точки зрения физики шаг  $\Delta$  должен быть пренебрежимо малым по отношению к характерным размерам задачи.

Таблица 1

## Соответствие математического и инженерного описаний задачи

Table 1

## Correspondence of mathematical and engineering task description

Инженерное описание	Математическое описание
Рабочая зона	$A$
Координаты точек рабочей зоны	$X=(x_1, x_2, x_3)$
Препятствия	$\bigcup_{i=1}^K B_i = B$
Зона видимости устройства, находящегося в точке $X$ , без учета влияния препятствий	$C(X)$
Зона видимости устройства, находящегося в точке $X$ , с учетом влияния препятствия под номером $n$	$f_n(C(X))$
Реальная зона видимости устройства, находящегося в точке $X$	$D(X) = \bigcap_{i=1}^K f_i(C(X))$

Итак, после перехода к описанию рабочей зоны в виде сетки начинается мышление в дискретных терминах. То есть, например, будем считать, что  $D(l_i)$  – дискретное множество, состоящее из элементов  $L$ , принадлежащих непрерывному множеству, соответствующему непрерывному  $D(l_i)$ , описанному ранее.

Поскольку рабочая зона  $A$  ограничена и  $\Delta$  – конечное число, можем считать, что задано конечное множество  $L=(l_1, l_2, \dots, l_M)$ . Действительно, ограниченность множества  $A$  означает, что  $\forall x \in A \exists M \in N \rightarrow \max_{y \in A} (dist(x, y)) \leq M\Delta$ , где  $N$  – множество натуральных чисел;  $dist(x, y)$  – расстояние между элементами  $x$  и  $y$ . Соответственно, общее число элементов сетки не может превышать  $M^2$  (при рассмотрении каждого элемента  $x \in A$ ). Это очень грубая оценка, но она в полной мере подтверждает конечность множества  $L$ . Обозначим множество подмножеств  $L$  как  $S=(S_1, S_2, \dots, S_p)$ . Здесь  $S_i=D(l_i)$  для всех  $l_i \in A \setminus B$ . Таким образом, надо найти покрытие множества  $A$  – подмножество  $S' \subseteq S$ , такое, что  $\bigcup_{S'_i \in S'} S'_i = L$ .

Покрывание  $S'$  называется минимальным, если не существует покрытия  $S'' \subseteq S'$ . Минимальных покрытий может быть несколько [7]. Покрытие  $S'$  называется наименьшим, если для любого минимального покрытия  $S''$  выполняется  $|S'| \leq |S''|$ , где  $|X|$  – мощность множества  $X$  [8]. Оптимально искать наименьшее покрытие  $S'$  [9].

**Лемма 1.**

Условие  $\bigcap_{S'_i \in S'} L \setminus S'_i = \emptyset$  является необходимым

и достаточным условием того, что  $S'$  – покрытие  $L$ .

**Доказательство.**

• **Необходимость.** Пусть  $S'$  – покрытие  $L$ . Следовательно, выполняется условие  $\bigcup_{S'_i \in S'} S'_i = L$ . Предположим, что  $\bigcap_{S'_i \in S'} L \setminus S'_i = F$ . Значит,

$L \setminus \left( \bigcap_{S'_i \in S'} L \setminus S'_i \right) = L \setminus F$ . Однако, с другой стороны,

$L \setminus \left( \bigcap_{S'_i \in S'} L \setminus S'_i \right) = \bigcup_{S'_i \in S'} S'_i$ . Значит,  $\bigcup_{S'_i \in S'} S'_i = L \setminus F$ , что может быть верным только в случае  $F=\emptyset$ . Что и требовалось доказать.

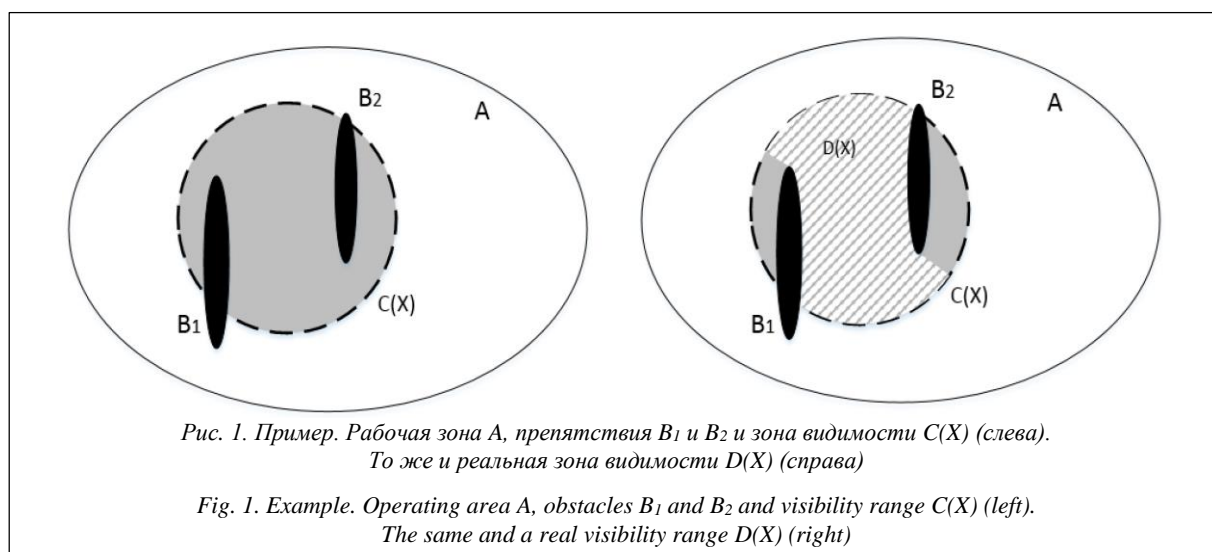


Рис. 1. Пример. Рабочая зона  $A$ , препятствия  $B_1$  и  $B_2$  и зона видимости  $C(X)$  (слева).  
То же и реальная зона видимости  $D(X)$  (справа)

Fig. 1. Example. Operating area  $A$ , obstacles  $B_1$  and  $B_2$  and visibility range  $C(X)$  (left).  
The same and a real visibility range  $D(X)$  (right)

• **Достаточность.**  $\bigcap_{S'_i \in S'} L \setminus S'_i = \emptyset$ . Из этого следует, что  $\bigcup_{S'_i \in S'} S'_i = L \setminus \left( \bigcap_{S'_i \in S'} L \setminus S'_i \right) = L \setminus \{\emptyset\} = L$ , то есть  $S'$  – покрытие  $L$ . Что и требовалось доказать.

**Лемма 2.**

Покрытие  $S'$  минимально тогда и только тогда, когда выполняется условие  $\forall S'_k \in S' \rightarrow \rightarrow \bigcap_{S'_i \in L \setminus \{S'_k\}} L \setminus S'_i \neq \emptyset$ , которое является необходимым и достаточным условием того, что покрытие  $S'$  – минимальное покрытие  $L$ .

**Доказательство.**

• **Необходимость.** Пусть  $S'$  – минимальное покрытие  $L$ . Предположим, что  $\exists S'_k \in S' \rightarrow \bigcap_{S'_i \in L \setminus \{S'_k\}} L \setminus S'_i = \emptyset$ . Тогда рассмотрим  $S'' = S' \setminus \{S'_k\}$ . Для этого множества выполняется  $\bigcap_{S'_i \in S''} L \setminus S'_i = \emptyset$ . Согласно лемме 1, это означает, что  $S''$  – покрытие  $L$ . Но  $S'' \subset S'$ , а значит,  $S'$  не является минимальным покрытием  $L$ . Что и требовалось доказать.

• **Достаточность.** Предположим, что условие  $\forall S'_k \in S' \rightarrow \bigcap_{S'_i \in L \setminus \{S'_k\}} L \setminus S'_i \neq \emptyset$  выполняется, но  $S'$  не является минимальным покрытием  $L$ . Значит,  $\exists S'' \subset S'$ , при этом  $S''$  – покрытие  $L$ . Тогда, согласно лемме 1,  $\bigcap_{S'_i \in S''} L \setminus S'_i = \emptyset$ . Однако из условия  $\forall S'_k \in S' \rightarrow \bigcap_{S'_i \in L \setminus \{S'_k\}} L \setminus S'_i \neq \emptyset$  напрямую следует, что  $\forall F \subset S' \rightarrow \bigcap_{F_i \in F} L \setminus S'_i \neq \emptyset$ . Значит,  $S''$  не может быть покрытием  $L$ , то есть получаем противоречие. Что и требовалось доказать.

**Метод построения полного нагруженного графа.** Строим полный нагруженный граф  $G=(V, E, L(V), L(E))$ . В нем множеству вершин  $V=\{v_i\}$ ,  $i \in \{1, 2, \dots, p\}$  взаимно однозначно сопоставляется множество  $L(V)=\{L(v_i)\}$ , где  $L(v_i)=L \setminus S_i$ . Соответственно, число вершин равно мощности множества  $S$ . Каждая вершина  $v_i$  нагружена множеством  $L(v_i)$ .

Множеству ребер  $E=\{e_{ij}\}=\{(v_i, v_j)\}$  сопоставляется множество  $L(E)=\{L(e_{ij})\}$ , где  $L(e_{ij})=L(v_i) \cap L(v_j)$ ,  $i \neq j$ . Множество всех ребер, инцидентных вершине  $v_i$ , будем обозначать  $E_i$ . Множество всех вершин, инцидентных ребрам из множества  $E_i$ , будем обозначать  $V(E_i)$ .

**Лемма 3.** Минимальное по мощности подмножество  $E' \subseteq E_i$  при выполнении условий  $L(v_i) \neq \emptyset$ ,  $\bigcap_{x \in E'_i} L(x) = \emptyset$  определяет минимальное покрытие  $S'$  множества  $L$ , однозначно соответствующее множеству  $V(E') \setminus \{v_i\}$  вершин, если  $\bigcap_{x \in V(E') \setminus \{v_i\}} L(x) = \emptyset$ , или множеству  $V(E')$ , если  $\bigcap_{x \in V(E') \setminus \{v_i\}} L(x) \neq \emptyset$ .

**Доказательство.**

$L(v_i) \neq \emptyset$ , следовательно,  $\{S_i\}$  не является покрытием множества  $L$ .

Условие  $\bigcap_{x \in E'_i} L(x) = \emptyset$  в силу того, что, по определению,  $L(e_{ij})=L(v_i) \cap L(v_j)$ ,  $i \neq j$ , можно записать как  $\bigcap_{x \in V(E'_i)} L(x) = \emptyset$ . Значит, согласно лемме 1, множество  $\hat{S}$ , соответствующее  $V(E')$ , определяет покрытие множества  $L$ .

Можно также утверждать, что при соблюдении условия  $\bigcap_{x \in V(E') \setminus \{v_i\}} L(x) = \emptyset$  множество  $\tilde{S}$ , однозначно соответствующее  $V(E') \setminus \{v_i\}$ , определяет покрытие множества  $L$ .

Установим теперь, являются ли эти покрытия минимальными. Исходя из предположения о том, что  $E' \subseteq E_i$  – минимальное по мощности подмножество, и рассматривая его, можно сделать вывод о возможности построения  $E'_i$ , удовлетворяющего условию леммы 2 (при рассмотрении отдельно множеств  $\hat{S}$  и  $\tilde{S}$  с соблюдением соответствующих условий). Что и требовалось доказать.

**Лемма 4.** Минимальное по мощности подмножество  $E' \subseteq E_i$  при выполнении условий  $|L(e_{ij})| = \min_{x \in E} |L(x)|$ ,  $\forall v \in V \rightarrow L(v) \neq \emptyset$ ,  $\bigcap_{x \in E'_i} L(x) = \emptyset$  определяет наименьшее покрытие  $S'$ , однозначно соответствующее множеству  $V(E_i) \setminus \{v_i\}$  вершин, если  $\bigcap_{x \in V(E_i) \setminus \{v_i\}} L(x) = \emptyset$ , или множеству  $V(E_i)$ , если  $\bigcap_{x \in V(E_i) \setminus \{v_i\}} L(x) \neq \emptyset$ .

**Доказательство.**

Поскольку  $\forall v \in V \rightarrow L(v) \neq \emptyset$ , тривиального решения не существует.

$E' \subseteq E_i$  выбирается таким образом, чтобы выполнялось условие  $\bigcap_{x \in E'_i} L(x) = \emptyset$ . Значит, используя лемму 3, можно сделать вывод о том, каким образом определяется минимальное покрытие. Однако, учитывая условие  $|L(e_{ij})| = \min_{x \in E} |L(x)|$ , можно сделать вывод о том, что данное покрытие также будет являться наименьшим. Что и требовалось доказать.

На основании всего вышесказанного, используя названные леммы, особенно лемму 4, можно выработать следующий алгоритм.

**Алгоритм построения наименьшего покрытия множества  $L$ .**

1. Имея множество  $L$ , задающее рабочую область, строим множество подмножеств  $S=(S_1, S_2, \dots, S_p)$ , где  $S_i=D(L_i)$  для всех  $L_i \in A \setminus B$ .

2. Ставим в соответствие этим множествам множество  $L(V^{(0)})=\{L \setminus S_i\}$ . Если один из его элементов – пустое множество, делаем вывод о том, что существует тривиальное решение, алгоритм закончен. Иначе – переход к шагу 3.

3. Строим полный нагруженный граф  $G^{(0)}=(V^{(0)}, E^{(0)}, L(V^{(0)}), L(E^{(0)}))$ . (О том, какие вершины и ребра какими множествами надо нагружать, уже говорилось.)

4. Проверка существования покрытия. Для произвольной вершины  $v_i^{(0)}$  определяется подмножество  $L(E_i^{(0)}) = \bigcap_{x \in E_i^{(0)}} L(x)$ . Если в результате получается непустое множество, делается вывод о том, что покрытия не существует, алгоритм закончен. В противном случае – переход к шагу 5.

5.  $k \leftarrow 0$ . Присвоение переменной  $k$  значения 0.

6. В полном нагруженном графе  $G^{(k)}$  осуществляется поиск ребра  $e_{ij}^{(k)}$ , для которого  $|L(e_{ij}^{(k)})| = \min_{x \in E^{(k)}} |L(x)|$ . Если  $\bigcap_j L(e_{ij}^{(k)}) = \emptyset$ , осуществляется переход к шагу 8, иначе – к шагу 7.

7. Строится полный нагруженный граф  $G^{(k+1)}=(V^{(k+1)}, E^{(k+1)}, L(V^{(k+1)}), L(E^{(k+1)}))$  следующим образом:  $V^{(k+1)} = V^{(k)} \setminus \{v_i^{(k)}\}$ ,  $E^{(k+1)} = E^{(k)} \setminus E_i^{(k)}$ ,  $L(v_i^{(k+1)}) = L(e_{ii}^{(k)})$ ,  $L(e_{ip}^{(k+1)}) = L(e_{ii}^{(k)}) \cap L(e_{ip}^{(k)})$  для всех  $v_i^{(k+1)} \in V^{(k+1)}$ ,  $e_{ip}^{(k+1)} \in E^{(k+1)}$ ,  $k \leftarrow k+1$ .

Переход к шагу 6.

8. Начало построения наименьшего покрытия  $V = V(\{e_{ij}^{(k)}\}) = \{v_i^{(k)}, v_j^{(k)}\}$ .

9. Если  $k=0$ , осуществляется переход к шагу 12. В противном случае  $k \leftarrow k-1$ , переход к шагу 10.

10. В графе  $G^{(k)}$  определяется подмножество  $L(V) = \bigcup_{x \in V} L(x)$ .

11. Если в  $G^{(k)}$  выполняется условие  $L(V) \neq \emptyset$ , то  $V = V \cup \{v_i^{(k)}\}$ . Переход к шагу 9.

12. Множество  $S' = \{L \setminus L(v_i^{(o)}): v_i^{(o)} \in V\}$  определяет наименьшее покрытие множества  $L$ .

Конец алгоритма.

#### Лемма 5.

Трудоёмкость алгоритма оценивается как  $O(p)=p^3+p^2$ .

#### Доказательство.

Вся вычислительная мощность реализуется при построении полного нагруженного графа. Число вершин графа совпадает с мощностью множества  $S$  и равняется  $p$ , как уже было определено ранее. Для  $p$  вершин графа имеем  $p$  действий по вычислению нагрузки в вершинах и  $\frac{(p-1)p}{2}$  действий для вычисления нагрузки ребер полного графа. Такие вычисления выполняются  $(p-2)$  раза. Таким образом, грубая оценка  $O(p)=p^3+p^2$ .

Величина  $p$  зависит от соотношения между размерами множества  $A \setminus B$  и величиной  $\Delta$ .

### Программное моделирование

Для апробации разработанных алгоритмов была построена программная (компьютерная) модель, позволяющая провести ряд экспериментов и рассчитать основные показатели, свидетельствующие о работоспособности и эффективности представленного алгоритма.

На вход программы поступает обработанный участок карты какой-либо местности, соответствующий зоне мониторинга (рис. 2).

Также на вход программы подается число устройств, которым предполагается решить задачу. Зоны покрытия (видимости) каждого абонента считаются заранее заданными окружностями определенного радиуса. Препятствия не влияют на зоны видимости.

Пример работы модели показан на рисунке 3, где найденные положения сенсоров обозначены зелеными точками. Для наглядности показано не только решение задачи покрытия для 12 абонентов, но и то, каким образом устройства, расположенные

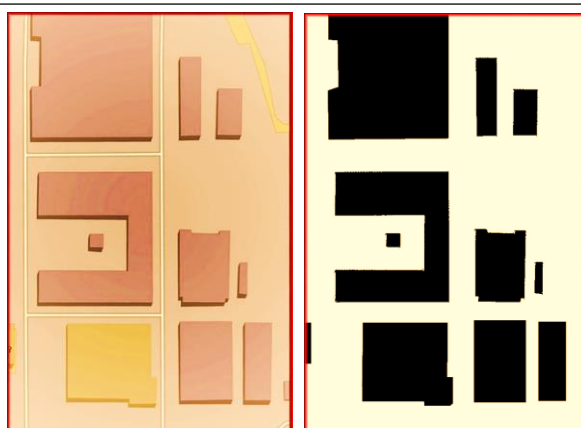


Рис. 2. Предварительная обработка карты местности

Fig. 2. Base map preprocessing

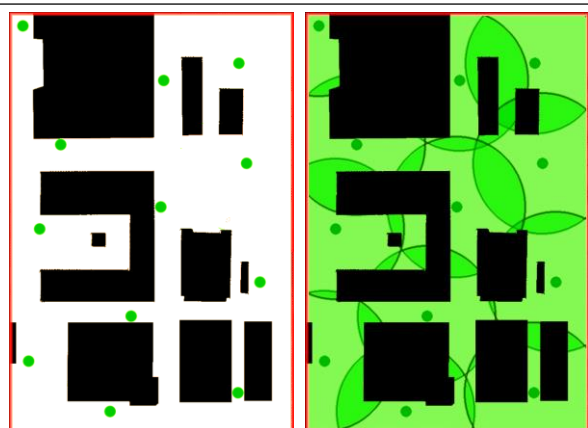


Рис. 3. Решение задачи покрытия для 12 абонентов с указанием зон покрытия (справа) и без указания (слева)

Fig. 3. Covering problem solving for 12 subscribers with determination of coverage area (right) and without (left)

в обозначенных точках, покрывали бы своими зонами видимости весь участок пространства, соответствующий зоне мониторинга. Отметим достаточно большое число попарных пересечений зон видимости абонентов, характерных для зон видимости такой формы.

Для проведения экспериментов использовались заранее обработанные участки карты Москвы. Поскольку здания имеют самые разные размеры и форму, для того чтобы охарактеризовать размер выбираемых участков, используется их отношение к радиусу зон видимости абонентов  $R$ .

В настоящей работе были рассмотрены 84 прямоугольных участка карт, линейные размеры которых  $4R$  и  $6R$ . Соответственно, площади рассматриваемых участков карты к площади зон видимости

относятся как  $\frac{4 \cdot 6}{\pi} \approx 7.64$ . Таким образом, минимальное

число устройств, для которых может быть решена задача покрытия при данных начальных условиях, равно 8.

По результатам моделирования были получены следующие результаты.

Программа успешно завершает свою работу (находит решение задачи) для 78 из 84 случайно выбранных карт, что составляет 93 %. Эта цифра постоянна для любого числа абонентов. Ситуации, при которых задача не может быть решена, возникают тогда, когда препятствия (в совокупности с границами зоны мониторинга) отсекают от выбранного участка карты связную область. В этом случае абонент никаким образом не может попасть туда и в некоторых случаях в силу этого не может осуществлять мониторинг в этой зоне. Также такая ситуация может возникнуть при слишком больших линейных размерах препятствий. Информация о количестве успешных решений задач покрытия для разного числа абонентов приведена в таблице 2.

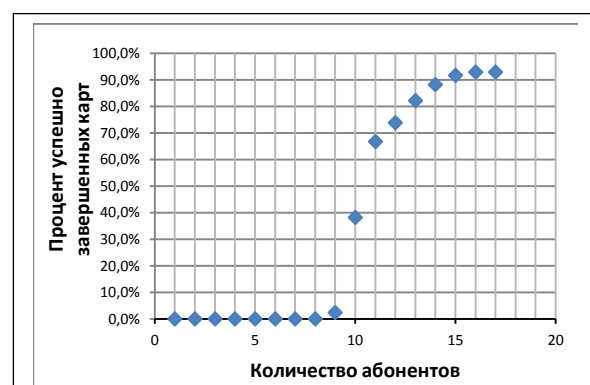


Рис. 4. Зависимость процента успешно завершённых карт при решении задачи покрытия от числа абонентов

Fig. 4. Dependence of a percentage of successfully covered maps when solving only covering problem from the number of subscribers

Таблица 2  
Зависимость решения задачи покрытия  
от числа абонентов

Table 2  
The dependence of covering problem solution  
on the number of subscribers

Число абонентов	Задача покрытия, %
1	0,0
2	0,0
3	0,0
4	0,0
5	0,0
6	0,0
7	0,0
8	0,0
9	0,0
10	2,4
11	38,1
12	66,7
13	73,8
14	82,1
15	88,1
16	91,7
17	92,9

Из таблицы видно, что при использовании 16 абонентов задача успешно решается на максимально возможном проценте карт, соответственно, дальнейшее увеличение числа абонентов не приведет к изменению этого процента и будет бессмысленным.

Зависимость процента карт, на которых успешно решается задача покрытия, от числа абонентов показана на графике (рис. 4), из которого следует, что уже для 13 абонентов задача покрытия решается более, чем в 80 % случаев, а для 15 – более, чем в 90 %. На основании этого графика можно делать вывод о вероятности успешного завершения случайной карты при выборе того или иного числа абонентов для решения задачи покрытия.

В целом результаты экспериментов показали работоспособность предложенных алгоритмов и подходов.

В заключение отметим, что в работе предложен и обоснован алгоритм, позволяющий решать задачу мониторинга пространства в частном случае. На его основе был разработан программный комплекс, позволяющий находить решение задачи мониторинга с помощью заданного числа мобильных сенсоров при известной карте пространства. Работоспособность алгоритма была подтверждена экспериментально. Общий подход, продемонстрированный в данной статье, позволяет применять разработанные методы для использования широкого спектра устройств в качестве субъектов мониторинга при решении соответствующих задач.

Разработанный программный комплекс имеет потенциал для применения при решении задач, связанных с мониторингом пространств децентрализованной сетью мобильных сенсоров. Данный ком-

плекс может войти в число системы программ, позволяющей эффективно решать такие задачи.

### Литература

1. Израэль Ю.А. Глобальная система наблюдений. Прогноз и оценка окружающей природной среды. Основы мониторинга // Метеорология и гидрология. 1974. № 7. С. 3–8.
2. Кондратьев А.Д., Королева Т.В., Пузанов А.В., Черницова О.В., Ефременков А.А., Шарапова А.В., Горбачев И.В., Двуреченская Е.Б. Совершенствование системы экологического мониторинга районов падения отделяющихся частей ракет-носителей // МНКО. 2012. № 6 (37). С. 483–486.
3. Кульба В.В., Сомов Д.С., Кочкаров А.А. Применение структурно-интегрированных индикаторов в мониторинге сложных технических систем // Изв. ЮФУ. Сер. Технич. науки. 2011. № 3 (116). С. 52–64.
4. Кочкаров А.А., Яцкин Д.В., Рахманов О.А. Особенно-

сти решения задачи геометрического мониторинга // Изв. ЮФУ. Сер. Технич. науки. 2016. № 2 (175). С. 158–168.

5. Yap P. Grid-based path-finding. *Advances in Artificial Intelligence*, Springer, Berlin Heidelberg, 2002, pp. 44–55.
6. Thompson J.F., Warsi Z.A., Mastin C.V. *Numerical grid generation, foundations and applications*. North-Holland, 1985, 438 p.
7. Cheng M., Ruan L., and Wu W. Achieving minimum coverage breach under bandwidth constraints in wireless sensor networks. *Proc. 24th Annual Joint Conf. of the IEEE Computer and Communications Societies*, 2005, vol. 4, pp. 2638–2645.
8. Еремеев А.В., Заозерская Л.А., Колоколов А.А. Задача о покрытии множества: сложность, алгоритмы, экспериментальные исследования // Дискретный анализ и исследование операций. 2000. Т. 7. № 2. С. 22–46.
9. Батищев Д.Д., Неймарк Е.А., Старостин Н.В. Применение генетических алгоритмов к решению задач дискретной оптимизации. Н. Новгород: Изд-во НГУ, 2006.

DOI: 10.15827/0236-235X.115.060-066

Received 08.07.16

### SEARCH ALGORITHM OF OPTIMAL SENSORS LOCATION FOR SOLVING THE SPACE MONITORING PROBLEM

(Acknowledgements. This work was financially supported by RFBR, grant no. 16-01-00342, 16-29-04268, and the grant of the President of the Russian Federation no. HIII-6831.2016.8)

<sup>1,2</sup>Kochkarov A.A., Ph.D. (Engineering), Deputy Director R&D centre, AKochkarov@oaorti.ru

<sup>3</sup>Yatskin D.V., Postgraduate Student, Danil@frtk.ru

<sup>1</sup>OJSC "RTI", March 8 St. 10, Bld. 1, Moscow, 127083, Russian Federation

<sup>2</sup>Financial University under the Government of the Russian Federation, Leningradsky Ave. 49, Moscow, 125993, Russian Federation

<sup>3</sup>Moscow Institute of Physics and Technology, Institutsky Line 9, Dolgoprudny, 141700, Russian Federation

**Abstract.** The paper considers the task of space monitoring, an algorithm and software package to solve it. From this problem the authors move to the problem of detection, and then to the problem of geometrical sensors location. It is proposed to use a decentralized robotic network to solve thy proposed problem. The authors set limitations and assumptions for sensors location problem leading to the problem of space covering. They also discretize problem parameters. The paper describes the problem in detail from a mathematical point of view, develops the algorithm to solve it, estimates its complexity. There mathematical and software models of the problem. The authors develop a software package that implements the specified algorithm. The map of the area and a number of physical parameters of the sensor serve as input data for a software package. The package allows calculating points in space, where the placed devices (detectors) help to solve the monitoring problem. This software package was used for a series of tests. The tests showed efficiency, correctness and optimality of the developed approaches and the algorithm. Due to the low polynomial computational complexity of the algorithm the software package can solve monitoring tasks for the particular case of large areas of monitoring, as well as for a big number of sensors.

**Keywords:** monitoring, detection, sensor, UAV, covering space, set theory, covering algorithms, mobile robot, robotic network, group management, software package.

### References

1. Izrael Yu.A. Global monitoring system. Forecast and assessment of environment. Monitoring basics. *Meteorologiya i gidrologiya* [Meteorology and Hydrology]. 1974, no. 7, pp. 3–8 (in Russ.).
2. Kondratyev A.D., Koroleva T.V., Puzanov A.V., Chernitsova O.V., Efremenkova A.A., Sharapova A.V., Gorbachev I.V., Dvurechenskaya E.B. Ecologic monitoring system development for booster rocket drop zones. *Mir nauki, kultury, obrazovaniya*. 2012, no. 6 (37), pp. 483–486 (in Russ.).
3. Kulba V.V., Somov D.S., Kochkarov A.A. Using structure integrated indicators in monitoring of complex technical systems. *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences]. 2011, no. 3 (116), pp. 52–64 (in Russ.).
4. Kochkarov A.A., Yatskin D.V., Rakhmanov O.A. Specific features of solving a geometrical monitoring problem. *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences]. 2016, no. 2 (175), pp. 158–168 (in Russ.).
5. Yap P. Grid-based path-finding. *Advances in Artificial Intelligence*. Berlin Heidelberg, Springer Publ., 2002, pp. 44–55.
6. Thompson J.F., Warsi Z.A., Mastin C.V. *Numerical Grid Generation, Foundations and Applications*. Amsterdam, North-Holland Publ., 1985.
7. Cheng M., Ruan L., Wu W. Achieving minimum coverage breach under bandwidth constraints in wireless sensor networks. *Proc. 24th Annual Joint Conf. of the IEEE Computer and Communications Societies*. 2005, vol. 4, pp. 2638–2645.
8. Eremeev A.V., Zaozerskaya L.A., Kolokolov A.A. A set covering problem: complexity, algorithms, experimental tests. *Diskretny analiz i issledovanie operatsiy* [Journal of Applied and Industrial Mathematics]. 2000, vol. 7, no. 2, pp. 22–46 (in Russ.).
9. Batishchev D.D., Neymark E.A., Starostin N.V. *Primenenie geneticheskikh algoritmov k resheniyu zadach diskretnoy optimizatsii* [Applying Genetic Algorithms to Solve Discrete Optimization Problems]. N. Novgorod, NGU Publ., 2006.

УДК 621.395.7, 519.854

Дата подачи статьи: 19.01.16

DOI: 10.15827/0236-235X.115.067-073

## **ПЧЕЛИНЫЙ АЛГОРИТМ ОПТИМИЗАЦИИ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ПЛАНИРОВАНИЯ БЕСПРОВОДНОЙ СЕТИ**

*Е.С. Скаков, аспирант, wallkiry@mail.ru*

*(Липецкий государственный педагогический университет,  
ул. Ленина, 42, г. Липецк, 398020, Россия);*

*В.Н. Малыш, д.т.н., профессор, vmalysh@mail.ru*

*(Липецкий филиал Российской академии народного хозяйства и государственной службы  
при Президенте Российской Федерации, ул. Интернациональная, 3, г. Липецк, 398050, Россия)*

Статья посвящена решению проблемы планирования беспроводной сети. В ней сформулирована задача размещения базовых станций, учитывающая межсотовые помехи, затухание при распространении сигнала, несколько типов базовых станций. Решение задачи представляется в виде массива структур, хранящих информацию об одном месте-кандидате (тип установленной базовой станции, список подключенных к ней клиентов).

Разработана модификация одного из метаэвристических методов – пчелиного алгоритма BCO (Bee Colony Optimization). Метод оптимизации подражением пчелиной колонии относится к мультиагентным методам, основанным на моделировании интеллектуального поведения колоний агентов, так называемым методом роевого интеллекта. В природе подобным интеллектом обладают группы общественных насекомых, например, колонии муравьев, пчел, термитов. В основе предлагаемого алгоритма лежит понятие окрестности текущего решения. Для формирования окрестности используются небольшие операции по изменению текущей конфигурации беспроводной сети. Новое решение из окрестности текущего может быть получено при помощи одной из двух операций: смена типа одной станции на более дешевый, удаление одной базовой станции.

Новый алгоритм реализован как ПО в среде разработки Delphi. Показано, что пчелиный алгоритм позволяет находить решение задачи размещения базовых станций за приемлемое время, на много порядков быстрее точного метода полного перебора. Выявлена зависимость качества решения поставленной задачи методом подражания пчелиной колонии от значений управляющих параметров алгоритма.

**Ключевые слова:** размещение базовых станций, пчелиный алгоритм, беспроводные сети, оптимизация, SIR.

Важным этапом при проектировании беспроводной сети передачи данных является синтез ее топологической структуры. Он подразумевает планирование территориального размещения базовых станций и подключение к ним клиентов.

Отметим работы [1] и [2], посвященные проблеме размещения базовых станций (БС) в сетях стандарта WiMAX. Статья [3] посвящена классификации алгоритмов размещения БС в сетях стандарта UMTS. В данной статье сформулирована задача размещения БС (ЗРБС), в основе которой лежит математическая модель, учитывающая уровень SIR для клиентов, уровень затухания сигнала при распространении от клиента к БС и обратно, наличие в сети нескольких типов БС. Сформулированная задача решается при помощи метаэвристического пчелиного алгоритма оптимизации, который демонстрирует высокую точность и скорость нахождения решений.

### **Постановка задачи размещения БС**

ЗРБС формулируется следующим образом. Пусть есть некоторое множество клиентов ( $N_{cp}$  штук). Все клиенты должны быть подключены к БС, причем есть заранее определенное множество ( $N_{ps}$  штук) потенциальных мест, куда может быть установлена БС (так называемых мест-кандидатов). Для установки доступны БС нескольких типов ( $N_{types}$  типов), отличающихся по своим характе-

ристикам. При этом должен выполняться ряд требований и ограничений [4].

Для описания решения ЗРБС воспользуемся такой единицей представления данных, как структура. Структура представляет собой композитный тип данных, который позволяет хранить набор переменных различного типа (т.н. полей), объединенных одним именем [5]. Для ЗРБС структура состоит из трех полей:

- целочисленная переменная *type*, отражающая тип БС, установленной на данном месте (считается, что БС не установлена, если *type* = 0);

- целочисленная переменная *cl\_nbr*, отражающая число клиентов, подключенных к данному месту-кандидату;

- вектор (одномерный массив) целых чисел **CL** размерности *cl\_nbr*, содержащий номера клиентов, подключенных к данному месту-кандидату.

Решение ЗРБС представляет собой вектор **Sol**, каждый элемент которого является структурой, соответствующей одному месту-кандидату (пояснение: здесь и в дальнейшем запись вида **Arr**[*j*] означает обращение к *j*-му элементу вектора **Arr**; элементы массивов нумеруются, начиная с 1). Обращение к полю структуры обозначается символом «.». То есть **Sol**[5].*type* означает тип БС пятого места-кандидата, а, например, **Sol**[3].**CL**[2] означает номер 2-го клиента, подключенного к станции на месте № 3.

Клиента можно подключить к месту-кандидату, только если на нем уже установлена БС, то есть для



места, у которого  $\text{Sol}[s].cl\_nbr \neq 0$ , должно выполняться  $\text{Sol}[s].type \neq 0$ .

Пусть  $\mathbf{b}$  – вектор, содержащий значения требуемых полос пропускания для клиентов;  $\mathbf{\beta}$  – вектор производительностей БС разного типа. Для каждой БС трафик всех клиентов не должен превышать значения производительности станции:

$$\sum_{i=1}^{\text{Sol}[s].cl\_nbr} \mathbf{b}[\text{Sol}[s].\text{CL}[i]] \leq \mathbf{\beta}[\text{Sol}[s].type], \quad (1)$$

$$\forall s \in \{1, 2, \dots, N_{ps}\}.$$

Пусть  $\mathbf{P}_{TP}^{\max}$  – вектор, элементами которого являются максимальные мощности клиентов;  $\mathbf{P}_{BS}^{\max}$  – вектор максимальных мощностей БС разного типа;  $\mathbf{P}_{TP}^{\text{tar}}$  – вектор, элементами которого являются чувствительности клиентов;  $\mathbf{P}_{BS}^{\text{tar}}$  – вектор чувствительностей БС разного типа;  $\mathbf{G}$  – двумерный массив ( $N_{tp} \times N_{ps}$ ), каждый элемент которого  $0 \leq \mathbf{G}[i][s] \leq 1$  показывает уровень затухания между  $s$ -м местом и клиентом с номером  $i$ .

Для режима downlink существует мощностное ограничение. Мощность, дошедшая к клиенту от БС, несмотря на затухание сигнала, должна превышать минимальную целевую мощность:

$$\frac{\mathbf{G}[\text{Sol}[s].\text{CL}[i]][s] \cdot \mathbf{P}_{BS}^{\max}[\text{Sol}[s].type]}{\mathbf{P}_{TP}^{\text{tar}}[\text{Sol}[s].\text{CL}[i]]} \geq 1, \quad (2)$$

$$\forall s \in \{1, 2, \dots, N_{ps}\}, \quad \forall i \in \text{Sol}[s].\text{CL}.$$

Мощностное ограничение для режима uplink:

$$\frac{\mathbf{G}[\text{Sol}[s].\text{CL}[i]][s] \cdot \mathbf{P}_{TP}^{\max}[\text{Sol}[s].\text{CL}[i]]}{\mathbf{P}_{BS}^{\text{tar}}[\text{Sol}[s].type]} \geq 1, \quad (3)$$

$$\forall s \in \{1, 2, \dots, N_{ps}\}, \quad \forall i \in \text{Sol}[s].\text{CL}.$$

Минимизируемая целевая функция имеет вид

$$F = \sum_{s=1}^{N_{ps}} \text{Cost}[\text{Sol}[s].type] + \sum_{s=1}^{N_{ps}} \sum_{i=1}^{\text{Sol}[s].cl\_nbr} \text{SIR}_{dB}^{si} \cdot k, \quad (4)$$

где **Cost** – массив стоимостей БС разного типа.

Первое слагаемое формулы (4) – стоимость обслуживания. Второе слагаемое учитывает уровень *SIR* для клиентов беспроводной сети. *SIR* (signal-to-interference ratio) – отношение уровня сигнала к уровню помех [6], рассчитываемое следующим образом:  $\text{SIR}_{dB} = 10 \lg(P_{\text{signal}}/P_{\text{interference}})$ .

В формуле (4) *SIR* рассчитывается для каждого клиента отдельно. Числитель содержит сигнал, дошедший от той БС, к которой подключен клиент, в знаменателе фигурируют сигналы от остальных БС, создающие помехи:

$$\text{SIR}_{dB}^{si} = 10 \lg \left( \frac{\mathbf{G}^{si} \cdot \mathbf{P}_{BS}^{\max}[\text{Sol}[s].type]}{\sum_{q=1}^{N_{ps}} \sum_{z=1}^{\text{Sol}[q].cl\_nbr} \mathbf{G}^{qz} \cdot \mathbf{P}_{BS}^{\max}[\text{Sol}[q].type] \cdot w_{siqz}} \right),$$

$$\mathbf{G}^{si} = \mathbf{G}[\text{Sol}[s].\text{CL}[i]][s], \quad \mathbf{G}^{qz} = \mathbf{G}[\text{Sol}[q].\text{CL}[z]][q],$$

где  $w_{siqz} = 0$ , если  $s = q$  и  $i = z$ , иначе  $w_{siqz} = 1$ .

Коэффициент  $k$  отвечает за одновременный учет в целевой функции затрат на создание комплекса БС и уровня межсетевых помех для клиентов.  $k$  должен иметь размерность стоимости БС, обеспечивая сохранение размерности в формуле (4), так как *SIR* – безразмерная величина. Коэффициент  $k$  должен обеспечивать «штраф» за низкий уровень *SIR* и «награду» за высокий. В данной статье принято, что  $k = -10$  условных единиц стоимости ( $k$  должен быть отрицательным, так как ЗРБС – минимизационная задача, а значит, высокий уровень *SIR* должен уменьшать целевую функцию).

Отсортируем типы БС по возрастанию стоимости, то есть сделаем так, что  $i$ -й тип БС дороже  $(i-1)$ -го и дешевле  $(i+1)$ -го.

### Общие сведения о роевом интеллекте

Метод оптимизации подражанием пчелиной колонии относится к мультиагентным методам, основанным на моделировании интеллектуального поведения колоний агентов, так называемым методом роевого интеллекта (Swarm Intelligence). В природе подобным интеллектом обладают группы общественных насекомых, например колонии муравьев, пчел, термитов [7]. Другие названия метода оптимизации подражанием пчелиной колонии – пчелиный алгоритм, метод/алгоритм пчелиного роя, метод/алгоритм пчелиной колонии. В настоящей статье для обозначения данного класса метаэвристик будет использоваться словосочетание «пчелиный алгоритм».

Отметим некоторые особенности роевого интеллекта. Каждый из агентов роя характеризуется стохастическим поведением на основе собственного восприятия окрестности. Агенты руководствуются локальными правилами, не имея никакого представления о глобальной цели. Взаимодействие между подобными самоорганизующимися агентами приводит к появлению некой формы коллективного разума под названием роевой интеллект [8]. Особи, входящие в рой, коллективно используют окружающую среду и ресурсы. Ключевой особенностью роевой системы является самоорганизация: низкоуровневые взаимодействия (так называемый микроскопический уровень) приводят к изменениям на глобальном (макроскопическом) уровне [8].

Под самоорганизацией роя в [9] подразумеваются следующие четыре особенности.

1. Положительная обратная связь. Например, на основе уровня концентрации феромона, отложенного другими муравьями, муравей выбирает себе путь от муравейника до источника питания (чем выше уровень феромона на тропе, тем с большей вероятностью муравей выберет эту тропу).

2. Отрицательная обратная связь. Уравновешивает положительную обратную связь и помогает стабилизировать коллективный шаблон поведения.



Например, на основе информации от других пчел пчела может решить, что ее текущий источник нектара значительно хуже других найденных источников, и оставить этот источник.

3. Случайность. Флуктуации (случайные блуждания, ошибки, вероятностный поиск) очень важны для функционирования роевого интеллекта, так как позволяют рою находить новые решения.

4. Множественность взаимодействия. Агенты роя используют информацию, поступающую от других агентов, таким образом, данные об окружающей среде распространяются по всему рою.

Особенности самоорганизации, отмеченные в [9], четко прослеживаются в пчелиных роях, однако решение оптимизационных задач на основе подражания поведению пчел – довольно молодая дисциплина: первые работы, посвященные пчелиному алгоритму, датированы началом 2000-х годов.

### Биологические основы пчелиного алгоритма

Пчелы в природе ищут пищу путем изучения пространства в окрестности их улья. Как правило, на начальной стадии несколько пчел-разведчиков исследуют окружающую среду. По завершении поиска пчелы-разведчики возвращаются в улей и информируют других членов роя о месте, количестве и качестве доступных источников питания, которые были найдены ими. Обмен информацией происходит при помощи танца на специально отведенной для этого площадке. Если пчела, наблюдавшая танцы скаутов, решает покинуть улей и собирать нектар, она будет следовать за одним из разведчиков к одному из ранее обнаруженных источников пищи. Такая пчела становится занятым фуражиром. Она занимается сбором нектара, при этом уточняя информацию о количестве нектара в окрестности найденного источника. После сбора фуражир возвращается в улей и оставляет там собранный нектар. Затем он может совершить одно из следующих действий [10]:

- стать незанятым фуражиром, оставив свой текущий источник нектара;
- продолжить добывать нектар из своего источника, не вербуя незанятых пчел посредством танца;
- продолжить добывать нектар из своего источника, при этом вербуя незанятых пчел.

Описанный процесс продолжается непрерывно, в то время как улей накапливает нектар и исследует новые области с потенциальными источниками питания.

### Алгоритм BCO

На данный момент одной из самых популярных концепций роевого интеллекта является метод

BCO (Bee Colony Optimization), предложенный в начале 2000-х годов [11]. Подробное описание алгоритма и его возможных разновидностей приведено в работах [10] и [12].

В данной статье рассматривается модифицированный алгоритм BCO (variant of BCO based on the improvement concept, BCOi) [13]. Особенностью BCOi является то, что рассматривается работа с полными (завершенными) решениями оптимизационной задачи, а не с частичными, как в классическом BCO.

Во всех разновидностях пчелиного алгоритма решение оптимизационной задачи рассматривается как источник нектара. Очевидно, что количество нектара в источнике пропорционально качеству представляющего его решения. В случае минимизационной задачи полезность источника нектара обратно пропорциональна значению целевой функции решения.

В алгоритме BCOi популяция агентов (искусственных пчел) состоит из  $B$  особей, совместно ищущих оптимальное решение. Каждая пчела ответственна за одно решение задачи. Имеются две чередующиеся фазы поиска: прямой проход (forward pass) и обратный проход (backward pass), которые вместе составляют один шаг алгоритма BCO [10–13].

В процессе прямого прохода каждая пчела исследует пространство поиска. Исследование подразумевает заданное число шагов по улучшению решения. Поиск осуществляется в окрестности текущего решения. В результате исследования получается новое решение. Отметим, что для расширения пространства поиска авторы работ [10–13] предлагают не выбирать лучшее решение в окрестности, а производить так называемую пропорциональную селекцию (другое название – метод рулетки). Данный метод для задач минимизации может быть кратко описан следующим образом: «Если значение целевой функции  $i$ -го решения в  $k$  раз меньше значения целевой функции  $j$ -го решения, то вероятность выбора  $i$ -го решения в качестве нового решения должна быть в  $k$  раз больше, чем аналогичная вероятность для  $j$ -го решения».

После получения новых решений начинается второй этап – обратный проход. В процессе обратного прохода все искусственные пчелы обмениваются информацией о качестве «своих» решений, то есть о вычисленных значениях целевой функции. Когда обмен информацией полностью завершен, каждая пчела решает (с определенной вероятностью), стоит ли ей отказаться от своего источника нектара. Пчелы, ассоциированные с лучшими решениями, имеют лучшие шансы на то, чтобы сохранить свои текущие решения и «разрекламировать» их среди других пчел.  $R$  пчел, сохраняющих свои решения, в BCOi называются рекрутами. Если пчела отказывается от своего текущего решения, то она должна выбрать одно из решений, предлагае-

мых другими пчелами. Очевидно, что у лучших решений наибольшая вероятность быть выбранными для дальнейшего исследования.

Две фазы поискового процесса, прямой и обратный проход, чередуются  $NC$  раз, то есть до тех пор, пока каждая пчела не произведет  $NC$  модификаций своего решения. Когда все шаги завершены, определяется лучшее из  $B$  имеющихся решений (оно используется для обновления глобального лучшего решения). После этого итерация считается завершенной. Все  $B$  решений удаляются, и стартует новая итерация. Критерием останова алгоритма может выступать, например, максимальное число итераций или суммарное время работы алгоритма.

Вероятность того, что  $b$ -я пчела перед началом очередного прямого прохода останется лояльной своему текущему решению (то есть не откажется от него), выражается формулой

$$p_b^{u+1} = \exp\left(-\frac{O_{\max} - O_b}{u}\right), \quad (5)$$

где  $O_b$  – нормализованное значение целевой функции решения, созданного  $b$ -й пчелой;  $O_{\max}$  – максимальное среди всех значений  $O_b$  ( $b = 1, 2, \dots, B$ );  $u$  – порядковый номер прямого прохода в пределах данной итерации ( $u = 1, 2, \dots, NC$ ).

Нормализованные значения целевой функции рассчитываются следующим образом. Пусть  $C_b$  – значение целевой функции для  $b$ -й пчелы,  $C_{\max}$  и  $C_{\min}$  – соответственно максимальное и минимальное значения целевой функции в пределах улья. Тогда для задачи минимизации [10]

$$O_b = \frac{C_{\max} - C_b}{C_{\max} - C_{\min}}, \quad b = 1, 2, \dots, B. \quad (6)$$

Для каждой пчелы, отказавшейся от своего решения (их  $(B-R)$  штук), необходимо найти рекрута, то есть выбрать одно из  $R$  решений, к которым их пчелы остались лояльными. Вероятность того, что решение данного конкретного рекрута будет выбрано любой из пчел, отказавшихся от своего решения, выражается формулой

$$p_b = \frac{O_b}{\sum_{k=1}^R O_k}, \quad b = 1, 2, \dots, R. \quad (7)$$

В работе [10] рассматривается концепция использования глобальной памяти в алгоритме BCO. Предлагается в конце каждой итерации не удалять все  $B$  решений, а лучшее из них делать стартовым для одной из пчел на следующей итерации. Такой подход иначе называется концепцией элитаризма. Для удобства обозначим как *use\_global* булеву переменную, показывающую, используется ли концепция элитаризма.

### Окрестность решения

Понятие «окрестность» является ключевым для метаэвристических алгоритмов, использующих

концепцию локального поиска. Для каждой конкретной оптимизационной задачи оно имеет свою специфику. В случае ЗРБС решение представляет собой множество БС с подключенными к ним клиентами. Задача построения окрестности состоит в определении множества решений, являющихся в некотором роде близкими к текущему решению.

Введем операцию «переподключение одного клиента» (имеется в виду подключение клиента к другой БС). Если мы ищем новую БС для  $i$ -го клиента, то мы должны последовательно, проверяя выполнение ограничений (1)–(3), пытаться подключить его к одному из мест-кандидатов (кроме его нынешнего места), начиная с самого ближнего к нему. Место  $w$ , к которому нужно подключить клиента, должно иметь активную БС ( $\mathbf{Sol}[w].type \neq 0$ ).

Предлагается формировать окрестность решения  $N(\mathbf{Sol})$  путем небольших изменений в текущем решении  $\mathbf{Sol}$ . Новое решение из окрестности текущего решения можно получить при помощи двух способов (операций).

1. *Смена типа одной БС на более дешевый*. Данная операция ( $\mathbf{Sol}[s].type = \mathbf{Sol}[s].type - 1$ ) возможна для каждого элемента вектора  $\mathbf{Sol}$ , для которого справедливо  $\mathbf{Sol}[s].type > 1$ . После смены типа БС мы должны проверить решение  $\mathbf{Sol}$  на выполнение ограничений (1)–(3).

2. *Удаление одной БС*. Подобная операция возможна для каждого места-кандидата с установленной БС. Для каждого клиента удаляемой станции  $s$  мы запускаем операцию «переподключение одного клиента». Удаление  $s$ -й станции возможно, если все клиенты  $s$ -й станции можно подключить к другим БС.

Очевидно, что окрестность, получаемая при помощи применения к решению  $\mathbf{Sol}$  операций «удаление одной БС» и «смена типа одной БС на более дешевый», содержит только решения с лучшей целевой функцией, чем у  $\mathbf{Sol}$ .

### Пчелиный алгоритм решения ЗРБС

Представим общую схему алгоритма BCOi для ЗРБС в виде псевдокода.

1. Введем обозначения: **Best** – лучшее решение задачи; *cost\_best* – целевая функция **Best**.
2. **Best** =  $\emptyset$ . *cost\_best* =  $\infty$ .
3. **If** (*use\_global*  $\neq$  *true*) **or** (**Best** =  $\emptyset$ ) **then**  
     **For**  $b = 1$  **to**  $B$   
         Построить начальное решение **Curr<sub>b</sub>**.  
     **End for**  
   **Else**  
     **Curr<sub>1</sub>** = **Best**  
     **For**  $b = 2$  **to**  $B$   
         Построить начальное решение **Curr<sub>b</sub>**.  
     **End for**  
   **End if**
4. **For**  $s = 1$  **to**  $NC$   
     **For**  $b = 1$  **to**  $B$  // *прямой проход*

1) Построим  $N(\mathbf{Curr}_b)$  окрестность решения  $\mathbf{Curr}_b$ .

2)  $W_b = \emptyset$ .

3) Цикл. Для каждого решения  $\mathbf{Sol}$  из построенной окрестности выполним:

3.1) Если решение  $\mathbf{Sol}$  недопустимое [не удовлетворяет одному из ограничений (1)–(3)], то переход к следующей итерации цикла.

3.2) Если  $F(\mathbf{Sol}) \geq F(\mathbf{Curr}_b)$ , то переход к следующей итерации цикла.

3.3) Добавить  $\mathbf{Sol}$  в множество  $W_b$ .

4) Если  $W_b \neq \emptyset$ , выберем одно решение из  $W_b$  при помощи метода рулетки и присвоим его  $\mathbf{Curr}_b$ .

5) Если  $F(\mathbf{Curr}_b) < cost\_best$ , то  $\mathbf{Best} = \mathbf{Curr}_b$ ,  $cost\_best = F(\mathbf{Curr}_b)$ .

**End for**

// обратный проход

$R = 0$ .

$G_{non\_loyal} = \emptyset$ . // множество пчел, которые отказались от своих текущих решений

$G_{loyal} = \emptyset$ . // множество пчел, которые не отказались от своих текущих решений

**If** ( $s \neq NC$ ) **then**

Рассчитать нормализованные значения целевой функции для каждой пчелы по формуле (6).

**For**  $b = 1$  **to**  $B$

1) Найти  $p_b^{u+1}$  по формуле (5).

2) Пусть  $z$  – случайное число из диапазона  $[0; 1]$ . Если  $z > p_b^{u+1}$ , добавим пчелу  $b$  в множество  $G_{non\_loyal}$ , иначе  $R++$  и добавим пчелу  $b$  в множество  $G_{loyal}$ .

**End for**

**For**  $b = 1$  **to**  $B$

**If** ( $b \in G_{non\_loyal}$ ) **then**

При помощи метода рулетки присвоим пчеле  $b$  решение одной из пчел из множества  $G_{loyal}$  согласно формуле (7).

**End if**

**End for**

**End if**

**End for**

5. Обнулیم все решения  $\mathbf{Curr}_b$  ( $b = 1, 2, \dots, B$ ).

6. Если время работы алгоритма ( $time$ ) истекло, то переход к п. 7, иначе – к п. 3.

7. Вернем  $\mathbf{Best}$  в качестве решения задачи.

### Компьютерное моделирование

Разработанный алгоритм был реализован в среде программирования Embarcadero Delphi 2010. С помощью данного ПО проведен ряд вычислительных экспериментов по нахождению оптимального решения ЗРБС. Моделирование проводилось на компьютере со следующими параметрами: процессор Intel Core i5-3470, 6 ГБ оперативной памяти.

Первая серия вычислительных экспериментов была посвящена исследованию быстрейшего и

точности предложенного пчелиного алгоритма ВСО путем сравнения его с методом *полного перебора* (ПП). Принцип ПП состоит в следующем: перебрать все возможные решения задачи, отбросить недопустимые, а среди оставшихся выбрать решение с минимальным значением целевой функции. Так как метод ПП не позволяет решать задачу за полиномиальное время, сравнение проводилось на задачах малой и средней размерности.

Зафиксируем параметры пчелиного алгоритма:  $B=5$ ,  $NC=10$ ,  $use\_global=true$ . Результаты работы алгоритмов приведены в таблице 1. Каждая ячейка таблицы содержит две строки: верхняя – время работы алгоритма в секундах при решении задачи определенной размерности, нижняя – значение соответствующей целевой функции в условных единицах стоимости. Время решения задачи для метода ВСО – среднее за 10 запусков алгоритма, значение целевой функции для ВСО – лучшее за 10 запусков алгоритма. Представленные в таблице 1 данные свидетельствуют о том, что на задачах малой размерности пчелиный алгоритм обеспечивает получение точных значений целевой функции, как и метод полного перебора, дающий точное решение при каждом запуске алгоритма. При этом предложенный пчелиный алгоритм справляется с решением ЗРБС за очень малое время.

Таблица 1

Сравнение пчелиного алгоритма и метода ПП

Table 1

Comparison of BCOi and exhaustive search

Размер задачи ( $N_p \times N_{ps} \times N_{types}$ )	Метод	
	Полный перебор	Пчелиный алгоритм
3×5×2	0,030 сек.	0,093 сек.
	27382 у.е.	27382 у.е.
3×7×2	0,710 сек.	0,091 сек.
	26784 у.е.	26784 у.е.
3×10×2	65,19 сек.	0,109 сек.
	26516 у.е.	26516 у.е.
5×5×2	0,606 сек.	0,101 сек.
	26776 у.е.	26776 у.е.
5×7×2	31,59 сек.	0,093 сек.
	26570 у.е.	26570 у.е.
5×10×2	5140,7 сек.	0,109 сек.
	26776 у.е.	26776 у.е.
7×5×2	14,65 сек.	0,097 сек.
	26458 у.е.	26458 у.е.
7×7×2	1390 сек.	0,107 сек.
	26396 у.е.	26396 у.е.
7×10×2	455798 сек.	0,112 сек.
	26230 у.е.	26230 у.е.

Далее был проведен ряд вычислительных экспериментов с целью выявления влияния использования глобальной памяти (концепции элитаризма) в методе ВСО на качество получаемых решений за-

дач средней и большой размерности. Зафиксированные параметры пчелиного алгоритма:  $B=5$ ,  $NC=10$ . Моделирование производилось следующим образом: каждая из задач решалась двумя способами ( $use\_global=true$  и  $use\_global=false$ ), нахождение решения отводилось одинаковое время (30 сек.), значения по итогам 100 запусков усреднялись. Результаты эксперимента приведены в таблице 2. Каждая ячейка таблицы содержит две строки: верхняя – время работы алгоритма (при данных значениях управляющих параметров) в секундах при решении задачи определенной размерности, нижняя – соответствующее значение оценки относительной погрешности решений в процентах. Относительная погрешность решений вычисляется как среднее отклонение полученных решений от лучшего из известных решений задачи. Из таблицы 2 видно, что алгоритм, подразумевающий использование элитаризма, решает ЗРБС лучше алгоритма без использования элитаризма.

Таблица 2

**Влияние использования глобальной памяти на решение задачи**

Table 2

**Impact of the global memory use on problem solving**

Размер задачи ( $N_{ip} \times N_{ps} \times N_{types}$ )	Тип ВСОi	
	С использованием памяти	Без использования памяти
50×50×3	30,112 сек.	30,109 сек.
	4,177 %	6,254 %
100×100×3	30,087 сек.	30,090 сек.
	4,521 %	6,387 %
150×150×3	30,098 сек.	30,101 сек.
	4,883 %	7,197 %
200×200×3	30,099 сек.	30,102 сек.
	5,691 %	7,366 %
500×500×3	30,092 сек.	30,103 сек.
	4,922 %	6,673 %

Таблица 3

**Влияние числа пчел и числа проходов на решение задачи**

Table 3

**Impact of the number of bees and the number of passes on problem solving**

B	NC			
	5	10	15	20
3	30,092 сек.	30,095 сек.	30,110 сек.	30,093 сек.
	5,785 %	5,123 %	6,007 %	5,997 %
5	30,110 сек.	30,087 сек.	30,096 сек.	30,098 сек.
	4,877 %	4,521 %	4,934 %	5,004 %
7	30,094 сек.	30,112 сек.	30,097 сек.	30,111 сек.
	5,376 %	4,793 %	5,446 %	6,106 %
10	30,089 сек.	30,098 сек.	30,105 сек.	30,107 сек.
	6,902 %	4,895 %	6,377 %	5,977 %
12	30,101 сек.	30,092 сек.	30,101 сек.	30,107 сек.
	5,854 %	4,992 %	6,326 %	6,542 %

Также был проведен ряд вычислительных экспериментов с целью выявления влияния числа особей в рое и числа проходов в пределах одной итерации на качество получаемых решений задач средней и большой размерности. В качестве тестового примера была выбрана задача следующей размерности:  $N_{ip}=100$ ,  $N_{ps}=100$ ,  $N_{types}=3$ . Зафиксированные параметры пчелиного алгоритма:  $use\_global=true$ . Результаты эксперимента приведены в таблице 3. Каждая ячейка таблицы содержит две строки: верхняя – время работы алгоритма ВСОi в секундах при определенных значениях  $B$  и  $NC$ , нижняя – соответствующее значение оценки относительной погрешности решений в процентах. Как и в предыдущем эксперименте, данные для каждой пары  $B$  и  $NC$  усреднялись по результатам 100 запусков алгоритма, один запуск подразумевал 30 секунд на нахождение решения. Из таблицы 3 видно, что наилучшие результаты получаются при  $B=5$ ,  $NC=10$ , что согласуется с данными, приводимыми в работе [13].

Проведенные вычислительные эксперименты позволяют сделать следующие выводы.

Предложенная модификация пчелиного алгоритма ВСО позволяет находить решение ЗРБС.

Для задач малой и средней размерности результаты пчелиного алгоритма совпадают с результатами алгоритма полного перебора. При этом предложенный метод решает задачу на много порядков быстрее.

Разработанное на основе созданного алгоритма ПО может находить решение задач средней и большой размерности за время, не превышающее 1 минуты.

Рекомендуемыми параметрами пчелиного алгоритма ВСО, позволяющими находить хорошее решение за приемлемое время, являются следующие:  $B=5$ ,  $NC=10$ ,  $use\_global=true$ .

### Литература

1. Вишневецкий В.М. Теоретические основы проектирования компьютерных сетей. М.: Техносфера, 2003. 512 с.
2. Ермолаев С.Ю. Оптимальное размещение базовых станций // Telecommunication Sciences. 2010. № 1. С. 349–355.
3. St-Hilaire M., Chinnnek J.W., Chamberland S., Pierre S. Efficient solution of the 3G Network planning problem. Computers & Industrial Engineering. 2012, no. 4, pp. 819–830.
4. Скаков Е.С., Малыш В.Н. Использование алгоритмов мултистарта и поиска с запретами для решения задачи размещения базовых станций // Информационно-управляющие системы. 2015. № 3. С. 99–106.
5. Архангельский А.Я. Программирование в Delphi 7. М.: Бином-Пресс, 2003. 1152 с.
6. Koskie S., Gajic Z. A nash game algorithm for SIR-based power control in 3G wireless CDMA networks. IEEE/ACM Transactions on Networking. 2005, no. 5, pp. 1017–1026.
7. Karaboga D., Basturk B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) Algorithm. Journ. of Global Optimization. 2007, no. 3, pp. 459–471.
8. Karaboga D., Akay B. A survey: algorithms simulating bee swarm intelligence. Artificial Intelligence Review. 2009, vol. 31, pp. 61–85.
9. Bonabeau E., Dorigo M., Theraulaz G. Swarm intelligence:

from natural to artificial systems. NY, Oxford Univ. Press, 1999, 320 p.

10. Davidovic T., Teodorovic D., Selmic M. Bee colony optimization, Part I: The algorithm overview. *Yugoslav Journ. of Operations Research*. 2015, no. 1, pp. 33–56.

11. Lucic P., Teodorovic D. Computing with bees: attacking complex transportation engineering problems. *Intern. Journ. on Artificial Intelligence Tools*. 2003, no. 3, pp. 375–394.

12. Teodorovic D., Selmic M., Davidovic T. Bee colony optimization. Part II: The application survey. *Yugoslav Journ. of Operations Research*. 2015, no. 2, pp. 185–219.

13. Davidovic T., Ramljak D., Selmic M., Teodorovic D. Bee colony optimization for the p-center problem. *Computers & Operations Research*. 2011, no. 10, pp. 1367–1376.

DOI: 10.15827/0236-235X.115.067-073

Received 19.01.16

## BEE COLONY OPTIMIZATION ALGORITHM FOR WIRELESS NETWORK PLANNING PROBLEM SOLVING

<sup>1</sup>Skakov E.S., Postgraduate Student, wallkiry@mail.ru

<sup>2</sup>Malysh V.N., Dr.Sc. (Engineering), Professor, vmalysh@mail.ru

<sup>1</sup>Lipetsk State Pedagogical University, Lenina St. 42, Lipetsk, 398020, Russian Federation

<sup>2</sup>Lipetsk Branch of the Russian Presidential Academy of National Economy and Public Administration under the President of the Russian Federation, Internatsional'naya St. 3, Lipetsk, 398050, Russian Federation

**Abstract.** The article is devoted to solving a wireless network planning task. The authors defined the problem of base stations' location, taking into account the level of attenuation of signal during propagation, the level of inter-cell interference and several types of base stations. A problem solution is represented as a vector of structures, each structure stores data about one potential site (an installed base station type, a list of connected clients).

A modification of one of metaheuristic methods (bee colony optimization algorithm) has been developed. The bee colony optimization method is a multi-agent method based on simulation of intelligent behavior of agent colonies (the so-called swarm intelligence methods). Naturally, some groups of social insects (for example, colonies of ants, bees, termites) have such intellect. The proposed algorithm is based on the concept of current solution neighborhood. Little changes in current wireless network configuration are used to form the neighborhood. To obtain a new solution from the neighborhood of the current solution we can use one of the following two operations: changing the type of a station to more cheap or removing a base station.

The developed algorithm is implemented in Delphi programming environment. The article shows that the new algorithm helps to find the solution for the base station location problem in a reasonable time, much faster than the exact method of exhaustive search. The authors found out how the quality of the problem solution obtained by bee colony optimization algorithm depends on values of control parameters.

**Keywords:** base station location, bee colony optimization, wireless networks, optimization, SIR.

## References

1. Vishnevsky V.M. *Teoreticheskie osnovy proektirovaniya kompyuternykh setey* [Theoretical Basics of Computer Networks Design]. Moscow, Tekhnosfera Publ., 2003, 512 p. (in Russ.).
2. Ermolaev S.Yu. Optimal Base Stations' Location. *Telecommunication Sciences*. 2010, no. 1, pp. 349–355 (in Russ.).
3. St-Hilaire M., Chinnek J.W., Chamberland S., Pierre S. Efficient Solution of the 3G Network Planning Problem. *Computers & Industrial Engineering*. 2012, no. 4, pp. 819–830.
4. Skakov E.S., Malysh V.N. Multi-Start and Tabu Search Algorithms for Base Station Location Problem. *Informatsionno-upravlyayushchie sistemy* [Information and Control Systems]. 2015, no. 3, pp. 99–106 (in Russ.).
5. Arkhangelsky A.Ya. *Programmirovaniye v Delphi 7* [Programming in Delphi 7]. Moscow, Binom-Press Publ., 2003, 1152 p. (in Russ.).
6. Koskie S., Gajic Z. A Nash Game Algorithm for SIR-based Power Control in 3G Wireless CDMA Networks. *IEEE/ACM Transactions on Networking*. 2005, no. 5, pp. 1017–1026.
7. Karaboga D., Basturk B. A Powerful and Efficient Algorithm for Numerical Function Optimization: Artificial Bee Colony (ABC) Algorithm. *Journ. of global optimization*. 2007, no. 3, pp. 459–471.
8. Karaboga D., Akay B. A Survey: Algorithms Simulating Bee Swarm Intelligence. *Artificial Intelligence Review*. 2009, vol. 31, pp. 61–85.
9. Bonabeau E., Dorigo M., Theraulaz G. *Swarm Intelligence: from Natural to Artificial Systems*. New York, Oxford Univ. Press, 1999, 320 p.
10. Davidovic T., Teodorovic D., Selmic M. Bee Colony Optimization Part I: The Algorithm Overview. *Yugoslav Journal of Operations Research*. 2015, no. 1, pp. 33–56.
11. Lucic P., Teodorovic D. Computing with Bees: Attacking Complex Transportation Engineering Problems. *International Journal on Artificial Intelligence Tools*. 2003, no. 3, pp. 375–394.
12. Teodorovic D., Selmic M., Davidovic T. Bee Colony Optimization Part II: The Application Survey. *Yugoslav Journal of Operations Research*. 2015, no. 2, pp. 185–219.
13. Davidovic T., Ramljak D., Selmic M., Teodorovic D. Bee Colony Optimization for the p-center Problem. *Computers & Operations Research*. 2011, no. 10, pp. 1367–1376.

УДК 004.912

Дата подачи статьи: 09.03.16

DOI: 10.15827/0236-235X.115.074-078

## **МЕТОД АВТОМАТИЧЕСКОГО ФОРМИРОВАНИЯ СЕМАНТИЧЕСКОЙ СЕТИ ИЗ СЛАБОСТРУКТУРИРОВАННЫХ ИСТОЧНИКОВ**

*(Работа выполнена в рамках темы НИР № 615869 «Методы проектирования ключевых систем информационной инфраструктуры»)*

*А.Е. Письмак, магистрант, alexey.pismak@cs.ifmo.ru;*

*А.Е. Харитонова, тьютор, nasty@tune-it.ru;*

*Е.А. Цопа, ассистент, evgenij.tsopa@cs.ifmo.ru;*

*С.В. Клименков, ассистент, serge.klimenkov@cs.ifmo.ru*

*(Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Кронверкский просп., 49, г. Санкт-Петербург, 197101, Россия)*

Автоматическая обработка текста – одно из основных научно-практических направлений в современных информационных технологиях. Важным предметом изучения в этой области остается возможность автоматического построения и обновления тезаурусов и семантических сетей. Семантическая сеть – это ориентированный граф; узлами его являются понятия, а ребрами – отношения между ними. Автоматическое построение сети требует наличия внешнего источника, из которого будут импортироваться узлы сети и на основе которого будут формироваться связи между ними. В качестве такого источника решено использовать внешний открытый словарь Wiktionary, формируемый сообществом пользователей сети Интернет. Внесение данных в семантическую сеть из вышеупомянутого источника может значительно повысить связность такой сети, однако, чтобы использовать этот подход, необходимо устранить структурный недостаток источника, состоящий в том, что разделы Wiktionary зачастую имеют ошибки в уровне вложенности. В ходе работы были исследованы существующие ошибки в сформированных словарных статьях и предложен метод их разрешения. Метод основывается на механизмах конечных автоматов, где выходной сигнал автомата – правильный уровень вложенности текущего раздела. На базе данного метода был разработан алгоритм, который стал основой программного модуля, осуществляющего автоматическую корректировку структуры словарных статей Wiktionary в процессе их импорта в семантическую сеть. Тестирование показало, что разработанный модуль обеспечивает им производительность и погрешность, достаточные для его использования в качестве составной части системы семантического анализа текста на естественном языке.

**Ключевые слова:** *семантические сети, автоматическая обработка текста, Wiktionary, словарные источники, конечные автоматы.*

Поверхностный семантический анализ – одно из наиболее активно развивающихся в настоящее время направлений в области автоматической обработки текста [1]. Практически всегда эта форма анализа текста предполагает использование некоторой базы знаний о реальных объектах и абстрактных понятиях. Такие базы играют ключевую роль в процессе анализа текста, так как содержат информацию о понятиях, отображенных на множества грамматик (алфавитов) различных языков. Подобные структуры знаний принято называть семантическими сетями.

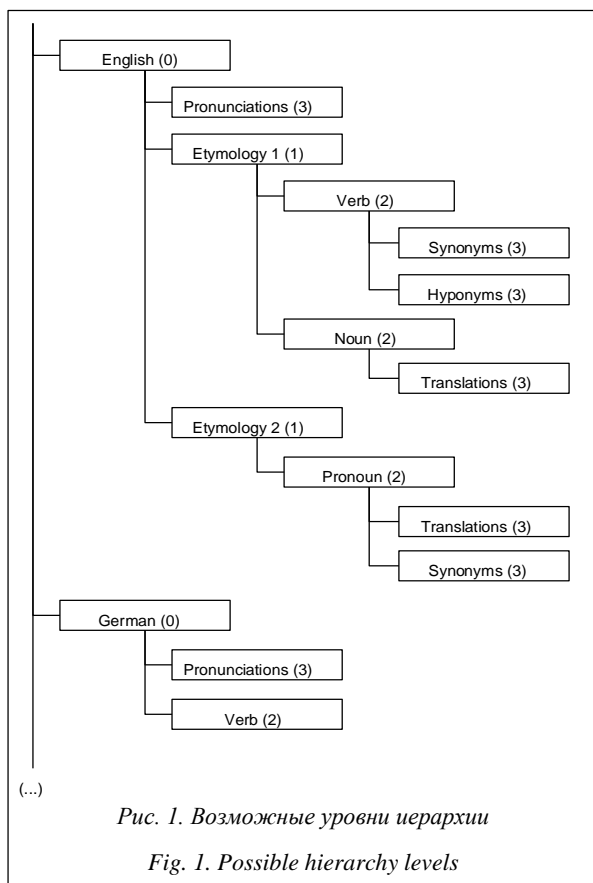
Семантическая сеть – это ориентированный граф, узлами которого являются понятия (или смысловые значения), а ребрами – отношения между ними. Такие сети можно использовать при решении задачи семантического анализа текста, так как они позволяют абстрагироваться от синтаксиса и рассматривать весь текст, составляющие его предложения, фразы и слова в виде набора смысловых значений [2]. При этом использование семантических сетей для решения задачи семантического анализа текстов приводит к появлению ряда дополнительных требований к ним. Основное требование в том, чтобы используемая семантическая сеть содержала достаточно большое количество взаимосвязанных узлов, а данные в этих узлах (зна-

чения слов и связи между ними) оперативно обновлялись (то есть база должна постоянно актуализироваться) [3]. Только в этом случае возможен эффективный семантический анализ текста.

Вследствие этих требований ряд популярных семантических сетей (например WordNet и BabelNet) становятся малоприменимыми для использования в качестве основы для построения системы семантического анализа [4], так как в них слишком мало узлов и данные либо вообще не актуализируются, либо актуализируются вручную администраторами системы. Таким образом, появляется необходимость формирования структуры знаний, хранящихся в семантической сети, из каких-либо внешних источников данных, обладающих достаточной полнотой и актуальностью. Одним из таких возможных источников данных является открытый словарь Wiktionary (викисловарь, проект Wikimedia Foundation [5]). Данные в него вносят обычные пользователи сети Интернет, англоязычный раздел этого словаря содержит в настоящий момент более 4 миллионов статей [6].

### **Постановка задачи**

Словарь Wiktionary представляет собой набор страниц, каждая из которых описывает одно или



несколько смысловых значений. Эти значения и их свойства в соответствии с рекомендациями по оформлению статей [7] должны быть расположены согласно некоторой иерархии, возможные уровни которой представлены на рисунке 1, где цифрами обозначены следующие секции: «Язык» – 0, «Этимология» – 1, «Часть речи» – 2, «Грамматические и синтаксические характеристики смыслового значения» (синонимы, произношение, гиперонимы, гипонимы, антонимы и так далее) – 3.

Каждая из секций может содержать другие секции (согласно определенным правилам [8]) либо непосредственно сами данные. Таким образом, если статья оформлена в соответствии с рекомендациями сообщества, иерархия позволяет однозначно выделить смысловые значения и свойства, относящиеся к найденным значениям.

Однако, как показали проведенные авторами эксперименты, иерархия более чем 80 % страниц Wiktionary не соответствует этим рекомендациям. Таким образом, корректный импорт статей Wiktionary в семантическую сеть возможен только после того, как структура секций в них будет изменена в соответствии с рекомендациями сообщества.

### Существующие подходы

Можно выделить следующие два подхода к решению данной задачи:

1. метод, основанный на правилах (rule based [9]); строится модель, представляющая собой детерминированный конечный автомат, переключение состояний которого осуществляется по заданным правилам, учитывающим текущее состояние автомата и входные сигналы;

2. машинное обучение с учителем (supervised machine learning [10]); системе необходимо предоставить эталонные страницы словаря, с помощью которых будут осуществляться обучение системы и дальнейшая обработка слабоструктурированных данных.

Сравнительный анализ данных подходов выявил следующее.

Метод, основанный на правилах, прост в реализации, не требует дополнительного времени на обучение системы и имеет вычислительную сложность  $O(1)$  [11]. Он имеет малую погрешность в заранее известных и изученных предметных областях. Метод машинного обучения с учителем может оказаться более точным при недостаточно подробном описании предметной области [12].

В результате для решения поставленной задачи был выбран метод, основанный на правилах, так как в данном случае предметная область детально описана, вследствие чего погрешность будет относительно невелика.

### Разработка алгоритма

Результатом формализации набора правил стал алгоритм, суть которого сводится к тому, что из исходного текста страницы Wiktionary формируется древовидная структура, содержащая секции страницы в логически правильном порядке. На рисунке 1 видно, что алгоритм производит перебор секций и в один проход осуществляет их перестроение в нужную структуру.

Алгоритм перебора секций и формирования для каждой секции ее уровня вложенности сводится к автомату Мили:  $A = \{S, S_0, X, Y, \lambda, \delta\}$ , где состояние автомата  $S_i$  характеризуется типом предыдущей секции, типом родительской секции и признаком наличия в тексте секций этимологии; набор входных символов  $X$  является множеством возможных типов секций; набор выходных символов  $Y$  – множеством коэффициентов, показывающих уровень вложенности текущей секции.

Так как всего типов секций 4, теоретических вариантов перехода от предыдущей к следующей получается 16, а с учетом флага этимологии – 32. Но фактически некоторые переходы невозможны по целому ряду причин. Например, невозможен переход от секции языка к секции языка (не несет смысловой информации). В связи с этим была проведена декомпозиция вариантов переходов. Все теоретические переходы показаны в таблице 1. В первой строке указан флаг наличия в тексте секции «Этимология». Вторая строка содержит обозначения те-

кущей обрабатываемой секции, третья – обозначения предыдущей секции: *L* (language), *E* (etymology), *P* (part of speech), *O* (option).

Фактически невозможны либо бессмысленны следующие переходы:

- семь случаев переходов, при которых отсутствует флаг этимологии, а в строках переходов есть

соответствующие обозначения (выделены в таблице жирным шрифтом и подчеркиванием);

- три случая переходов lang-lang и ety-ety (то есть секции, не содержащие необходимую информацию);

- один случай перехода lang-pos при наличии этимологии.

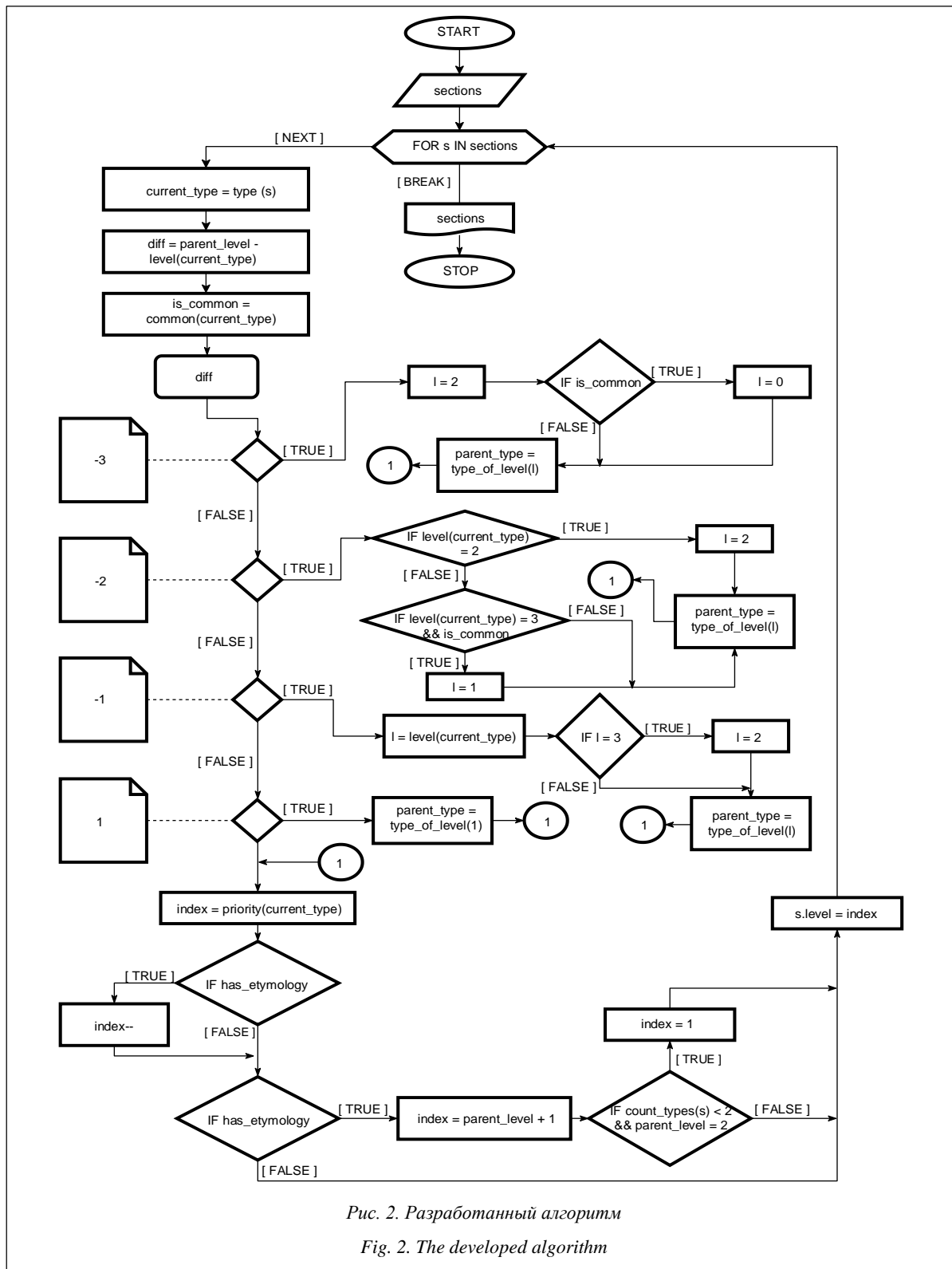




Таблица 1  
Таблица возможных переходов  
Table 1  
A table of possible transitions

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	E	P	O	L	E	P	O	L	E	P	O	L	E	P	O
L	L	L	L	E	E	E	E	P	P	P	P	O	O	O	O

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
L	E	P	O	L	E	P	O	L	E	P	O	L	E	P	O
L	L	L	L	E	E	E	E	P	P	P	P	O	O	O	O

После исключения этих переходов получим таблицу фактических переходов (табл. 2).

Таблица 2  
Таблица фактических переходов  
Table 2  
A table of actual transitions

1	0	1/0	1	1	1	1/0	1	1/0	1/0	1/0	1	1/0	1/0
E	P	O	L	P	O	L	E	P	O	L	E	P	O
L	L	L	E	E	E	P	P	P	P	O	O	O	O

Также стоит обратить внимание на переходы lang-option и ety-option. Дело в том, что в этих случаях допускается вложенность в родительскую секцию семантических и грамматических характеристик, но эти характеристики должны быть из следующего списка: Alternative form, Pronunciation, Anagram, Quotation.

С учетом всех указанных особенностей был разработан алгоритм, блок-схема которого показана на рисунке 2.

Выходным сигналом спроектированного конечного автомата является уровень вложенности для текущей обрабатываемой секции, входным – очередная обрабатываемая секция со всем необходимым набором параметров.

Состояние конечного автомата характеризуют следующие величины:

- current\_type – тип секции, обрабатываемой в данный момент;
- diff – коэффициент перепада, фиксирующий разницу между предыдущим значением уровня вложенности и текущим;
- is\_common – флаг, обозначающий принадлежность текущей секции одному из следующих типов: Alternative form, Pronunciation, Anagram, Quotation;
- parent\_type – тип секции, являющейся родительской по отношению к текущей обрабатываемой секции;
- has\_etymology – флаг наличия в статье секции «Этимология».

### Результаты экспериментальных исследований

Результатом практической реализации предложенного алгоритма стало приложение на языке

Java, осуществляющее преобразование статей Wiktionary в процессе их импорта в семантическую сеть [13]. Пример перестроения структуры словарной статьи с помощью разработанного модуля приведен на рисунке 3.

В данном случае была перестроена секция «Etymology 4»: две секции с грамматическими и синтаксическими характеристиками смыслового значения были перемещены внутрь секции «Часть речи».

Тестирование приложения на выборке в 1 000 статей словаря показало его достаточно высокую эффективность – корректно была восстановлена структура 875 статей; таким образом, процент корректно восстановленных статей составил 87,5.

Тестирование производительности разработанного модуля также показало его высокую эффективность – восстановление одной словарной статьи занимало от 4 до 15 мс на тестовом стенде со следующими характеристиками: CPU Intel Core i7-3612QM 2.10GHz, 8 Gb RAM, OS Linux x64 без виртуализации, с реализацией алгоритма на языке программирования Java и выделенной памятью для виртуальной машины Java 2GB.

В заключение отметим, что в работе был предложен и реализован алгоритм восстановления структуры статей словаря Wiktionary. Предложенный алгоритм основан на правилах (rule-based) – такой подход позволяет добиться более высокой эффективности алгоритма. Разработанный алгоритм был применен для восстановления структуры статей словаря Wiktionary. Экспериментальное исследование показало, что разработанный алгоритм имеет приемлемую погрешность, обеспечивая



Рис. 3. Перестроение структуры словарной статьи

Fig. 3. Dictionary entry structure rearrangement

быстродействие, достаточное для его применения в системе семантического анализа текста на естественном языке.

### Литература

1. Bessmertny I. Knowledge visualization based on semantic networks. *Programming and Computer Software*, 2010, no. 7, pp. 197–204.
2. Hoffman T. Probabilistic latent semantic analysis. *Proc. 15th Conf. on Uncertainty in Artificial Intelligence (UAI'99)*, Stockholm, 1999, pp. 289–296.
3. Budanitsky A., Hirst G. Semantic distance in WordNet: An experimental, application-oriented evaluation of five measures. *Workshop on WordNet and Other Lexical Resources*, 2nd Meeting of the NAACL, Pittsburg, 2001.
4. Navigli R., Ponzetto S.P. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *AIJ*, 2012, no. 12, pp. 217–250.
5. Zesch T., Muller C., Gurevych I. Using wiktionary for computing semantic relatedness. *Proc. 23rd AAAI Conf. on Artificial Intelligence*, 2008.

6. Статистика Wiktionary. URL: <https://en.wiktionary.org/wiki/Special:Statistics>. (дата обращения: 20.02.2016).

7. Рекомендации по написанию статей для Wikipedia и Wiktionary. URL: [https://en.wikipedia.org/wiki/Wikipedia:Manual\\_of\\_Style](https://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style) (дата обращения: 20.02.2016).

8. Правила разметки статей Wikipedia и Wiktionary. URL: [https://en.wikipedia.org/wiki/Wikipedia:Manual\\_of\\_Style/Layout](https://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style/Layout) (дата обращения: 20.02.2016).

9. Brownston L., Farrell R., Kant E., Martin N. *Programming expert systems in OPS5: An introduction to rule-based programming*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1985, 490 p.

10. Vapnik V.N. *The nature of statistical learning theory*. Springer Verlag, 2000, 314 p.

11. Kowalski T.J., Levy L.S. *Rule-based programming*. Springer Science & Business Media, 2012, 306 p.

12. Mohri M., Rostamizadeh A., Talwalkar A. *Foundations of machine learning*. The MIT Press, 2012, 432 p.

13. Письмак А.Е., Харитонов А.Е., Цопа Е.А., Клименков С.В. Оценка семантической близости предложений на естественном языке методами математической статистики // Науч.-технич. вестн. ИТМО. 2016. Т. 16. № 2. С. 324–330.

DOI: 10.15827/0236-235X.115.074-078

Received 09.03.16

## METHOD OF AUTOMATIC GENERATION OF SEMANTIC NETWORK FROM SEMI-STRUCTURED SOURCES

(Acknowledgements. The work has been performed as a part of a research topic no. 615869 "Methods of information infrastructure key systems designing")

<sup>1</sup>Pismak A.E., Student, alexey.pismak@cs.ifmo.ru

<sup>1</sup>Kharitonova A.E., Tutor, nasty@tune-it.ru

<sup>1</sup>Tsopa E.A., Assistant, evgenij.tsopa@cs.ifmo.ru

<sup>1</sup>Klimenkov S.V., Assistant, serge.klimenkov@cs.ifmo.ru

<sup>1</sup>The National Research University of Information Technologies, Mechanics and Optics, Kronverksky Ave. 49, St. Petersburg, 197101, Russian Federation

**Abstract.** Natural language processing is one of the most rapidly growing areas in current IT-related research. An important task in this area is the ability to automatically build and update thesauri and semantic networks. Semantic network is a directed graph with concepts as nodes and relations between them as edges. Automatic semantic network generation requires some external dictionary nodes and relations source. It was decided to use an external source of Wiktionary dictionary for this task. Wiktionary articles, which are effectively imported into semantic network, can significantly increase this network's completeness and coherence. Wiktionary is open for editing by anyone, so there are some typical problems in Wiktionary articles markup that must be solved to effectively import it into semantic network. The main problems are errors in article sections nesting structure. The authors propose a novel approach for automatic nesting structure errors resolving. The proposed method is based on finite automata approach. The output signal of the automata is the correct level of nesting of the current section. A new Wiktionary article processing algorithm was developed based on the proposed approach and a new software module based on this algorithm was developed. Test results showed the applicability of the developed software module for using in modern complex NLP systems.

**Keywords:** semantic networks, natural language processing, Wiktionary, dictionary, thesauri, finite automata.

### References

1. Bessmertny I. Knowledge visualization based on semantic networks. *Programming and Computer Software*. 2010, no. 7, pp. 197–204.
2. Hoffman T. Probabilistic latent semantic analysis. *UAI'99 Proc. 15th Conf. on Uncertainty in Artificial Intelligence*. Stockholm, 1999, pp. 289–296.
3. Budanitsky A., Hirst G. Semantic distance in WordNet: An experimental, application-oriented evaluation of five measures. *Workshop on WordNet and Other Lexical Resources*. 2nd Meeting of the NAACL, Pittsburg, 2001.
4. Navigli R., Ponzetto S.P. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence Journ.* 2012, no. 12, pp. 217–250.
5. Zesch T., Muller C., Gurevych I. Using Wiktionary for Computing Semantic Relatedness. *Proc. 23rd AAAI Conf. on Artificial Intelligence*. 2008.
6. *Wiktionary Statistics*. Available at: <https://en.wiktionary.org/wiki/Special:Statistics> (accessed February 20, 2016).
7. *Wikipedia & Wiktionary Manual of Style*. Available at [https://en.wikipedia.org/wiki/Wikipedia:Manual\\_of\\_Style](https://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style) (accessed February 20, 2016).
8. *Wikipedia & Wiktionary Articles Layout Rules*. Available at: [https://en.wikipedia.org/wiki/Wikipedia:Manual\\_of\\_Style/Layout](https://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style/Layout) (accessed February 20, 2016).
9. Brownston L., Farrell R., Kant E., Martin N. *Programming Expert Systems in OPS5: an Introduction to Rule-based Programming*. Addison-Wesley Longman Publ., Boston, MA, 1985.
10. Vapnik V.N. *The Nature of Statistical Learning Theory*. Springer Verlag Publ., 2000.
11. Kowalski T.J., Levy L.S. *Rule-Based Programming*. Springer Science & Business Media Publ., 2012.
12. Mohri M., Rostamizadeh A., Talwalkar A. *Foundations of Machine Learning*. MIT Press, 2012.
13. Pismak A.E., Kharitonova A.E., Tsopa E.A., Klimenkov S.V. Evaluation of semantic similarity in sentences in natural language by using descriptive statistics method. *Nauch.-tekhnich. vestn. ITMO* [Scientific and Technical Journal of Information Technologies, Mechanics and Optics]. 2016, vol. 16, no. 2, pp. 324–330 (in Russ.).

УДК 004.89

Дата подачи статьи: 23.12.15

DOI: 10.15827/0236-235X.115.079-084

## **ПРОТОТИП СИСТЕМЫ ПОДДЕРЖКИ ПРОЕКТИРОВАНИЯ АГЕНТОВ ДЛЯ ИМИТАЦИОННЫХ МОДЕЛЕЙ СЛОЖНЫХ СИСТЕМ**

(Работа выполнена при частичной финансовой поддержке РФФИ,  
проекты №№ 14-07-31298, 14-07-0222, 15-07-05641)

*А.И. Павлов, к.т.н., старший научный сотрудник, asd@icc.ru;*

*А.Б. Столбов, младший научный сотрудник, stolboff@icc.ru*

*(Институт динамики систем и теории управления СО РАН,  
ул. Лермонтова, 134, г. Иркутск, 664033, Россия)*

В работе рассматривается реализация прототипа системы поддержки проектирования агентов имитационных моделей. Приведена последовательность этапов процесса разработки имитационной модели, использованная в качестве основы для построения алгоритма формирования агентной модели, – от концептуального описания исследуемой области до ее программной реализации. Рассмотрена программная реализация компонентов создания концептуальной модели, визуального конструирования баз знаний продукционного типа и реализации агентов в среде моделирования Madkit. Компонент создания концептуальной модели предназначен для формирования иерархической структуры понятий, определения свойств понятий и отношений между ними. При этом описание отношений осуществляется путем построения набора семантических сетей, вершинами которых являются понятия концептуальной модели, а связями – отношения между ними. Каждая из семантических сетей поясняет отдельный аспект моделируемой области. Компонент визуального конструирования баз знаний обеспечивает конструирование правил вида «ЕСЛИ условия ТО действия» и генерацию кода базы знаний на языках Jess и Clips. Компонент реализации агентов в среде моделирования Madkit включает в себя библиотеку среды моделирования, классы, использующиеся в процессе агентного моделирования («Типовой агент», «Планировщик» и т.д.), а также классы для выполнения служебных функций (взаимодействие с БД или с внешней системой). Основные функции данного компонента – выполнение агентной модели, а также информирование пользователя о текущем состоянии процесса моделирования и параметров отдельных агентов. Реализация всех компонентов осуществлена в виде веб-приложений.

**Ключевые слова:** *агентное имитационное моделирование, продукционная экспертная система, системы, основанные на знаниях.*

Разнообразие структур и поведения моделируемых объектов реального мира в зависимости от задач исследования обуславливает высокую сложность автоматизации процесса разработки имитационных моделей, в том числе агентных, поэтому в большинстве современных инструментальных средств моделирования данный процесс осуществляется путем написания кода на императивном языке программирования с невысоким уровнем автоматизации. В работе [1] был предложен подход к обеспечению поддержки пользователя в процессе создания агентных имитационных моделей сложных систем. Особенностью его является нацеленность на обеспечение специалистов-предметников, не имеющих существенных навыков программирования, возможностью создания имитационных моделей на основе информации из *концептуальной модели* (КМ) предметной области, а также использования формализма продукционных правил и парадигмы визуального программирования для описания поведения элементов модели (агент, ресурс, среда). Для реализации предложенного подхода была разработана концептуальная архитектура системы поддержки проектирования агентов для имитационных моделей сложных систем [2].

Целью данной работы является описание промежуточных результатов, полученных в процессе реализации архитектуры системы. В частности, приводится более подробное описание последовательности этапов разработки модели, а также рас-

смотрены разработанные к настоящему времени модули системы поддержки: компонент создания КМ, компонент визуального конструирования баз знаний продукционного типа, компонент реализации агентов в среде моделирования Madkit [3].

### **Основные этапы процесса разработки агентной модели**

Эффективность применения инструментария имитационного моделирования зависит от способности исследователя сформулировать задачу и описать способ ее решения. В рамках применяемого подхода целевым пользователем системы является специалист-предметник, не владеющий навыками программирования. Поэтому неизбежно появляются определенные ограничения на способы описания элементов модели (агент, ресурс, среда). Во-первых, формулировка задач должна осуществляться с помощью содержательной терминологии, что, в свою очередь, определяет необходимость явного применения КМ исследуемой области в процессе проектирования агентной модели. Во-вторых, для описания закономерностей исследуемой области целесообразно использование интуитивно понятных инструментов: продукционных правил и элементов визуального программирования. С учетом указанных ограничений процесс проектирования агентной модели можно представить в виде следующих основных этапов.

1. Разработка КМ исследуемой области. Результат данного этапа – КМ, включающая формальное описание понятий, их свойств, отношений между ними, а также вычислительных процедур, реализующих известные закономерности исследуемой области, которые могут быть использованы в процессе моделирования.

2. Постановка задачи имитационного моделирования. Результат – начальное описание разрабатываемой имитационной модели, включающее название, описание и перечень понятий КМ, которые в дальнейшем будут использованы в качестве агентов или ресурсов имитационной модели.

3. Проектирование среды. Результат – структура, описывающая наблюдаемые агентами параметры внешней среды, а также набор ограничений, налагаемых средой на действия агентов (взаимодействие между агентами, доступ к ресурсам и т.д.). Предполагается, что ограничения формируются в виде продукционной базы знаний.

4. Проектирование агента. Результат – описание типа архитектуры агента, сенсорной функции, способов взаимодействия с другими агентами, а также продукционная база знаний, описывающая поведение определенного класса агентов модели.

5. Спецификация модели для вычислительного эксперимента. Результатом данного блока является количество агентов каждого класса в текущем эксперименте, начальные значения свойств агентов, параметров среды и другие исходные данные для моделирования (критерии оценки результатов моделирования).

6. Программная реализация имитационной модели. Результат – программная система, позволяющая осуществлять рабочие расчеты и представлять результаты моделирования.

Последовательность определяет список основных задач, которые должна решать разрабатываемая система поддержки проектирования агентных моделей. Для реализации 1-го этапа используется компонент создания КМ. Компонент визуального конструирования баз знаний продукционного типа применяется на этапах 3 и 4, а компонент реализации агентов в среде моделирования Madkit – на этапе 6. Опишем данные компоненты более подробно.

### Компонент создания КМ

Согласно требованиям к программной системе, исходным для построения модели является описание предметной области в форме КМ. В данной работе структура этой КМ имеет следующий вид:

КМ={Понятие, Атрибут, Отношение, Спецификация внешнего метода};

Понятие={Имя, Описание, Список атрибутов};  
Атрибут={Имя, Тип атрибута, Значение по умолчанию};

Тип атрибута={Литеральный, Объектный};

Отношение={Понятие 1, Тип отношения, Имя, [Понятие 2]};

Внешний метод={Имя, Описание, Список параметров, Тип реализации};

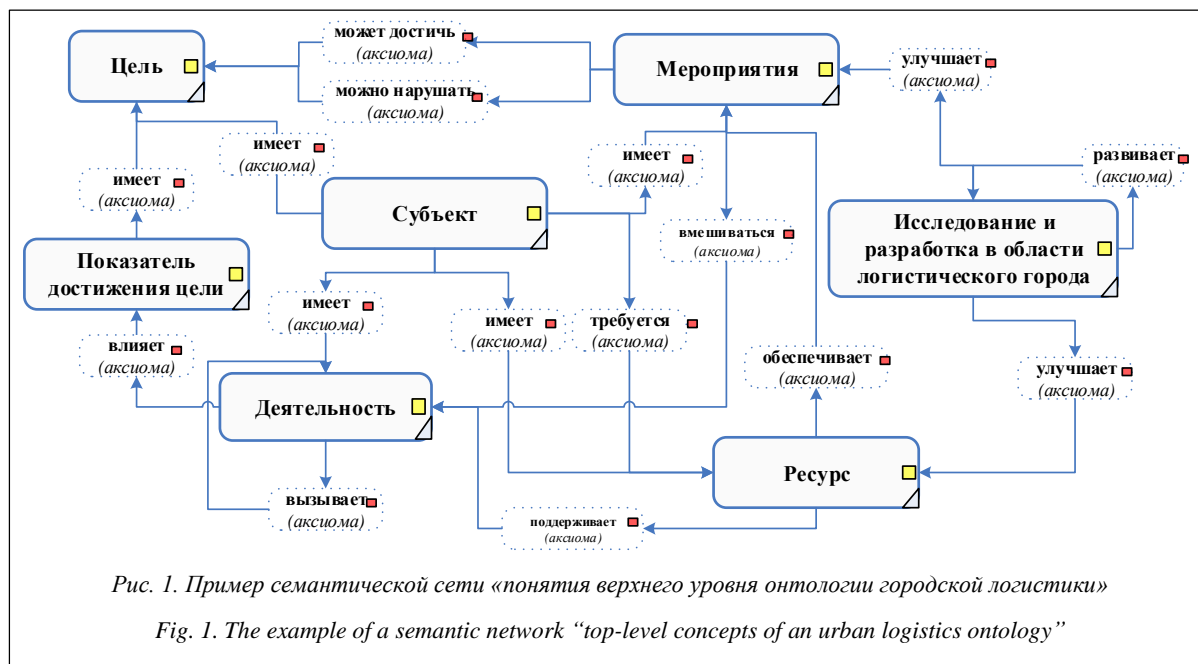
Поясним некоторые определения, связанные с предложенной структурой КМ. Под литеральным типом атрибута понимаются текст, число, дата и т.п., а под объектным – ссылка на другое понятие. Отношения задаются между понятиями и являются как бинарными, так и унарными. Типами отношений могут быть как общепринятые (причинно-следственное отношение, ассоциация, агрегация и др.), так и связанные с предметной спецификой и заданные пользователем. Например, для моделирования поведения роботов пользователь может явно ввести такой тип отношения, как «движение», со следующими возможными названиями для этого отношения: вправо, влево. Спецификация внешнего метода – это способ описания в КМ, разработанных сторонними исследователями программного кода. При этом необходимо указать спецификацию входных и выходных параметров, а также тип реализации (например, функция Jess [4], метод JAVA, скрипт PHP и т.п.).

Компонент создания КМ реализован в виде веб-приложения и обладает следующей функциональностью: создание КМ исследуемых областей, создание перечня возможных свойств понятий, формирование иерархической структуры понятий КМ от абстрактных к более конкретным, определения свойств понятий и отношений между ними. При этом описание отношений осуществляется путем построения набора семантических сетей, вершинами которых являются понятия КМ, а связями – отношения между ними. Каждая из семантических сетей поясняет отдельный аспект моделируемой области. Результат применения компонента для описания онтологии городской логистики [5] представлен на рисунке 1.

Несмотря на наличие редакторов, обладающих схожей функциональностью, например Protégé [6] или TopBraid Composer [7], было принято решение о разработке данного компонента ввиду приблизительно одинаковых трудозатрат на интеграцию модулей работы с форматами OWL или RDF и программой реализации данного компонента. Кроме того, обеспечение коллективной работы с КМ более эффективно реализуется в случае использования СУБД, а не XML-файла.

### Компонент визуального конструирования баз знаний продукционного типа

Данный компонент реализован в виде веб-приложения и обладает следующей функциональностью: создание баз знаний для выбранной КМ, формирование перечня шаблонов фактов на основе существующих понятий КМ, визуальное конструирование правил вида «ЕСЛИ условия ТО дей-



ствия», ввод начальных условий для запуска процедуры логического вывода, генерация и тестирование кода базы знаний на языках Jess и Clips [8]. Пример построения правила для описания развития деградационного процесса [9] приведен на рисунке 2.

В рамках системы предложена графическая схема отображения правил. В настоящее время при конструировании правил можно использовать следующие условия:

- наличие факта указанного шаблона;
- ограничение на значение слота факта (текст, число);
- сравнение значения слота с существующим фактом;

– сравнение со значением слота другого факта.

Для формирования правой части правила используются следующие виды действий:

- создание нового факта;
- модификация существующего факта;
- удаление факта;
- вызов внешней процедуры.

Генерация кода может быть осуществлена как для базы знаний в целом, так и для ее отдельных элементов (шаблон факта, правило, описание начальных условий). Для тестирования кода базы знаний используется специальный терминал, обеспечивающий возможность редактирования сгенерированного кода и его выполнения с помощью ин-

Рис. 2. Пример экранной формы конструирования правила

Fig. 2. The example of a screen form for rule design

терпретатора Jess. Терминал обеспечивает отображение результатов логического вывода: состояние рабочей памяти, список сработавших правил, сообщения об ошибках.

Для хранения информации баз знаний используется СУБД PostgreSQL. Серверная часть включает в себя генератор кода, реализованный на языке PHP, и блок интерпретации, обеспечивающий доступ к машине вывода. Блок интерпретации реализован в виде веб-сервиса на языке Java, который обеспечивает доступ к системе Jess.

### Компонент реализации агентов в среде моделирования Madkit

Данный компонент обеспечивает автоматическое создание программных объектов Java, которые, в свою очередь, выполняют программную реализацию агентов имитационной модели. Реализация агентов осуществляется с помощью специального базового класса (далее – типовой агент). Применение типового агента обеспечивает унификацию процесса создания агентов, а также возможность декларативного определения состояния и поведения агентов. Типовой агент обладает следующей структурой:

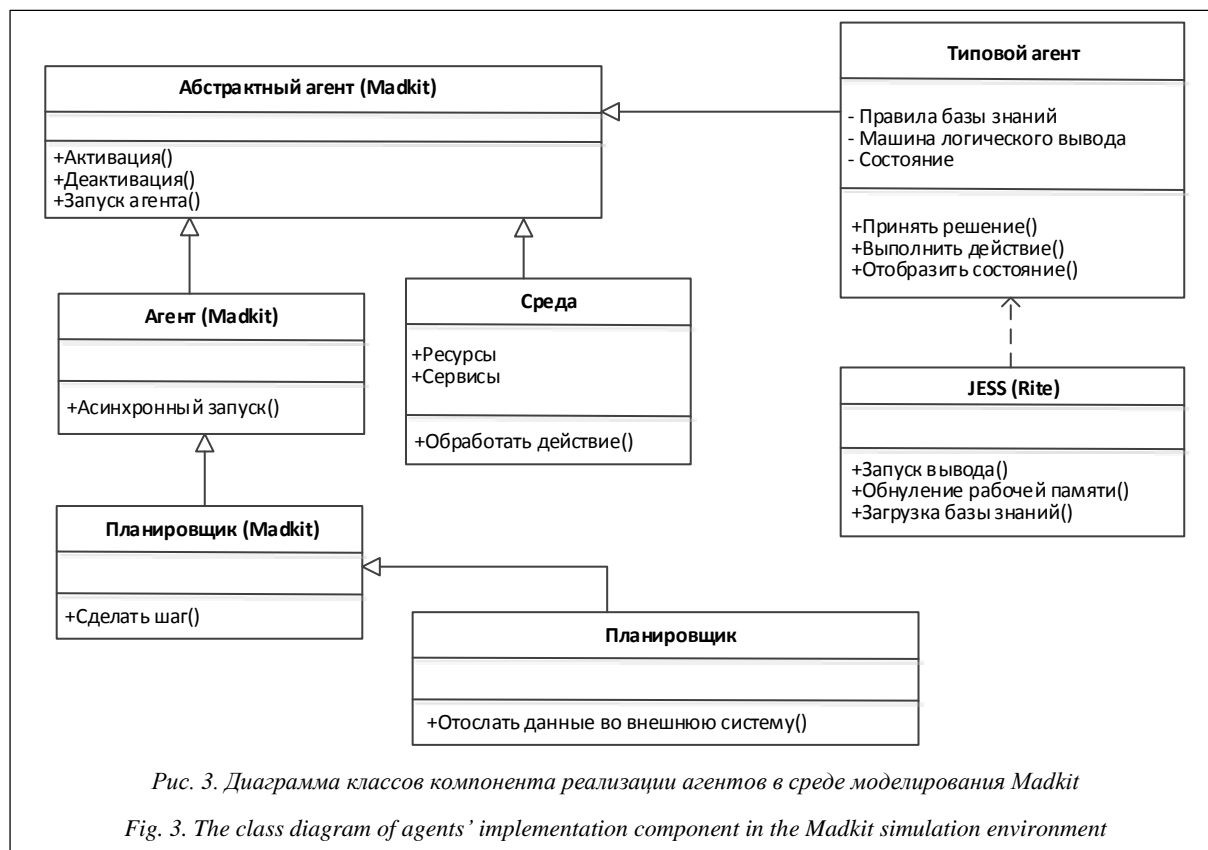
– блок состояния, описывающий свойства агента и содержащий значения этих свойств в некоторую единицу модельного времени, причем свойства имеют объектно-ориентированную структуру;

– блок поведения, описывающий закономерности изменения состояний агента (стратегию поведения), в котором выделяются блоки, ответственные за обработку разных типов информации: декларативный блок, содержащий стратегию поведения агента, формализованную в виде знаний, представленных продукционными правилами, прецедентами и др.; императивный блок, содержащий информацию, представленную в виде программной реализации расчетных процедур, используемых в процессе логического вывода.

Программная реализация типового агента в среде Madkit (рис. 3) осуществляется путем наследования от класса «Абстрактный агент», предназначенного для организации синхронного типа агентного взаимодействия.

При этом в типовом агенте создаются специальные методы, ответственные за выполнение на каждом шаге моделирования следующих операций: *принять решение, выполнить действие, отобразить состояние*. Далее создается специальный планировщик, в качестве которого используется агент асинхронного типа, обеспечивающий в рамках своего жизненного цикла тактирование шагов моделирования.

В текущей реализации в процессе принятия решений типовой агент использует машину вывода Jess для интерпретации своего поведения. В связи с предлагаемой структурой поведения для реализации императивного блока в систему Jess с использованием встроенного интерфейса Userfunction



внедрен общий механизм вызова внешних программных процедур, реализованных на языке Java.

В результате получается следующий принципиальный алгоритм работы типового агента на некотором шаге моделирования  $t$ .

1. Получение сообщений.
2. Обнуление состояния рабочей памяти Jess.
3. Загрузка правил и шаблонов фактов в машину вывода Jess.
4. Загрузка описания текущей ситуации в машину вывода Jess:
  - преобразование литеральных свойств агента в слоты шаблона «мое состояние»;
  - преобразование объектных свойств в факты соответствующих шаблонов;
  - преобразование агентных свойств в факты соответствующих шаблонов (при этом используются состояния агентов на момент времени  $t-1$ ).
5. Преобразование информации из сообщений в факты шаблона.
6. Запуск логического вывода.
7. Поиск в рабочей памяти фактов шаблонов, связанных с событием и действием.
8. Запуск метода среды «Обработать действие».
9. Сохранение состояния из соответствующих фактов рабочей памяти.
10. Обработка действий, связанных с посылкой сообщений между агентами.

Программная реализация компонента осуществлена путем создания веб-сервиса на языке Java, который включает в себя библиотеку среды моделирования Madkit, классы, использующиеся в процессе агентного моделирования («Типовой агент», «Планировщик» и т.д.), а также классы для выполнения служебных функций (взаимодействие с БД или с внешней системой). Основной функцией сервиса является выполнение агентной модели.

Кроме того, учитывая специфику удаленного исполнения агентной модели, необходимо решить задачу информирования пользователя о текущем состоянии процесса моделирования и параметров отдельных агентов. Для решения этой задачи разработан коммуникационный модуль. Модуль реализован на основе WebSocket-сервера [10] программной платформы Node.js и обеспечивает трансляцию данных о состоянии процесса моделирования всем заинтересованным пользователям. В общем случае процесс передачи данных инициируется планировщиком в конце каждого шага моделирования. Передаваемая информация кодируется в JSON-формат и отправляется на HTTP-сервер платформы Node.js, в контексте которого запущен WebSocket-сервер, осуществляющий отправку данных пользователю в соответствии с информацией об адресатах. При этом вопросы отображения полученных данных на компьютере пользователя не относятся к сфере ответственности компонента.

В заключение отметим, что формализация процесса проектирования агентной модели является обязательным элементом для разработки системы поддержки проектирования агентных моделей. В работе дано определение основной функциональности разрабатываемой системы и предложена последовательность этапов, которая является основой для построения алгоритма поддержки действий пользователя при формировании агентной модели от этапа концептуального описания исследуемой области до ее программной реализации. Отличительной особенностью этого варианта формализации процесса проектирования является явное включение этапа создания КМ исследуемой области.

Программная реализация компонента создания КМ обеспечивает пользователей системы возможностью создания согласованной КМ сложной системы. Полученные модели в дальнейшем будут использованы в качестве единой спецификации объектов и процессов исследуемой области для построения логических моделей хранения данных, формирования баз знаний и имитационных моделей.

Программная реализация компонента визуального конструирования баз знаний продукционного типа обеспечивает пользователей системы возможностью создания баз знаний с помощью визуального конструирования, а не написания кода. Следует отметить, что в настоящее время при создании базы знаний используется ограниченное подмножество языка представления знаний (Jess, Clips), так как необходимо обеспечить некоторый компромисс между выразительной силой и сложностью устройства логической модели БД для хранения элементов базы знаний, а также генератора кода. В дальнейшем возможны постепенное увеличение набора доступных условий и действий, а также обеспечение использования машины вывода Clips.

С помощью компонента реализации агентов в среде моделирования Madkit был апробирован предложенный подход к автоматизации программной реализации агентов имитационной модели. В данной реализации обеспечена возможность создания реактивных агентов без построения плана действий. В дальнейшем предполагается осуществить оценку быстродействия генерируемых агентных моделей, а также совершенствование архитектуры типового агента с целью обеспечения возможности реализации агентов с более сложным поведением.

### Литература

1. Столбов А.Б., Павлов А.И. Разработка системы поддержки проектирования имитационных моделей сложных систем на основе декларативного метода описания агентов // Имитационное моделирование: теория и практика. ИММОД-2013: тр. Всерос. науч.-практич. конф. СПб, 2013. С. 267–270.

2. Павлов А.И., Столбов А.Б. Архитектура системы поддержки проектирования агентов для имитационных моделей сложных систем // Программные продукты и системы. 2015. № 2 (109). С. 12–16.
3. Michel F., Ferber J., Gutknecht O. The Multiagent Development Kit. URL: <http://www.madkit.org/> (дата обращения: 21.12.2015).
4. Jess, the rule engine for the Java™ platform. URL: <http://www.jessrules.com/> (дата обращения: 21.12.2015).
5. Anand N., Yang M., J.H.R. van Duin, Tavasszy L. GenCLON: An ontology for city logistics. *Expert Systems with Applications*, 2012, no. 39, pp. 11944–11960.
6. Protégé is a free, open source ontology editor and knowledge-base framework. URL: <http://protege.stanford.edu/> (дата обращения: 21.12.2015).
7. TopBraid composer standard edition. URL: <http://www.topquadrant.com/tools/modeling-topbraid-composer-standard-edition/> (дата обращения: 21.12.2015).
8. CLIPS, a tool for building expert systems. URL: <http://clipsrules.sourceforge.net> (дата обращения: 21.12.2015).
9. Берман А.Ф., Николайчук О.А., Павлов Н.Ю., Юрин А.Ю. Методы и средства автоматизированного построения деревьев событий и отказов // Автоматизация и современные технологии. 2013. № 9. С. 8–16.
10. RFC 6455 – The WebSocket Protocol. URL: <https://tools.ietf.org/html/rfc6455> (дата обращения: 21.12.2015).

DOI: 10.15827/0236-235X.115.079-084

Received 23.12.15

### A PROTOTYPE OF AN AGENTS DESIGN SUPPORT SYSTEM FOR COMPLEX SYSTEM SIMULATION MODELS

(Acknowledgements. The research was partially supported by Russian Foundation for Basic Research, projects no. 14-07-31298, 14-07-0222, 15-07-05641)

<sup>1</sup>Pavlov A.I., Ph.D. (Engineering), Senior Researcher, [asd@icc.ru](mailto:asd@icc.ru)

<sup>1</sup>Stolbov A.B., Junior Researcher, [stolboff@icc.ru](mailto:stolboff@icc.ru)

<sup>1</sup>Institute for System Dynamics and Control Theory Siberian Branch of RAS, Lermontov St. 134, Irkutsk, 664033, Russian Federation

**Abstract.** The paper describes the details of implementation of a support system for designing simulation model agents. It also considers the sequence of stages for developing a simulation model that was used as a basis for constructing an agent model algorithm from the conceptual model of domain to the software implementation. The paper considers software implementation of the components for a conceptual model, for visual design of knowledge bases and for implementation of agents in the Madkit simulation environment. A conceptual model component provides facilities for a hierarchical structure of concepts, determining concepts' properties and relations between them. This description of the relationships is performed by constructing a set of semantic networks, whose vertices are the concepts of a conceptual model, and links are relationships between them. Each of semantic networks explains a different aspect of a considered domain. A component for visual design of knowledge base provides an opportunity for rule construction in the form "IF conditions THEN actions" and code generation in Jess and Clips languages. A component for agent implementation in Madkit includes a simulation library; specially designed software classes ("Typical Agent", "Scheduler", etc.), as well as classes to perform service functions (interaction with a database or an external system). The main function of this component is providing simulation of an agent-based model, as well as informing the user about the current state of a modeling process and the parameters of particular agents. All components are implemented as web applications.

**Keywords.** agent-based simulation, expert system, knowledge-based system.

### References

1. Pavlov A.I., Stolbov A.B. Development of a design support system for complex system simulation models based on a declarative method of describing agents. *Vseross. nauch.-praktich. konf. "Imitatsionnoe modelirovanie. Teoriya i praktika" IMMOD-2013* [All-Russian Science and Practice Conf. "Simulation Modeling. Theory and Practice" IMMOD-2013]. 2013, St. Petersburg, 2013, pp. 267–270 (in Russ.).
2. Pavlov A.I., Stolbov A.B. Architecture of agents design support system for complex systems simulation models. *Programmye produkty i sistemy* [Software & Systems]. 2015, no. 109, pp. 12–16 (in Russ.).
3. Michel F., Ferber J., Gutknecht O. *The Multiagent Development Kit*. Available at: <http://www.madkit.org/> (accessed December 21, 2015).
4. *Jess, the Rule Engine for the Java™ Platform*. Available at: <http://www.jessrules.com/> (accessed December 21, 2015).
5. Anand N., Yang M., van Duin J.H.R., Tavasszy L. GenCLON: An ontology for city logistics. *Expert Systems with Applications*. 2012, no. 39, pp. 11944–11960.
6. *Protégé is a free, open source ontology editor and knowledge-base framework*. Available at: <http://protege.stanford.edu/> (accessed December 21, 2015).
7. *TopBraid Composer Standard Edition*. Available at: <http://www.topquadrant.com/tools/modeling-topbraid-composer-standard-edition/> (accessed December 21, 2015).
8. *CLIPS, a Tool for Building Expert Systems*. Available at: <http://www.clipsrules.sourceforge.net> (accessed December 21, 2015).
9. Berman A.F., Nikolaychuk O.A., Pavlov N.Yu., Yurin A.Yu. Methods and tools for automated construction of event and failure trees. *Avtomatizatsiya i sovremennye tekhnologii* [Automation and Modern Technologies]. 2013, no. 9, pp. 8–16 (in Russ.).
10. *RFC 6455 – The WebSocket Protocol*. Available at: <https://tools.ietf.org/html/rfc6455> (accessed December 21, 2015).



УДК 004.891.2

Дата подачи статьи: 30.11.15

DOI: 10.15827/0236-235X.115.085-089

## **ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА АВТОМАТИЧЕСКОГО ОПРЕДЕЛЕНИЯ КАТЕГОРИИ ПОТЕНЦИАЛЬНЫХ АДРЕСАТОВ ТЕКСТА**

*А.В. Глазкова, аспирант, anya\_kr@aol.com  
(Тюменский государственный университет,  
ул. Володарского, 6, г. Тюмень, 625003, Россия)*

Статья посвящена описанию программного комплекса – интеллектуальной системы автоматического определения категории потенциальных адресатов текста. Основная функция рассматриваемой системы – классификация текстов на основании возрастных категорий их потенциальных читателей.

Автором сделан обзор существующих работ по схожей тематике – классификации документов на естественном языке. Особое внимание уделено исследованиям, посвященным выявлению характеристик авторов или адресатов текста. Сделан вывод о высокой актуальности разработок в данном направлении в связи с их востребованностью в качестве элементов поисковых систем, систем электронного обучения и т.д.

В статье описан подход к моделированию текста, в рамках которого текст представляется в виде набора признаков. Выбор классификационных признаков осуществлен на основании анализа текстов, входящих в Национальный корпус русского языка.

На основе разработанной модели создан программный комплекс, состоящий из трех подсистем: модуля семантико-синтаксического анализа, модуля хранения текстов и модуля классификации. Модуль классификации использует нейронную сеть типа многослойный персептрон. Разработанная интеллектуальная система успешно протестирована на текстах, возрастная категория читателей для которых определена на основании мнений экспертов.

**Ключевые слова:** классификация текстов, обработка естественного языка, нейронная сеть, извлечение информации, ПО.

Решение задач автоматической классификации текстов в последние годы стало одним из приоритетных направлений развития исследований в областях информационного поиска и искусственного интеллекта. Средства автоматической классификации текстов находят применение не только при отборе наиболее релевантных результатов поисковых запросов, но и при решении таких прикладных задач, как фильтрация спама, составление персональных подборок новостей, автоматическое аннотирование, снятие неоднозначности при автоматическом переводе, определение языка текста.

Задача классификации текстов подразделяется на две подзадачи: обучение классификатора и непосредственная классификация текстов. Наибольшую трудность при этом составляет первая подзадача, от успешности решения которой в итоге зависит достоверность проведенной классификации.

Для обучения классификатора прежде всего проводится построение модели текста – представление в виде набора признаков. Решение задачи поиска классификационных признаков зависит от тематики текстов и цели классификации. Выделению значимых классификационных признаков посвящены работы многих исследователей. Например, проблема выявления наиболее информативных признаков описана в работе [1]. В статье [2] рассмотрена задача выделения дифференцирующих признаков при классификации текстовых сообщений. При этом авторами исследования описан метод классификации текстов на естественных языках, основанный на методе взаимной информации. В работе [3] выделены типы признаков на основании шкал, в которых они измеряются, приводятся

примеры признаков каждого типа, а также обсуждаются возможные методы их анализа.

На основании выбранных признаков выделяются тематические классы, на которые будут подразделяться тексты, подлежащие классификации, и проводится обучение классификатора, которое бывает трех видов: с учителем, без учителя и смешанное. Самым качественным считается обучение с учителем, однако оно осуществимо только в случае, когда возможно заранее получить выборку объектов со знанием их классов. Обучение с учителем использует байесовские и линейные методы, методы, построенные на применении деревьев решений или нейронных сетей. Модификации и применению классических методов решения задачи классификации текстов посвящены работы [4–6]. В настоящее время все большую распространенность в качестве средства решения слабоформализуемых задач приобретают искусственные нейронные сети. Рост популярности нейросетевых технологий обусловлен способностью нейронных сетей относить объект, представленный в виде набора классификационных признаков, к соответствующей ему категории на основании найденных скрытых закономерностей в данных обучающей выборки. При наличии достаточного числа примеров, по которым производится обучение, классификаторы, использующие нейронную сеть, часто являются оптимальным решением задач, слабо поддающихся формализации. Среди работ, представленных за последнее время, применению нейросетевых технологий в интеллектуальных системах классификации различных объектов посвящены статьи [7–9]. Перспективы и основные аспекты применения нейросетевых классификато-

ров, а также использования нейронных сетей как составной части систем искусственного интеллекта рассмотрены в работе [10].

В рамках задач автоматической классификации текстов все большую актуальность приобретают проблемы установления характеристик автора и адресата текста, механизмы решения которых реализуются в поисковых системах, системах определения плагиата, построения авторских профилей (выявления возрастных, гендерных и других характеристик автора), электронных библиотек и каталогах.

Подходы к решению задачи установления характеристик автора текста довольно широко освещаются в работах российских и зарубежных ученых. Так, методы идентификации авторства подробно рассмотрены в статье [11]. Различные подходы к выявлению характеристик автора предложены в работах [12–15]. В статье [16] описывается процесс формирования комплексной модели авторского текста. Задача определения адресата текста в настоящее время затрагивается преимущественно зарубежными исследователями. Например, в работах [17–19] рассмотрена задача разработки диалоговых систем, в контексте которой проанализированы признаки, характеризующие текст с точки зрения его ориентации на различные категории читателей. В настоящее время разработка средств решения задачи выявления характеристик адресата текста становится все более актуальной в связи с введением возрастных ограничений на контент интернет-ресурсов, развитием систем электронного обучения, а также с малой освещенностью данной проблемы в работах российских ученых.

Данная работа посвящена описанию программного комплекса – интеллектуальной системы автоматического определения категории потенциальных адресатов текста, которая классифицирует тексты общей тематики на основании возрастной категории их предполагаемых читателей. В рамках исследования классификация текстов осуществляется по двум возрастным категориям – тексты, адресованные взрослым читателям, и тексты детской литературы, что обусловлено аналогичным делением текстов, входящих в обучающую выборку.

### Представление текстов

Пусть имеются текст  $T$  и множество категорий  $K=\{K_1, K_2, \dots, K_n\}$ , с которыми он может быть соотнесен. В процессе классификации текст  $T$  представляется в виде набора (вектора) классификационных признаков:  $T=\{q_j\}$ ,  $1 \leq j \leq L$ , где  $q_j^K$  – классификационный признак;  $L$  – общее число классификационных признаков.

Соответственно категория может быть представлена в следующем виде:  $K_i=\{q_j^K, w_j^K\}$ ,  $1 \leq j \leq L$ ,  $1 \leq i \leq n$ , где  $w_j^K$  – весовой коэффициент классифика-

ционного признака. Весовые коэффициенты характеризуют степень влияния каждого признака на вероятность отнесения объекта к той или иной категории. Предлагаемый подход подробно описан в работе [20], в данной статье сделан акцент на описании программной реализации.

Для дальнейшего представления текста в виде набора признаков был проведен анализ текстов, входящих в Национальный корпус русского языка [21], с целью выявления базовых признаков, на основании которых может быть проведена классификация [22]. Для выявления данных признаков использовались две выборки – художественные тексты различных жанров (историческая проза, приключения, документальная проза и т.д., кроме детской литературы, – всего 5 902 документа, 9 332 659 предложений, 94 538 056 слов) и детская литература (всего 632 документа, 547 735 предложений, 4 742 627 слов). Были выявлены классификационные признаки, отражающие лексические (например, частотность слов или наличие специальных терминов) и синтаксические (частотность частей речи, сложность структуры предложений и синтаксических конструкций) особенности текста. Диапазон значений выделенных признаков может состоять из пары бинарных значений, конечного упорядоченного или неупорядоченного множества значений или бесконечного множества количественных значений.

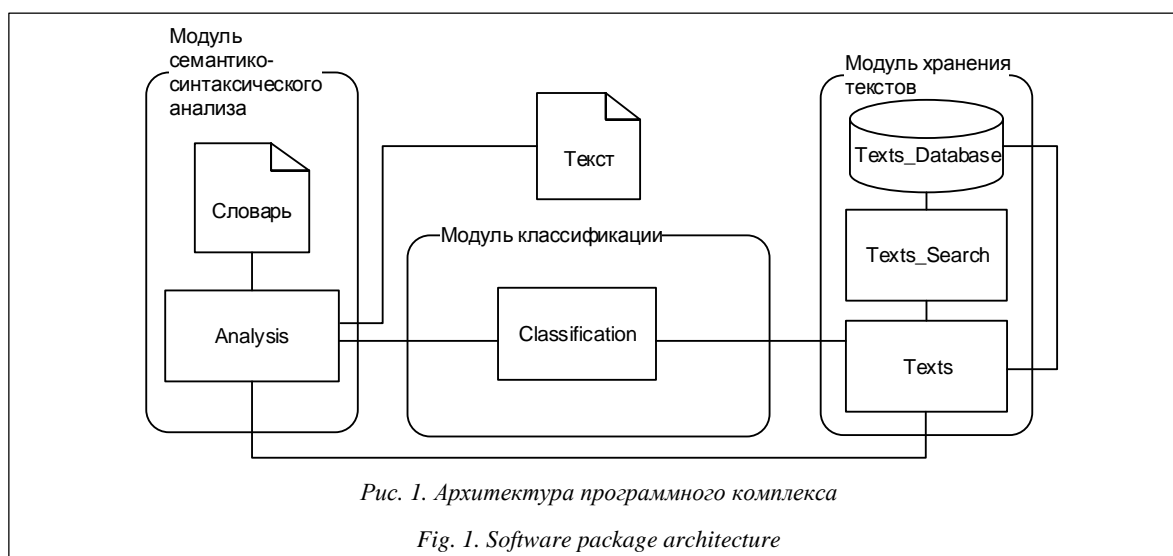
### Программная реализация

Интеллектуальная система разработана в среде Visual Studio 2010 на языке C#. Программный комплекс позволяет вводить или загружать и редактировать текст, предназначенный для классификации. Далее система проводит анализ полученного текста с целью нахождения значений классификационных признаков. Выявленные значения признаков используются в качестве основания для отнесения текста к той или иной категории – определения его возрастной аудитории. Данные о текстах, поступающих в систему, и сопоставленных им категориях могут быть сохранены в реляционной БД.

Рассматриваемая интеллектуальная система состоит из трех подсистем: модуля семантико-синтаксического анализа, модуля хранения текстов и модуля классификации. Архитектура программного комплекса представлена на рисунке 1, где элементы диаграммы названы в соответствии с наименованиями элементов программного комплекса.

Основная функция модуля семантико-синтаксического анализа состоит в представлении текстов в виде наборов характеризующих их признаков.

Загрузка текста для анализа осуществляется из БД или из файла формата .xml или .txt. После загрузки или создания документа пользователю предлагается провести анализ текста на основании



его лексических и синтаксических характеристик. В процессе анализа текст разбивается на предложения, а также представляется в виде модели bag-of-words – множества пар лексем и соответствующих им частотностей. Результатом проведенного анализа текста является набор значений классификационных признаков, который впоследствии может быть использован при работе модуля классификации. Данные о тексте, а также соответствующие ему значения характеристик могут быть сохранены в БД.

Модуль классификации осуществляет определение соответствия текста существующим категориям. Соотнесение текста с категориями выполняется на основе классификационных признаков, полученных в ходе работы модуля семантико-синтаксического анализа.

Для компьютерной реализации классификации была использована нейронная сеть – многослойный персептрон. Использование данного типа сети обусловлено ее способностью к решению слабоформализуемых типов задач на основании имеющихся примеров и выявлению закономерностей в связи входных и выходных данных.

Входной слой нейронной сети содержит количество нейронов, равное количеству классификационных признаков, а выходной слой – количество нейронов, соответствующее количеству категорий (рис. 2). Текст поступает на вход классификатора в виде набора классификационных признаков, выявленных в ходе работы модуля семантико-синтаксического анализа. При этом множество значений признаков для классифицируемого текста следует интерпретировать как входные значения нейронной сети, а весовые коэффициенты – как веса межнейронных связей.

Обучение персептрона производилось по алгоритму обратного распространения ошибки.

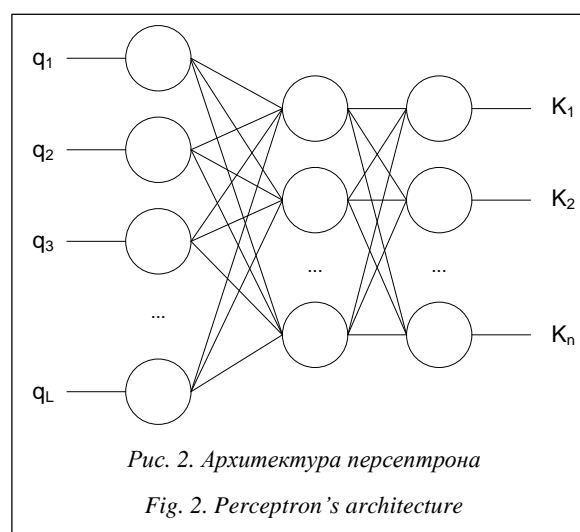
Модуль хранения текстов предназначен для добавления информации, организации хранения данных в реляционных таблицах БД, организации

взаимодействия всех модулей системы. В БД предусмотрена возможность хранения значений признаков различных типов.

Хранение данных о текстах, поступающих в систему, и назначенных им категориях осуществляется в БД, использующей систему управления БД Microsoft SQL Server 2012 Express. Поиск информации о текстах по заданным критериям организован при помощи SQL-запросов.

### Проверка результатов

Для обучения и тестирования нейронной сети использовались тексты, входящие в «Базу данных метатекстовой разметки Национального корпуса русского языка» [23]. База содержит заведомо качественные и максимально разнообразные тексты на русском языке, возрастная категория потенциальных читателей которых (взрослая или детская) определена на основании мнений экспертов. Объем обучающей выборки – 532 текста художественной литературы для взрослых и 510 текстов детской литературы.



Точность классификации для обучающей выборки составила 98 %, для тестовой – 72 %.

В заключение отметим следующее. Автоматическая классификация текстов является важной задачей информационного поиска и обработки естественного языка. В данной статье задача классификации рассмотрена на примере отнесения текстов к той или иной возрастной категории адресатов и представлена интеллектуальная система, протестированная на русскоязычных текстах.

В перспективе разработанный программный комплекс может найти практическое применение в поисковых системах (для отбора релевантного контента), системах обучения, электронных библиотеках и каталогах, системах автоматического реферирования и рецензирования.

### Литература

1. Мангалова Е.С., Агафонов Е.Д. О проблеме выделения информативных признаков в задаче классификации текстовых документов // Вестн. Томск. гос. ун-та. Сер.: Управление, вычислительная техника и информатика. 2013. № 1 (22). С. 96–103.
2. Поляков И.В., Соколова Т.В., Чеповский А.А., Чеповский А.М. Проблема классификации текстов и дифференцирующие признаки // Вестн. Новосиб. гос. ун-та. Сер.: Информационные технологии. 2015. № 2. С. 55–63.
3. Колесникова С.И. Методы анализа информативности разнотипных признаков // Вестн. Томск. гос. ун-та. Сер.: Управление, вычислительная техника и информатика. 2009. № 1. С. 69–80.
4. Виноградов С.Ю. Применение байесовской сети в задаче классификации структурированной информации // Программные продукты и системы. 2013. № 2. С. 155–158.
5. Кубарев А.И., Кукушкина О.В., Поддубный В.В., Шевелев О.Г. Построение таблиц стилей текстовых произведений с использованием алгоритмов классификации на основе деревьев решений // Вестн. Томск. гос. ун-та. Сер.: Управление, вычислительная техника и информатика. 2012. № 4. С. 79–88.
6. Фальк В.Н., Шаграев А.Г., Бочаров И.А. Трансдуктивное обучение логистической регрессии в задаче классификации текстов // Программные продукты и системы. 2014. № 2. С. 114–118.
7. Головкин В.А., Крошенин А.А. Применение нейронных сетей глубокого доверия для выделения семантически значимых признаков // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2015): матер. конф. 2015. С. 481–486.
8. Станкевич Л.А., Сонькин К.М., Нагорнова Ж.В., Хоменко Ю.Г., Шемякина Н.В. Классификация электроэнцефалографических паттернов воображаемых движений пальцами руки для разработки интерфейса мозг–компьютер // Труды СПИИРАН. 2015. № 3. С. 163–182.
9. Палюх Б.В., Шпрехер Д.М., Богатиков В.Н. Диагностика электромеханических систем на основе нейросетевых технологий // Программные продукты и системы. 2015. № 3. С. 5–11.
10. Мелихова О.А. Нейронные сети как составная часть систем искусственного интеллекта // Информатика, вычислительная техника и инженерное образование. 2015. № 1. С. 40–51.
11. Батура Т.В. Формальные методы установления авторства текстов и их реализация в программных продуктах // Программные продукты и системы. 2013. № 4. С. 286–295.
12. Argamon S., Koppel M., Pennebaker J., Schler J. Automatically profiling the author of an anonymous text. Communications of the ACM, 2009, no. 2, pp. 119–123.
13. Burger J., Henderson J., Kim G., Zarella G. Discriminating gender on twitter. Proc. Conf. on Empirical Methods in Natural Language Processing, 2011, pp. 1301–1309.
14. Gillam L. Readability for author profiling? Proc. CLEF, 2013, Valencia, Spain, 4 p.
15. Орлов Ю.Н., Осминин К.П. Определение жанра и автора литературного произведения статистическими методами // Прикладная информатика. 2010. № 2. С. 95–108.
16. Муха А.В., Розалиев В.Л., Орлова Ю.А., Заболевая-Зотова А.В. Автоматизированный подход к определению авторства текста // Изв. ВолГТУ. 2013. № 14. С. 51–54.
17. Akker R. op den, Traum D. A comparison of addressee detection methods for multiparty conversations. Proc. of methods for multiparty conversations. Amsterdam, 2009, pp. 99–106.
18. Baba N., Huang H.-H., Nakano Y.I. Addressee identification for human-human-agent multiparty conversations in different proxemics. Proc. 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction. ICMI. 2012, Santa Monica, USA, 6 p.
19. Lee H., Stolcke A., Shriberg E. Using out-of-domain data for lexical addressee detection in human-human-computer dialog. Proc. North American ACL/Human Language Technology Conf. Atlanta, 2013, pp. 215–219.
20. Глазкова А.В., Захарова И.Г. Подход к моделированию задачи автоматической классификации текстов (на примере их отнесения к определенной возрастной аудитории) // Вестн. Тюмен. гос. ун-та. Сер.: Физ.-мат. моделирование. Нефть, газ, энергетика. 2014. № 7. С. 205–211.
21. Национальный корпус русского языка. 2015. URL: ruscorgpora.ru (дата обращения: 08.10.2015).
22. Глазкова А.В. Проверка информативности классификационных признаков в задаче автоматической классификации текстов на естественном языке // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2015): матер. конф. 2015. С. 541–544.
23. База данных метатекстовой разметки Национального корпуса русского языка (коллекция детск. лит-ры). 2014. URL: ruscorgpora.ru (дата обращения: 08.10.2015).

DOI: 10.15827/0236-235X.115.085-089

Received 30.11.15

### INTELLIGENT SYSTEM FOR AUTOMATIC IDENTIFICATION OF TEXT ADDRESSEE CATEGORY

<sup>1</sup>Glazkova A.V., Postgraduate Student, anya\_kr@aol.com

<sup>1</sup>Tyumen State University, Volodarskogo St. 6, Tyumen, 625003, Russian Federation

**Abstract.** Automatic text classification is an important task of information retrieval and natural language processing. The article describes the intelligent system for automatic identification of the categories of text potential recipients. The main function of the system is text classification based on text addressee category.

The author provides an overview of existing work on similar subjects, which is document classification in natural language. Particular attention he gives to research related to the identification of text authors' or recipients' characteristics. There is a conclusion that developments in this field are highly relevant due to the demand for them as elements of the search engines, e-learning systems, etc.

The paper presents an approach to text modeling when the text is presented as a set of features. A selection of classification features based on the analysis of texts belonging to the Russian National Corpus.

The intelligent system based on the proposed mathematical model of the text consists of three subsystems: a semantic-syntactic parsing module, a text storage module and a classification module. The classification module uses a neural network belonging to the type of multi-layer perceptron. The developed intelligent system has been successfully tested on texts with potential recipients' category defined on the basis of expert opinion.

**Keywords:** text classification, natural language processing, neural network, data mining, software.

### References

1. Mangalova E.S., Agafonov E.D. On features selection approach for text mining problem. *Vestn. Tomsk. gos. un-ta. Ser.: Upravlenie, vychislitel'naya tekhnika i informatika* [The Bulletin of Tomsk State Univ. Management, computer engineering and computer science]. 2013, no. 1 (22), pp. 96–103 (in Russ.).
2. Polyakov I.V., Sokolova T.V., Chepovsky A.A., Chepovsky A.M. Text classification problem and features set. *Vestn. Novosib. gos. un-ta. Ser.: Informatsionnye tekhnologii* [The Bulletin of Novosibirsk State University. Series: Information Technology]. 2015, no. 2, pp. 55–63 (in Russ.).
3. Kolesnikova S.I. Methods of different-type features informativity analysis. *Vestn. Tomsk. gos. un-ta. Ser.: Upravlenie, vychislitel'naya tekhnika i informatika* [The Bulletin of Tomsk State University. Management, Computer Engineering and Computer Science]. 2009, no. 1, pp. 69–80 (in Russ.).
4. Vinogradov S.Yu. The use of Bayesian network in structured text classification. *Programmnye produkty i sistemy* [Software & Systems]. 2013, no. 2, pp. 155–158 (in Russ.).
5. Kubarev A.I., Kukushkina O.V., Poddubny V.V., Shevelev O.G. Building a style sheet using text classification algorithms based on decision trees. *Vestn. Tomsk. gos. un-ta. Ser.: Upravlenie, vychislitel'naya tekhnika i informatika* [The Bulletin of Tomsk State University. Management, Computer Engineering and Computer Science]. 2012, no. 4, pp. 79–88 (in Russ.).
6. Falk V.N., Shagraev A.G., Bocharov I.A. Transductive learning logistic regression for the text classification problem. *Programmnye produkty i sistemy* [Software & Systems]. 2014, no. 2, pp. 114–118 (in Russ.).
7. Golovko V.A., Kroschenko A.A. Applying deep belief neural networks to extract valuable semantic features. *Otkrytye semanticheskie tekhnologii proektirovaniya intellektualnykh sistem (OSTIS-2015): mater. konf.* [Proc. Open Semantic Technologies for Intelligent Systems (OSTIS-2015)]. 2015, pp. 481–486 (in Russ.).
8. Stankevich L.A., Sonkin K.M., Nagornova Zh.V., Homenko Yu.G., Shemyakina N.V. Classification of electroencephalographic patterns of imaginary one-hand finger movements for brain-computer interface development. *Trudy SPIIRAS* [Proc. of SPIIRAS]. 2015, no. 3, pp. 163–182 (in Russ.).
9. Palyukh B.V., Shprekher D.M., Bogatkov V.N. Electromechanical systems diagnosis based on neural network technologies. *Programmnye produkty i sistemy* [Software & Systems]. 2015, no. 3, pp. 5–11 (in Russ.).
10. Melikhova O.A. Neural networks as a part of artificial intelligence systems. *Informatika, vychislitel'naya tekhnika i inzhenernoe obrazovanie* [Informatics, Computer Science and Engineering Education]. 2015, no. 1, pp. 40–51 (in Russ.).
11. Batura T.V. Formal methods of text authorship attribution and their usage in software products. *Programmnye produkty i sistemy* [Software & Systems]. 2013, no. 4, pp. 286–295 (in Russ.).
12. Argamon S., Koppel M., Pennebaker J., Schler J. Automatically profiling the author of an anonymous text. *Communications of the ACM*. 2009, no. 2, pp. 119–123.
13. Burger J., Henderson J., Kim G., Zarella G. Discriminating gender on twitter. *Proc. Conf. on Empirical Methods in Natural Language Processing*. 2011, pp. 1301–1309.
14. Gillam L. Readability for author profiling? *Proc. CLEF*, 2013, Valencia, Spain, 4 p.
15. Orlov Yu.N., Osminin K.P. The definition of the genre and the author of a literary work with statistical methods. *Prikladnaya informatika* [Applied Informatics]. 2010, no. 2, pp. 95–108 (in Russ.).
16. Mukha A.V., Rozaliev V.L., Orlova Yu.A., Zaboloeva-Zotova A.V. Automated approach to authorship attribution. *Izv. VolgGTU* [News of Volgograd State Technical Univ.]. 2013, no. 14, pp. 51–54 (in Russ.).
17. Akker R. op den, Traum D. A comparison of addressee detection methods for multiparty conversations. *Proc. of Methods for Multiparty Conversations*. Amsterdam, 2009, pp. 99–106.
18. Baba N., Huang H.-H., Nakano Y.I. Addressee identification for human-human-agent multiparty conversations in different proxemics. *Proc. 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction. ICMI*. 2012, Santa Monica, USA, 6 p.
19. Lee H., Stolcke A., Shriberg E. Using out-of-domain data for lexical addressee detection in human-human-computer dialog. *Proc. North American ACL/Human Language Technology Conf.* Atlanta, 2013, pp. 215–219.
20. Glazkova A.V., Zakharova I.G. An approach to modeling automatic text classification problem (case study of the audience age prediction). *Vestn. Tyumen. gos. un-ta. Ser.: Fiz.-mat. modelirovanie. Neft, gaz, energetika* [Herald of Tyumen State Univ. Physical and Mathematical Modeling. Oil, Gas, Energy]. 2014, no. 7, pp. 205–211 (in Russ.).
21. *Natsionalny korpus russkogo yazyka* [Russian National Corpus]. 2015. Available at: <http://ruscorpora.ru/> (accessed October 8, 2015).
22. Glazkova A.V. Classification features informational content testing for automatic natural texts classification task. *Otkrytye semanticheskie tekhnologii proektirovaniya intellektualnykh sistem (OSTIS-2015): mater. konf.* [Proc. Open Semantic Technologies for Intelligent Systems (OSTIS-2015)]. 2015, pp. 541–544 (in Russ.).
23. *Baza dannykh metatekstovoy razmetki Natsionalnogo korpusa russkogo yazyka* [Database metatextual marking Russian National Corpus]. Available at: <http://ruscorpora.ru/> (accessed October 8, 2015).

## **A NUMERICAL SIMULATION OF MULTI-SPECTRAL PHOTO-ACOUSTIC IMAGING TOWARD HEMATOLOGICAL ANALYSIS**

<sup>1</sup>*Trinh Quang Duc, Dr.Sc., duc.trinhquang@hust.edu.vn;*

<sup>2</sup>*Nguyen Hoai Giang, Dr.Sc., giangnh@hou.edu.vn*

<sup>1</sup>*Hanoi University of Science and Technology, 1 Dai Co Viet Road, Hanoi, Viet Nam*

<sup>2</sup>*Hanoi Open University, B101, Nguyen Hien Street, Hanoi, Viet Nam*

**Abstract.** This paper presents pilot studies to assert the feasibility of Photo-Acoustic Tomography combined with Photo-Acoustic Spectroscopy applied to quantitative clinical examination as noninvasive measurement for blood analysts. To estimate the parameters of practical measurement systems in the future, we simulate obtaining the results using a program that was built in Matlab environment. For the assessment, we examined the technique in the case of glucose measurement as an example. With assumption of an aqueous mixture that includes: Glucose, Urea, Albumin,  $\gamma$ -Globulin evenly dissolved in water, the glucose concentration was extracted from the mixture under excitation light of 300mW and the Fourier analysis of a photo-acoustic signal. Throughout the simulation, the glucose concentration was determined as 2.8 g/l corresponding to photo-acoustic pressure of 2.27 Pa in the condition of 10 mm soft tissue depth and 1mm diameter vessel under the light excitation power.

**Keywords:** natural frequency, photo-acoustic spectroscopy, concentration measurement.

Nowadays, a number of people afflicted with diabetes mellitus is increasing from 382 million to 592 million as a forecast for a period from 2013 to 2035 [1]. In 2011, 1.4 million people around the world had died due to diabetes mellitus being the one of the 8 top causes of death [2]. Diabetes is a disease causing serious long-term influence. The most significant consequences are cardiovascular, retinal, renal and neuropathic pathologies. Determination glucose concentration in blood is an important method to diagnose the patients who are afflicted by diabetes mellitus. Conventional blood examination methods using drawn blood such as test-strip, electronic glucose meter, interstitial blood fluid and enzyme electrodes are invasive techniques leading to the risk of infection. They are also painful and uncomfortable in several situations. Noninvasive techniques were developed for glucose monitoring as well as other blood analysis. They eliminated the disadvantages of invasive methods. Many optical techniques were employed to overcome the drawback such as polarized light measurement, Raman spectroscopy, near infrared spectroscopy and others. They provide non-ionized and reduce the risk of radiation for medical applications. However, using each method depends on aims of studies such as: optical polarization provides information about structurally anisotropic tissue components: collagen and muscle fiber; optical spectra based on absorption, Raman scattering provides biochemical information because it is related to molecular; optical absorption provides contrast for functional image, etc. Moreover, there were several popular reconstruction image methods that were also applied: Diffuse Optical Tomography (DOT) with excellent contrast and good imaging depth (~50mm), but poor resolution (~5mm) [3]; Ultrasound with excellent resolution (~0.3mm), good imaging depth (~60mm), limited by poor contrast [3]; Photo Acoustic Tomography with excellent contrast, good imaging

depth (~60mm) and excellent resolution (0.01mm) [3]. With the aims of precise determination targeting to glucose concentration in blood, good glucose imaging depth with excellent resolution images is required, a Photo Acoustic Tomography method can be considered. This method is based on photo acoustic effects. When the light passes a biological tissue, an acoustic wave is induced due to elastic thermal caused by interaction of photo streams with the tissues. These ultrasound waves were acquired and then reconstructed with the Time Reversal Algorithm to be an image which the pattern performs in high contrast [3]. The intensity values of the photo acoustic image presented as overlap absorption of the components, leading to imprecise measurement of distribution of each component concentration through the absorption coefficient. The Photo-Acoustic Spectroscopy method can be a solution to extract photo-acoustic spectrum of the components based on their natural frequencies from the measured complex signal. Using the spectrum analysis it is possible to calculate the precise values of the component concentration in the image pattern.

This paper presents pilot studies to assert the feasibility of Photo-Acoustic Tomography combined with Photo-Acoustic Spectroscopy applied to quantitative clinical examination as noninvasive measurement for blood analysts. For the assessment, we examined the technique in the case of glucose measurement as an example. To estimate the parameters of the practical measurement systems in the future, we firstly simulate to obtaining the results using a program that was built in Matlab environment.

### **Theoretical basis**

The photo-acoustic imaging is usually obtained with illumination from an optical pulses train with the duration of  $r(t - \tau)$  because TOF (time-of-flight) of

photo-acoustic waves is used to determine the spatial information of wave sources. Assuming that, a photo-acoustic image for a blood vessel, which has been shown in figure 1, will be imaged with the condition of the light energy as

$$E_0(t) = \int_0^t I_0 r(t-\tau) d\tau, \quad (1)$$

where  $I_0$  is the intensity amplitude of the optical pulses.

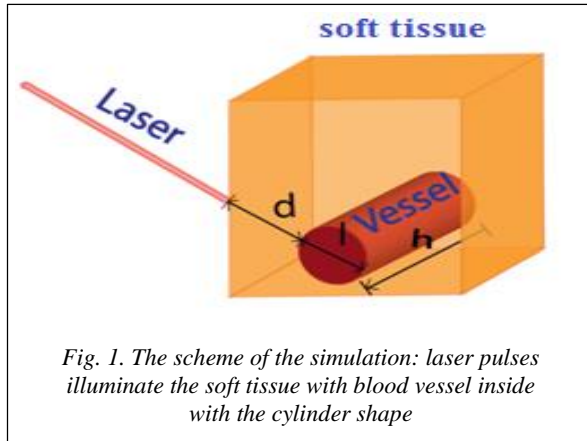


Fig. 1. The scheme of the simulation: laser pulses illuminate the soft tissue with blood vessel inside with the cylinder shape

When the soft tissues with the absorption coefficient  $a_{st}$  are illuminated by optical pulses, a part of light energy is absorbed and at the boundary of the vessel, the intensity amplitude of optical pulses are reduced in Beer-Lambert law [3] so that

$$I_{01}(t) = I_0 e^{-a_{st}d} r(t-\tau), \quad (2)$$

where  $d$  represents the distance from the illuminated point posed on the boundary of the biological tissue to the boundary of blood vessel. After pass through the blood vessel, the absorption energy can be written as

$$E_{ht}(t) = \int_0^t I_{ht}(t) dt = \int_0^t I_0 e^{-a_{st}d} (1 - e^{-\alpha l}) r(t-\tau) dt \quad (3)$$

with  $\alpha$ ,  $l$  as the absorption coefficient and the diameter of the blood vessel (optical path length), respectively.

Taking into account a length of time from 0 to  $\tau$ , Equation 3 can be rewritten as

$$E_{ht} = \int_0^\tau I_0 e^{-a_{st}d} (1 - e^{-\alpha l}) r(t-\tau) dt = I_0 e^{-a_{st}d} (t - e^{-\alpha l}). \quad (4)$$

Beside it, the converted heat energy in blood vessel is also presented as

$$Q_{ht} = m C_v T. \quad (5)$$

With  $m$  is mass of blood (Kg);  $T$  is a temperature (K);  $C_v$  is specific heat capacity (J/KgK).

Assuming that, the efficiency of conversation between absorption light energy and heat energy in blood vessel is  $\eta_{th}$ . The heat energy can be presented as  $Q_{ht} = \eta_{th} E_{ht}$ ,

$$Q_{ht} = \eta_{th} I_0 e^{-a_{st}d} (1 - e^{-\alpha l}). \quad (6)$$

Inserting equation 6 into equation 5, we obtain

$$T = \eta_{th} \frac{I_0 e^{-a_{st}d} (1 - e^{-\alpha l})}{m C_v}. \quad (7)$$

On the other hand, the equation of the state applied to ideal gas law is expressed by Clapeyron as [4]

$$PV = \frac{m}{\mu} RT. \quad (8)$$

In which,  $P$  is pressure (Pa);  $V$  is a volume of blood ( $m^3$ );  $T$  is a temperature (K);  $m$  is a mass (Kg);  $\mu$  is a molar mass (Kg);  $R$  is a constant 8.31 (Pa  $m^3$ /mol K).

Replacing equation 7 by equation 8 with consideration of a volume unit, the pressure induced by the heat can be expressed as:

$$P = \eta_{th} \frac{R I_0 e^{-a_{st}d} (1 - e^{-\alpha l})}{V \mu C_v}. \quad (9)$$

Equation 9 presents a relationship between the localized photo-acoustic pressure and the incident light intensity.

Taking into account a position  $x$  inside the blood vessel at a point of time  $t$ , the photo-acoustic wave equation can be rewritten in Feynman derivation [5] as

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{v_s^2} \frac{\partial^2 p}{\partial t^2}. \quad (10)$$

The photo-acoustic pressure  $p(x, t)$  is the solution of the wave equation [5] and presented as:

$$p(x, t) = P \cos(\omega t \pm kx - \theta), \quad (11)$$

here  $P$  is amplitude of a localized photo-acoustic pressure;  $\omega$  angular natural frequency of the acoustic medium;  $k = \frac{\omega}{v}$  with  $v = 1540$  m/s wave velocity;  $\theta$  is an initial angular phase.

Inserting the equation 9 into equation 11, the photo-acoustic pressure function can be presented as

$$p(x, t) = \eta_{th} \frac{R I_0 e^{-a_{st}d} (1 - e^{-\alpha l})}{V \mu C_v} \cos(\omega t \pm kx - \theta). \quad (12)$$

The biological tissue is structured from sophisticated components including organelles in which each component owns a natural frequency. Thus, the equation 12 can be considered as a complex acoustic wave at the position  $x$  and time  $t$  which is synthesized from the component waves generated by a single laser pulse. According to above mentioned, assuming that inside the vessel, the blood contains components whose each  $i^{th}$  agent has a specific molar absorptivity  $\epsilon_i$ , absorption coefficient  $a_i$  with  $a_i = \epsilon_i c_i$  ( $c_i$  is concentration of the  $i^{th}$  agent in the blood vessel), angular natural frequency  $\omega_i$  and molar mass  $\mu_i$ . The photo-acoustic component wave generated by each component medium is written as

$$p_i(x, t) = \eta_{th} \frac{R I_0 e^{-a_{st}d} (1 - e^{-\epsilon_i c_i l})}{V \mu_i C_v} \cos(\omega_i t \pm k_i x - \theta_i), \quad (13)$$

here the term  $\eta_{th} \frac{R I_0 e^{-a_{st}d} (1 - e^{-\epsilon_i c_i l})}{V \mu_i C_v}$  mentions to the amplitude of the each component photoacoustic wave,  $P_i$ . Under this consideration the measurement photoacoustic wave is synthesized from  $M$  component

photoacoustic waves generated by the  $M$  agents in the blood vessel, equation 11 can be rewritten as

$$p(x, t) = \sum_{i=0}^M p_i(x, t). \quad (14)$$

Due to the synthesis of the photo-acoustic waves, the concentrations of the agents measured through the light absorbance of the medium are superposed, then lead to the inaccuracy. To measure the agent concentration precisely, the photo-acoustic complex wave should be extracted to be individual acoustic waves based on natural frequency. For this purpose the Discrete Fourier transform can be considered and the transform equation for  $p(x, t)$  signal with  $N$  samples in a period as [6] will be

$$p(x, \omega) = \sum_{t=0}^{N-1} p(x, t) e^{-i\omega t}. \quad (15)$$

After using the Discrete Fourier transform the amplitude  $P_i$  of each component of a photoacoustic wave is determined. The concentration of each agent is obtained as

$$c_i = \frac{1}{\varepsilon_i I} \ln \left( \frac{\eta_{th} R I_o e^{-\alpha_{ad}}}{\eta_{th} R I_o e^{-\alpha_{ad}} - P_i V \mu_i C_v} \right). \quad (16)$$

Figure 2 illustrates a scheme for a photo acoustic Spectroscopy method.

### Simulation

The concentration of blood agents can be calculated directly from equation 16 after receiving  $P_i$  from the Fourier analysis through equation 15. However, the images showing the concentration distribution dynamically with the data of the individual photoacoustic wave should be examined to assert the feasibility for the development based on this technique toward clinical application as well as the visualization provision for the observers. For simplicity, here, we assume that the conversation between light energy and heat energy is 100 %. So, equation 12 can be rewritten as:

$$p(x, t) = \frac{R I_o e^{-\alpha_{ad}} (1 - e^{-at})}{V \mu C_v} \cos(\omega t \pm kx - \theta). \quad (17)$$

To simulate the scheme of the measurement, we used an aqueous mixture pattern in a blood vessel that includes 5 agents: Glucose, Urea, Albumin,  $\gamma$ -Globulin. Some reports [3] discuss the simulation of the light propagation in biological tissues proposed under consideration of the energy transfer approximated as diffusion equation. It is called a forward problem. Due to the small acoustic scattering coefficient of the

biological tissue [7], the attenuation of the photoacoustic wave is neglected. In the condition of scattering dominated media, the value of  $a_{st}$  in equation 17 corresponding to the effective attenuation coefficient in soft tissue was estimated at  $6.88 \text{ cm}^{-1}$  [8, 9].

According to some reports, at the wavelength of 1 600 nm the molar absorptivity of glucose was the most different in comparison with other agents [10, 11]. Therefore, this wavelength was considered to determine precisely the glucose concentration precisely.

Table 1

The parameters of blood agents in simulation conditions including  $a_i$ ,  $\mu_i$ ,  $f_i$ , and  $\varepsilon_i$

	Glucose	Urea	Albumin	$\gamma$ -Globulin	Water
$a$ $\text{cm}^{-1}$	0.01	$1.4 \cdot 10^{-4}$	$4.18 \cdot 10^{-3}$	$4.18 \cdot 10^{-2}$	$1.34 \cdot 10^{-3}$
$\mu$ $\frac{\text{kg}}{\text{mol}}$	0.18 [12]	0.06 [13]	66.5 [14]	150 [15]	0.018 [16]
$f$ Hz	$1 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$5 \cdot 10^{-7}$	$7 \cdot 10^{-7}$	$10 \cdot 10^{-7}$
$\varepsilon$ $\frac{1}{\text{molcm}}$	0.644	0.045	9.28	24.19	0.005

Table 2

Physical properties of the simulated samples

Conditions	Values
Thickness of mixture pattern (1)	0.1 cm [11]
Effective attenuation coefficient of soft tissue	$6.88 \text{ cm}^{-1}$ [8, 9]
Initial phase angular $\theta$	$\pi/4$
Position $x$	0.05 cm
Specific heat capacity $C_v$	4000 (J/KgK) [3]
Velocity wave in mixture $v$	1540 m/s [3]
The effective height of vessel ( $h$ )	$h=1 \text{ mm}$
The volume of blood vessel ( $V$ )	$1.57 \times 10^{-9} \text{ m}^3$

At 1 600 nm specific absorption ( $a_i$ ) [10, 11], molar mass ( $\mu_i$ ), molar absorptivity ( $\varepsilon_i$ ), natural frequencies ( $f_i$ ) of Glucose, Urea, Albumin,  $\gamma$ -Globulin and Water were addressed via the references [10, 11] and assumed as shown in table 1. Table 2 describes physical properties of the simulated samples. The scattering of components also can be neglected because they are too small in 1mm thickness of the mixture pattern.

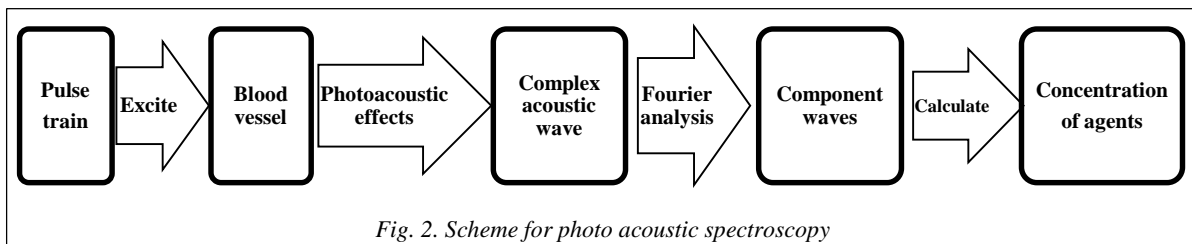


Fig. 2. Scheme for photo acoustic spectroscopy



Based on equation 1, a pulse train which is used as excitation light to the mixture and has duration  $\tau: (t - \tau)$ , is generated from two step pulses  $r(t - \tau) = u_1(t) + u_2(t)$  with  $u(t) = \begin{cases} 1, & t \geq 0, \\ 0, & t < 0, \end{cases}$

$u_1(t) = u(t - t_1)$ ,  $u_2(t) = -u(t - t_2)$ , duration  $\tau = t_2 - t_1$ . In this simulation with optical length of 0.1 cm  $\tau$  was calculated as  $4 \times 10^{-7}$  s [3].

With sampling frequency  $100 \times 10^7$  Hz we generated the waves for each component in a mixture, then they was superposed leading to a complex acoustic wave:

$$p(x, t) = \sum_{i=1}^{i=5} p_i(x, t) = p_{\text{glucose}}(x, t) + p_{\text{urea}}(x, t) + p_{\text{albumin}}(x, t) + p_{\gamma\text{-globulin}}(x, t) + p_{\text{water}}(x, t).$$

This wave were analyzed based on natural frequencies by Discrete Fourier transform in Matlab environment ( $N=10\ 000$  samples), the component photo-acoustic waves were separated. Using peaks detection algorithm, we obtained the individual photo-acoustic pressure amplitudes of component agents. The concentration of agents is calculated as

$$c_i = \frac{1}{\varepsilon_i l} \ln \left( \frac{R I_o e^{-\alpha_{ad}}}{R I_o e^{-\alpha_{ad}} - P V \mu_i C_v} \right). \quad (18)$$

To visualize the concentration distribution, the numerical results of this simulation were assigned using a color. Each color represents each agent and a mixed color expresses a mixture.

## Results and discussion

Figure 3 shows the relationship between photoacoustic pressure generated by glucose and the depth of soft tissue under excitation power of laser.

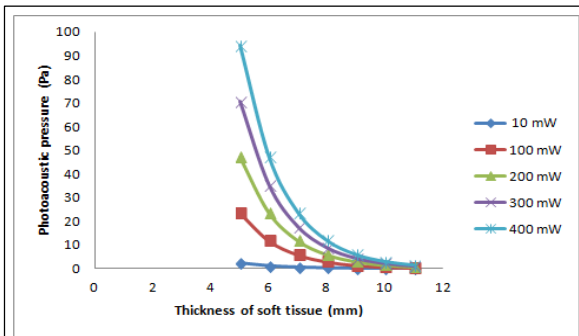


Fig. 3. Relationship between photoacoustic pressures generated by glucose and depth of biological tissue

As shown in figure 3, with excitation laser power of 300 mW, at the soft tissue depth of 10 mm, the photoacoustic pressure corresponds approximately to 2.27 Pa. This pressure is available for measurement using a typical ultrasonic transducer in sensitivity of 40 nV/Pa.

Figure 4 presents square pulses with period time  $T$  equal to  $10 \times 10^{-7}$  s, duration  $\tau$  equal to  $4 \times 10^{-7}$  s were presented in figure 4.

When pulse train excited a mixture, complex acoustic waves were recorded and expressed using  $t = (0:35) \times 10^{-7}$  s as shown in figure 5.

Using Fourier analysis for this complex acoustic wave, we obtained the results the spectrum of photo-acoustic pressure generated from the mixture under excitation of the light pulse train. Then we detected the peak values represented in the photo-acoustic pressure amplitudes of each component using the peak searching algorithm. The results were indicated as shown in figure 6.

At frequency of  $10^7$  Hz, the photo-acoustic pressure amplitude of Glucose reached to 2.27 Pa. Using this value we calculated glucose concentration based on equation 18 to obtain the concentration value of 0.016 mol/L corresponding to 2.8 g/L.

Correspondingly, the values of the agent concentrations were assigned to intensity the components' color in range of human physiological conditions. In normal human blood concentration ranges of Glucose [17], Urea [17], Albumin [18],  $\gamma$ -Globulin [18] were estimated in table 3. Furthermore, with molar absorptivity of water is 0.005 and specific absorption is

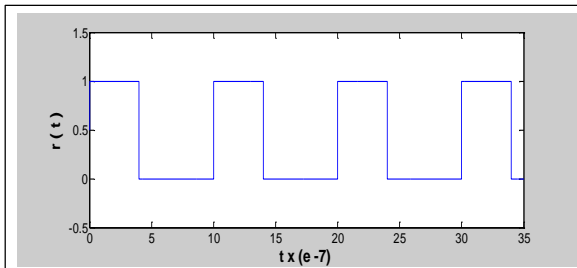


Fig. 4. The pulse train

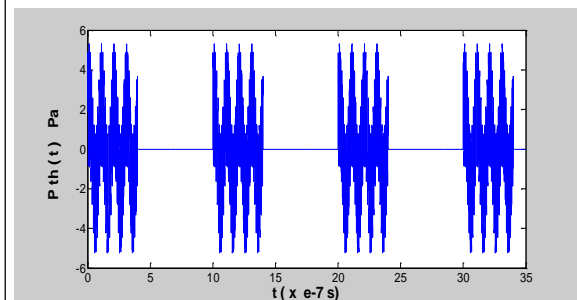


Fig. 5. The complex acoustic wave

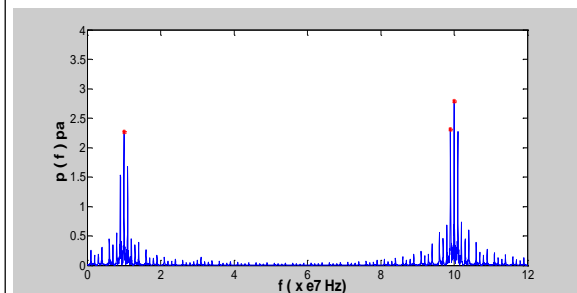


Fig. 6. The photo acoustic spectrum and peak detection of photo-acoustic pressure amplitude

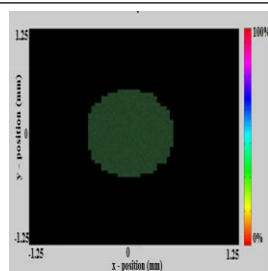


Fig. 7. The color of the mixture

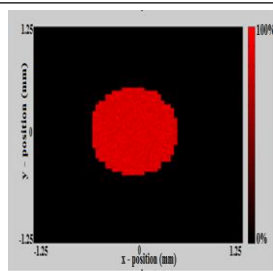


Fig. 8. Glucose concentration

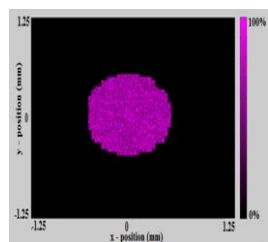


Fig. 9. Urea concentration

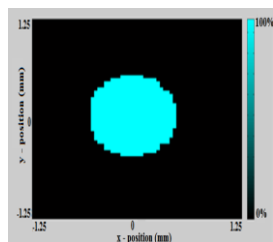


Fig. 10. Water

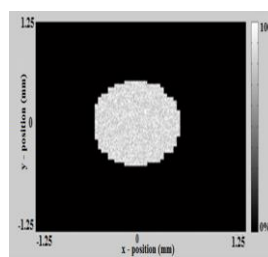


Fig. 11. Albumin concentration

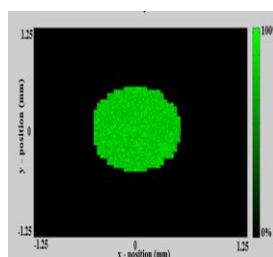


Fig. 12. Globulin concentration

$1.34 \times 10^{-3} \text{ cm}^{-1}$  [10, 11], we also have volumetric distribution of water 0.268 mol/L corresponding to 482.4 mg/dL. A mixture was produced by dissolving from the above agents.

Figure 7 shows the color of the mixture. To separate the agents for concentration distribution observation, the color of the mixture is presented as the component's colors which are used to synthesize the mixture color. It is easy when component agents are extracted from the mixture.

Table 3

**Concentration ranges of agents in normal human blood**

Agents	Concentration ranges
Glucose	65–105 mg/dL
Urea	7–18 mg/dL
Albumin	3200–4800 mg/dL
$\gamma$ -Globulin	500–1500 mg/dL
Water	482.4 mg/dL

Each color represents one agent: Glucose-red, Urea-pink lotus, Albumin-bone,  $\gamma$ Globulin-green, Wa-

ter-turquoise. Brightness of colors (dark to light) is based on concentration of component agents as shown in figures 8–12.

Concentration ranges of Glucose, Urea: 65–105 mg/dL, 7–18 mg/dL with 0–100 % respectively.

Concentration ranges of  $\gamma$ -Globulin, Albumin: 500–1 500 mg/dL, 3 200–4 800 mg/dL with 0–100 % respectively.

## Conclusion

The technique being a combination of Photo Acoustic Tomography and Photo Acoustic Spectroscopy applied to blood analysis was demonstrated using a numerical simulation. Throughout an example of glucose measurement scheme, the precise glucose concentration was extracted from an imitated blood compound including 5 different components. Under the condition of 1 mm vessel diameter and 10 mm thickness of soft tissue with the light excitation of 300 mW generated through the pulse train with its width 0.4 microseconds, the 2D cross-sectional images were reconstructed. The intensities at image point data can be used to indicate the concentration quantitatively by Fourier analysis. With the sensitivity of few pascals for the photo-acoustic pressure level and several hundreds milliwatts for the light excitation, this technique promises the developments concerning to clinical applications of blood analysis practically in the near future.

## References

1. Intern/ Diabetes Foundation: *Diabetes Atlas*. Available at: <http://www.diabetesatlas.org/> (accessed April 4, 2014).
2. *Diabetes Fact sheet N°312. WHO*. 2013. Available at: <http://www.who.int/mediacentre/factsheets/fs312/en/> (accessed March 25, 2014).
3. Wang L.V., Wu H. *Biomedical Optics: Principles and Imaging*. Wiley-Interscience Publ., Hoboken, NJ, 2007.
4. Balasubramanian V. *Chemistry higher secondary – first year*. Tamilnadu Textbook Corp., College Road, Chennai, 2004, no. 1, Unit 10.
5. Feynman R., Leighton R.B., Sands M. *The Feynman Lectures on Physics*. Addison Publ., 1969, no. 1, pp. 47–1–47–8.
6. Li Tan. *Digital Signal Processing Fundamentals and Applications*. Academic Press, Elsevier, 2008, pp.105–152.
7. Shung K.K., Thieme G.A. *Ultrasonic Scattering in Biological Tissues*. CRC Press, America, 1993, 512 p.
8. Bashkatov N.A. Optical Properties of the Subcutaneous Adipose Tissue in the Spectral Range 400–2500 nm. *Optics and Spectroscopy*. 2005, vol. 99, iss. 5, pp. 836–842.
9. Khalil O.S. Spectroscopic and Clinical Aspects of Noninvasive Glucose Measurements. *Clinical Chemistry*. 1999, vol. 45, iss. 2, pp. 165–177.
10. Amerov K., Chen J., Arnold M.A. Molar Absorptivities of Glucose and Other Biological Molecules in Aqueous Solutions over the First Overtone and Combination Regions of the Near-Infrared Spectrum. *Applied Spectroscopy*. 2004, no. 58, pp. 1199–1201.
11. Kasemsumran S. *Analytica Chimica Acta*. 2004, vol. 512, pp. 223–230.
12. Brown Th.E., Eugene H., LeMay H. *Chemistry The Central Science*. Prentice Hall Publ., Pearson, 2011, 89 p.
13. Zumdahl S., DeCoste D.J. *Chemical principles*. Brooks/Cole, Davis Drive Belmont, 7<sup>th</sup> ed., 2011, pp. 865–866.
14. Strega M.A., Lagu A.L. *Capillary Electrophoresis of Proteins and Peptides*. Human Press, NJ, 2004, 332 p.

15. Rogers K. *The components of Life from Nucleic Acids to Carbohydrats*. Britanica Educational Publ., NY, 1<sup>st</sup> ed., 2011, 242 p.
16. Brown L., Holme T. *Chemistry for Engineering*. Cengage Learning, 2013, 608 p.
17. Shaw A.R., Mantsch H.H. *Encyclopedia of Analytical*

*Chemistry*. John Wiley & Sons Ltd. Publ., Chichester, 2000, pp. 83–102.

18. Faller A., Schuenke M. *The Human Body: An Introduction to Structure and Function*. Georg Thieme Verlag Publ., Germany, 2004, 720 p.

DOI: 10.15827/0236-235X.115.090-095

Дата подачи статьи 14.06.16

## ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ МУЛЬТИСПЕКТРАЛЬНОГО ФОТОАКУСТИЧЕСКОГО ИЗОБРАЖЕНИЯ ДЛЯ ГЕМАТОЛОГИЧЕСКОГО АНАЛИЗА

**Чинь Куанг Дык**, доктор наук, [duc.trinhquang@hust.edu.vn](mailto:duc.trinhquang@hust.edu.vn)

(Ханойский университет науки и технологий, ул. Дай Ко Вьет, 1, Ханой, Вьетнам);

**Нгуен Хоай Жанг**, доктор наук, [giangnh@hou.edu.vn](mailto:giangnh@hou.edu.vn)

(Ханойский открытый университет, ул. Нгуен Хиен, Б101, Ханой, Вьетнам)

**Аннотация:** Статья посвящена исследованиям возможности применения фотоакустической томографии и фотоакустической спектроскопии для неинвазивного клинического анализа крови. Для определения параметров, получаемых перспективными системами анализа крови, была разработана программа симуляции в среде Matlab. Оценка работоспособности предлагаемой техники проводилась на примере измерения глюкозы. Предположив, что глюкоза, карбамид, альбумин и  $\gamma$ -глобулин равномерно растворены в воде, была определена концентрация глюкозы в смеси при помощи возбуждающего источника света мощностью 300 мВт и выполнен анализ фотоакустического сигнала по методу Фурье. По результатам симуляции концентрация глюкозы под возбуждающим источником света составила 2,8 г/л при соблюдении следующих условий: фотоакустическое давление – 2.27 Па, высота мягких тканей – 10 мм, диаметр кровеносного сосуда – 1 мм.

**Ключевые слова:** частота собственных колебаний, фотоакустическая спектроскопия, измерение концентрации.

### Литература

1. Intern/ Diabetes Foundation: Diabetes Atlas. Available at: <http://www.diabetesatlas.org/> (accessed April 4, 2014).
2. Diabetes Fact sheet N°312. WHO. 2013. Available at: <http://www.who.int/mediacentre/factsheets/fs312/en/> (accessed March 25, 2014).
3. Wang L.V., Wu H. *Biomedical Optics: Principles and Imaging*. Wiley-Interscience Publ., Hoboken, NJ, 2007.
4. Balasubramanian V. *Chemistry higher secondary – first year*. Tamilnadu Textbook Corp., College Road, Chennai, 2004, no. 1, Unit 10.
5. Feynman R., Leighton R.B., Sands M. *The Feynman Lectures on Physics*. Addison Publ., 1969, no. 1, pp. 47–1–47–8.
6. Li Tan. *Digital Signal Processing Fundamentals and Applications*. Academic Press, Elsevier, 2008, pp. 105–152.
7. Shung K.K., Thieme G.A. *Ultrasonic Scattering in Biological Tissues*. CRC Press, America, 1993, 512 p.
8. Bashkatov N.A. Optical Properties of the Subcutaneous Adipose Tissue in the Spectral Range 400–2500 nm. *Optics and Spectroscopy*. 2005, vol. 99, iss. 5, pp. 836–842.
9. Khalil O.S. Spectroscopic and Clinical Aspects of Noninvasive Glucose Measurements. *Clinical Chemistry*. 1999, vol. 45, iss. 2, pp. 165–177.
10. Amerov K., Chen J., Arnold M.A. Molar Absorptivities of Glucose and Other Biological Molecules in Aqueous Solutions over the First Overtone and Combination Regions of the Near-Infrared Spectrum. *Applied Spectroscopy*. 2004, no. 58, pp. 1199–1201.
11. Kasemsumran S. *Analytica Chimica Acta*. 2004, vol. 512, pp. 223–230.
12. Brown Th.E., Eugene H., LeMay H. *Chemistry The Central Science*. Prentice Hall Publ., Pearson, 2011, 89 p.
13. Zumdahl S., DeCoste D.J. *Chemical principles*. Brooks/Cole, Davis Drive Belmont, 7<sup>th</sup> ed., 2011, pp. 865–866.
14. Strega M.A., Lagu A.L. *Capillary Electrophoresis of Proteins and Peptides*. Human Press, NJ, 2004, 332 p.
15. Rogers K. *The components of Life from Nucleic Acids to Carbohydrats*. Britanica Educational Publ., NY, 1<sup>st</sup> ed., 2011, 242 p.
16. Brown L., Holme T. *Chemistry for Engineering*. Cengage Learning, 2013, 608 p.
17. Shaw A.R., Mantsch H.H. *Encyclopedia of Analytical Chemistry*. John Wiley & Sons Ltd. Publ., Chichester, 2000, pp. 83–102.
18. Faller A., Schuenke M. *The Human Body: An Introduction to Structure and Function*. Georg Thieme Verlag Publ., Germany, 2004, 720 p.

УДК 681.518

Дата подачи статьи: 13.04.16

DOI: 10.15827/0236-235X.115.096-100

**ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ ПРИ ВЫБОРЕ ЛЕКАРСТВЕННЫХ ПРЕПАРАТОВ НА ОСНОВЕ МЕТОДА АНАЛИЗА ИЕРАРХИИ***Д.В. Бутенко, к.т.н., доцент, gindenburg@mail.ru;**А.Н. Бутенко, д.х.н., профессор;**А.А. Большаков, аспирант**(Волгоградский государственный технический университет,  
просп. Ленина, 28, г. Волгоград, 400005, Россия)*

Статья посвящена решению проблемы безопасности применения лекарственных препаратов в аспектах повышения эффективности принятия клинических решений, снижения риска клинических ошибок и усиления обоснованности назначения препарата и его дозы. Суть проблемы состоит в том, что в современной фармакологической индустрии с огромным количеством однотипных лекарственных средств различных производителей, а также различного качества задача анализа множества препаратов с целью выбора наиболее безопасного из них является крайне сложной. По результатам проведенных предпроектных исследований описан ряд основных задач, которые должна решать автоматизированная система по данной проблематике. Разработана модель поддержки принятия решения для выбора наиболее безопасных лекарственных препаратов. Модель основана на методе анализа иерархии и позволяет учитывать индивидуальные особенности конкретного пациента. Показан обобщенный алгоритм выбора наиболее безопасного лекарственного средства. На этой основе разработана программная система.

**Ключевые слова:** модель, принятие решения, лекарственное средство, нежелательная побочная реакция, выбор лекарственного препарата.

Проблема безопасности лекарственных препаратов на сегодняшний день является одной из основных в клинической практике. Фармакологический рынок интенсивно развивается, количество лекарственных средств (ЛС) постоянно увеличивается. Вместе с расширением диапазона ЛС усложняется выбор наиболее эффективных, а также наиболее безопасных лекарственных препаратов. Врачу необходимо ежедневно читать в среднем около двадцати медицинских статей, чтобы быть в курсе актуальной информации по своей специальности, однако время, которым он располагает для чтения на работе, – менее одного часа в неделю. В связи с этим адекватно оценить безопасность препарата для конкретного пациента становится для врача все более и более сложной задачей и как следствие – затруднения в принятии клинических решений, повышение риска клинических ошибок, необоснованность назначения препарата. Одна из первостепенных проблем лекарственной терапии – правильный выбор лекарственного препарата и его дозы, так как именно ошибка выбора занимает лидирующее место среди врачебных ошибок. Опасность развития тяжелых, зачастую непоправимых осложнений вследствие неверной лекарственной терапии привлекает внимание все большего количества врачей и пациентов. Как правило, причиной ошибок при назначении лекарственного препарата является отсутствие у врача своевременной актуальной информации о назначаемом ЛС, а именно о его вновь выявленных побочных действиях.

Анализ врачебной деятельности показал, что при принятии решения о выборе и дозировке ЛС врач, как правило, руководствуется информацией из инструкции к средству, из настольных справочников, сетевых электронных банков данных, меди-

цинских изданий, но эта информация не успевает актуально обновляться в соответствии с регистрируемыми новыми побочными действиями ЛС. Таким образом, анализ информации о безопасности лекарственных препаратов целиком ложится на практикующего врача. Однако в силу современных масштабов фармакологической индустрии, огромного количества однотипных ЛС разных производителей различного качества задача анализа всего множества этих препаратов с целью выбора наиболее безопасного становится все более сложной для врача, а зачастую просто невыполнимой.

В данной работе для облегчения ее решения предлагается использовать созданную авторами информационно-аналитическую систему, которая обеспечивает врача поддержкой при принятии решения при выборе наиболее безопасного ЛС с учетом индивидуальных особенностей пациента. Информация, предоставляемая системой, должна быть в привычном для врача виде, то есть соответствовать современным стандартам описания и классификации ЛС, таким как анатомо-терапевтически-химическая классификация лекарств, классификация в соответствии с международной классификацией болезней, классификация препаратов по активному веществу (по основному химическому компоненту препарата), классификация по фармакологическим группам и т.д.

В процессе предпроектных исследований были проанализированы и учтены современные потребности в информации при ведении клинической практики и выделен ряд основных задач, решаемых автоматизированной системой:

- описание ЛС;
- классификация ЛС в соответствии с современными стандартами и методиками;

- сбор информации о вновь выявляемых побочных действиях ЛС;
- описание различных нежелательных лекарственных взаимодействий;
- описание существующих схем лечения, а также выбор наиболее оптимальной из них;
- поддержка принятия решения при выборе ЛС.

Таким образом, основной целью работы по созданию автоматизированной системы являлось применение методов теории принятия решения для поддержки выбора препарата. Разработанная методика актуальна при автоматизации задач выбора врачом наиболее безопасного ЛС для конкретного пациента в условиях информатизации общества и врачебной практики в частности.

Принятие врачебного решения традиционно ассоциируют с процессами поиска и связанным с ними перебором вариантов лечения – подбором ЛС. После анализа работы с привлечением экспертов [1] был обобщен и формализован процесс принятия врачебного решения, который можно представить в виде следующей последовательности:

- описание состояния пациента (проведение анализов, постановка диагноза);
- анализ ограничений (наличие аллергии на компоненты препарата, нарушение функции почек и т.п.);
- выбор метода лечения/лекарственного препарата;
- прогноз вариантов результата лечения (оценка безопасности препарата, возможности возникновения нежелательных побочных реакций).

Более подробно следует рассмотреть такую задачу системы, как поддержка принятия решения. Она основана на трех основных принципах: контроль допустимой дозы назначаемого препарата, контроль нежелательных лекарственных взаимодействий между назначаемым препаратом и препаратами, принимаемыми совместно или ранее (на основе данных об известных лекарственных взаимодействиях), анализ безопасности применения назначаемого препарата на основе данных о случаях нежелательных побочных действий, вызванных им.

Наибольший интерес представляют задача анализа безопасности применения назначаемого препарата, а также задача выбора наиболее безопасного из некоторого множества доступных препаратов. Для решения этой задачи авторы предлагают концепцию, которая основана на взаимодействии факторов, влияющих на проявление побочного действия ЛС. Предлагается информационная модель, с помощью которой возникновение побочного действия представляется как результат взаимодействия таких факторов, как назначаемая длительность и доза применения препарата, лечимая болезнь, препараты, применяемые совместно, а также индивидуальные особенности пациента (воз-

раст, пол, вес, сопутствующие болезни, различные аллергические показатели и т.д.). В системе ведется сбор информации о случаях возникновения побочных действий при применении ЛС, а также значений факторов, в результате взаимодействия которых они проявились. При помощи системы во время регистрации случая нового побочного действия устанавливается степень достоверности связи «лекарство–побочное действие» посредством различных оценочных шкал (шкала Karch F., Lasagna L. [2] или алгоритм Наранжо [3]). В дальнейшем при анализе безопасности применения и выборе наиболее безопасного препарата, используя эту оценку, можно будет относиться с разной степенью доверия к каждому зарегистрированному случаю побочного действия.

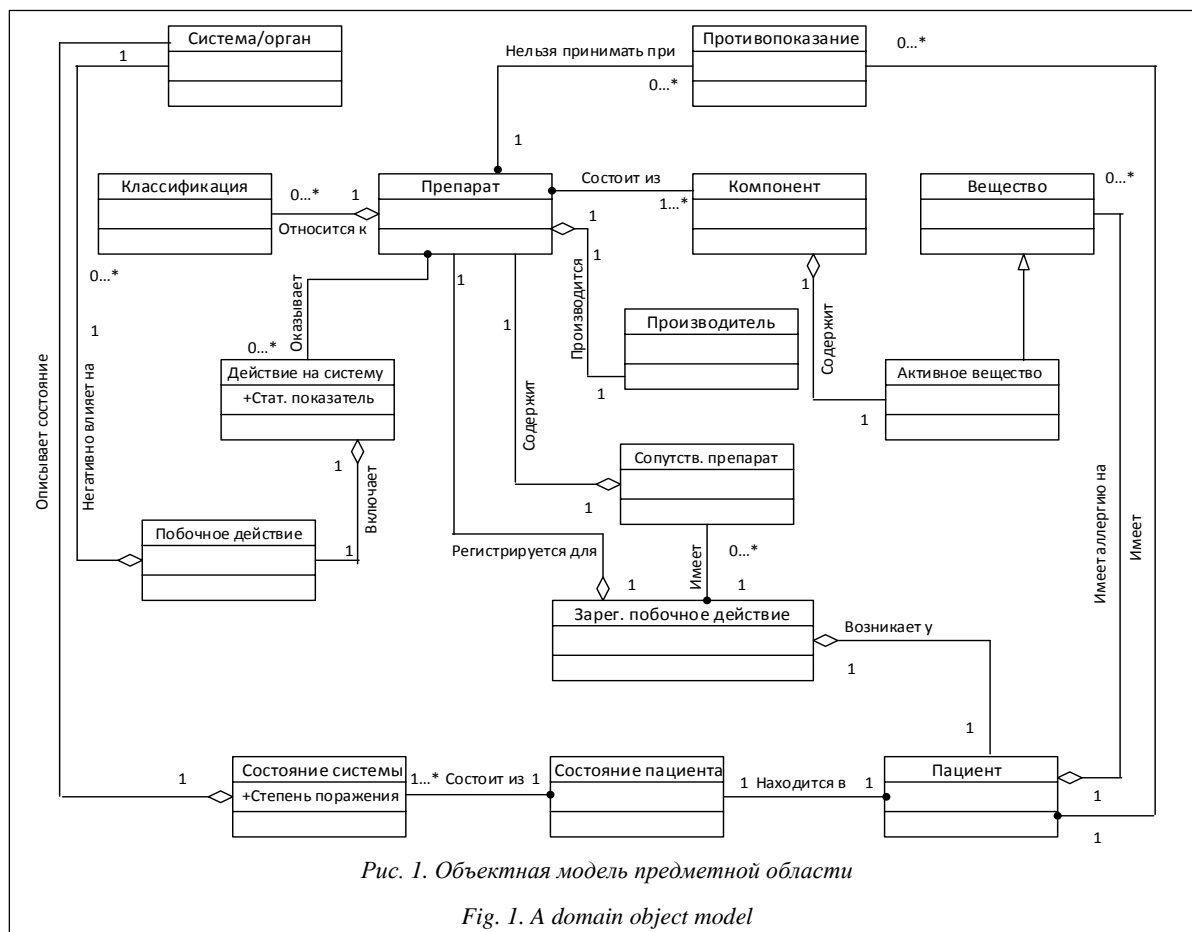
С точки зрения теории принятия решений процесс сводится к определению множества критериев выбора, множества альтернатив, алгоритма выбора альтернатив согласно заданным критериям. В задаче выбора наиболее безопасного ЛС можно выделить следующие критерии:

- диагноз, при котором назначается ЛС (согласно международной классификации болезней);
- фармакологическая группа ЛС;
- присутствие аллергических реакций на компоненты ЛС.

По данным критериям выполняется отбор подмножества альтернатив – ЛС из всего множества для дальнейшего выбора наиболее подходящего из них. В качестве критериев, по которым осуществляется окончательный выбор наилучшей альтернативы ЛС, предлагается использовать вероятности возникновения нежелательного влияния препарата на органы или системы организма пациента. Эти критерии могут иметь различную значимость и выбираются врачом-экспертом для каждого конкретного пациента после проведения соответствующей диагностики [4]. Например, для пациента с нарушением функции почек и печени критерии, соответствующие этим системам и органам, будут иметь наивысший приоритет среди остальных, так как даже малое поражение этих систем вследствие применения ЛС для данного пациента недопустимо.

Учитывая основные особенности предметной области, выполнена ее формализация и построена объектная модель при помощи унифицированного языка моделирования – UML [5]. Результат моделирования изображен на рисунке 1. Объектная модель позволяет учитывать при поддержке принятия решения параметры лекарственных препаратов с целью оценки их безопасности.

За основу модели поддержки принятия решения выбран метод анализа иерархии, так как он позволяет отбирать альтернативы по нескольким критериям, важность которых различна [6, 7]. Общая схема иерархии задачи поддержки принятия решения при выборе наиболее безопасного ЛС изображена на рисунке 2.



На первом уровне иерархии находится цель поддержки принятия решения – выбор наиболее безопасного ЛС, на втором уровне – критерии выбора ( $O1-O_n$ ), которые описывают состояние органов и систем организма конкретного пациента, где  $B1-B_n$  – важность этих критериев в зависимости от состояния тех или иных органов.

На третьем уровне находятся альтернативы (ЛС)  $D1-D_n$ ;  $A_{ji}$  – оценки вероятности нежелательного влияния этих ЛС на системы и органы организма ( $O1-O_n$ ), которые устанавливаются экспертами – клиническими фармакологами для каждого ЛС в ходе клинических испытаний или практики.

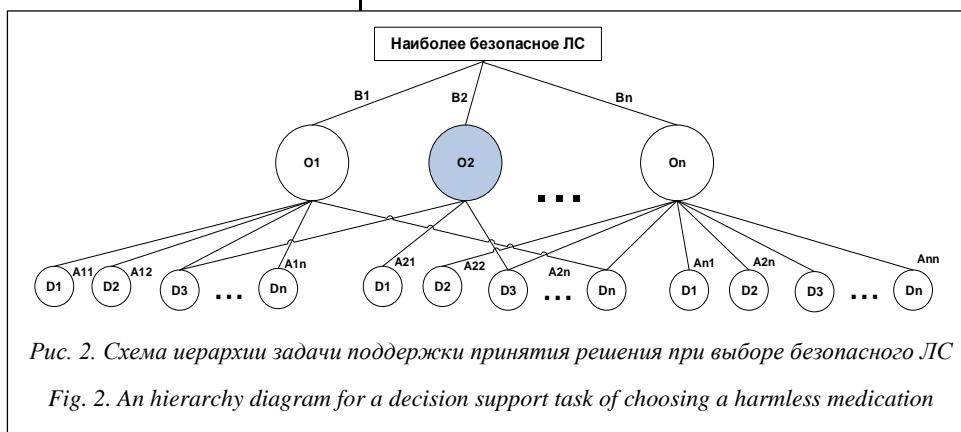
В общем виде алгоритм выбора наиболее безопасного ЛС согласно описанной модели поддержки принятия решения состоит из следующих этапов:

1) оценка вероятности нежелательного влияния  $A_{ji}$  для каждого препарата – альтернативы на каждую систему и органы организма;

2) расстановка критериев ( $O_n$ ) и их важности ( $B_n$ ) для каждого органа и системы организма конкретного пациента;

3) построение матрицы парных сравнений для критериев  $O_n$ , столбцы и строки матрицы – критерии  $O_n$ , значения элементов матрицы отношения значимости критериев  $B_{ni}/B_{nj}$ , если  $B_{ni} \geq B_{nj}$ , и  $B_{nj}/B_{ni}$  в противном случае;

4) построение матрицы парных сравнений альтернатив для каждого критерия  $O_n$  второго уровня, в которых сравниваются значения  $A_{ij}$  каждой альтернативы; столбцы и строки матрицы – альтернативы  $D_n$ , значения элементов матрицы отношения





$Ani/Anj$ , если  $Ani \geq Anj$ , и  $Anj/Ani$  в противном случае;

5) нахождение для каждой матрицы вектора приоритетов  $Q = (W1, ..., Wn)$ , где  $Wi$  рассчитывается по формуле  $Wi = q_i/r$ ,  $q_i = \sqrt[n]{a_{i1} \cdot a_{i2} \cdot ... \cdot a_{in}}$  и  $r = q_1 + ... + q_n$ ,  $a_i$  – значение матрицы  $i$ -й строки  $n$ -го столбца;

6) расчет показателя безопасности  $Sj$  для каждой альтернативы препарата:  $S_j = \sum_{i=1}^n W_i \cdot V_{ji}$ , где  $Wi$  – вес  $i$ -го критерия, а  $V_{ji}$  – важность  $j$ -й альтернативы по  $i$ -му критерию;

7) сортировка альтернатив по возрастанию значения  $Sj$ .

Описанная методика поддержки принятия решения используется в разработанной автоматизированной системе поддержки принятия врачебного решения по выбору ЛС.

В ходе проектирования программной системы рассмотрены принципы построения и особенности реализации банка данных лекарственных препаратов и подсистемы поддержки принятия решения при выборе ЛС. На рисунке 3 представлены концептуальная структура банка данных лекарственных препаратов и место подсистемы поддержки принятия решения (подсистема анализа безопасности и выбора препарата) в его структуре. Банк данных состоит из трех основных слоев: слой хранения данных о препаратах, веществах, других объектах предметной области; слой отображения данных, представляющий собой пользовательский интерфейс – набор необходимых элементов управления; слой бизнес-логики, в котором находится специфичная для каждой подсистемы логика.

В результате проектирования программной среды была достигнута цель, заключающаяся в повышении обоснованности выбора ЛС за счет обеспечения автоматизированной поддержки принятия решения при выборе ЛС с учетом множества факторов. Выполнен следующий комплекс работ:

- исследованы современные методы описания и классификации ЛС и их побочных действий, критериев безопасности ЛС;
- проведен анализ современных информационных систем и банков данных по лекарственным препаратам, выявлены основные недостатки;
- формализована предметная область и построена ее модель;
- в качестве метода поддержки принятия решения выбран метод анализа иерархии и адаптирован под задачу выбора лекарственного препарата;
- спроектирован и разработан банк данных ЛС с подсистемой поддержки принятия решения с использованием разработанной модели предметной области и адаптированного метода анализа иерархии.

Таким образом, при назначении лекарственного препарата пациенту, указав в системе значения факторов, влияющих на возникновение побочного действия, можно получить либо статистические сведения о случаях возникновения побочных действий данного препарата при таких же значениях факторов или близких к ним, либо конкретные предостережения со стороны системы с указанием наиболее вероятного побочного действия или действий данного препарата. Имея возможность оценить безопасность применения сразу для нескольких ЛС с учетом индивидуальных особенностей

пациента, врач может принять объективное решение при выборе наиболее безопасного препарата и тем самым снизить риск возникновения нежелательных побочных реакций.

Используя предложенную концепцию, можно выбрать наиболее безопасный препарат для конкретного пациента, основываясь на статистике случаев возникновения побочных действий ЛС. Таким образом, статистика случаев возникновения побочных действий препаратов, анализ роли факторов, повлиявших на эти действия, коэффициент достоверности каждого из случаев побочных действий предоставляют поддержку для принятия решения о выборе наиболее безопасного препарата.

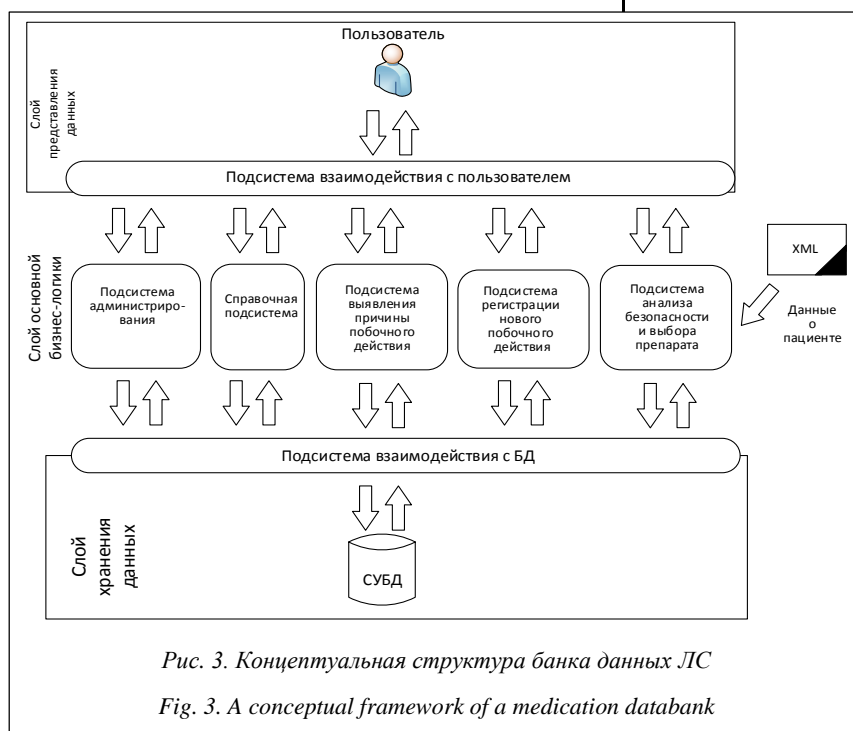


Рис. 3. Концептуальная структура банка данных ЛС

Fig. 3. A conceptual framework of a medication databank

Данная система, внедренная в лечебные учреждения, предоставит врачу комплекс необходимой информации для выбора наиболее безопасного ЛС, а также наиболее оптимальной схемы лечения с учетом индивидуальных особенностей пациента.

Точность результатов, получаемых при анализе системой статистической информации о новых побочных действиях ЛС, напрямую зависит от количества поступающей информации о случаях регистрации побочных реакций, поэтому внедрение данной системы целесообразно проводить централизованно на областном или на федеральном уровне. То есть данная информационно-аналитическая система должна обладать единой БД, информацию из которой врачи могли бы использовать в любое время.

Авторы считают, что новым в данной работе является предлагаемая концепция обеспечения поддержки принятия врачебного решения, а также метод анализа безопасности ЛС, использование которых поможет врачу принимать решение в сложной ситуации в условиях недостатка информации о клинической практике применения того или иного лекарственного препарата.

Внедрение данной информационно-аналитической системы позволит снизить риск причинения вреда здоровью пациента при назначении ЛС, по-

высить уровень фармакологического контроля и сократить расходы на лекарственную терапию.

### Литература

1. Большаков А.Л., Бутенко Д.В. Применение метода анализа иерархии для поддержки принятия решения при выборе лекарственных препаратов // Программные продукты и системы. 2010. № 2. С. 148–149.
2. Edwards I.R., Lindquist M., Wiholm B.E., et al. Quality criteria for early signals of possible adverse drug reactions. *Lancet*, 1990, vol. 336, pp. 156–158.
3. Астахова А.В., Лепяхин В.К. Неблагоприятные побочные реакции и контроль безопасности лекарственных средств. М.: Когито-Центр, 2004. 200 с.
4. Хомская Е.Д. X76 Нейропсихология: Сер. «Классический университетский учебник». 4-е изд. СПб: Питер, 2005. 496 с.
5. Ларман К. Применение UML 2.0 и шаблонов проектирования. М.: Вильямс, 2006. 3-е изд. 736 с.
6. Ларичев О.И. Теория и методы принятия решений, а также хроника событий в волшебных странах. М.: Логос, 2000. 296 с.
7. Саати Т.Л. Принятие решений при зависимостях и обратных связях: Аналитические сети. М.: Изд-во ЛКИ, 2008. 360 с.
8. Большаков А.Л., Бутенко Д.В. Информационно-аналитическая система поддержки принятия решения при выборе лекарственного средства // Вестн. новых медицинских технологий. Тематич. вып. Избранные технологии диагностики и лечения. 2009. № 1. С. 60–61.
9. Бутенко Д.В., Ананьев А.С., Попов К.В. Методика концептуального проектирования программных информационных систем // Программные продукты и системы. 2012. № 2. С. 101–104.

DOI: 10.15827/0236-235X.115.096-100

Received 13.04.16

### DECISION SUPPORT WHEN CHOOSING MEDICATIONS BASED ON THE HIERARCHY ANALYSIS METHOD

<sup>1</sup>*Butenko D.V., Ph.D. (Engineering), Associate Professor, gindenburg@mail.ru*

<sup>1</sup>*Butenko L.N., Dr.Sc. (Dr.Sc. (Economics)), Professor*

<sup>1</sup>*Bolshakov A.L., Postgraduate Student*

<sup>1</sup>*Volgograd State Technical University, Lenin Ave. 28, Volgograd, 400005, Russian Federation*

**Abstract.** The article is devoted to solving a problem of safe medicine usage in the aspects of increasing the efficiency of clinical decision-making, reducing the risk of clinical errors and enhancing the validity of drug prescription and its dose. The key part of the problem is that the safety analysis of many drugs to choose the safest one is extremely complicated task in modern pharmacological industry with a huge number of similar drugs from different manufacturers and of different quality. According to the initial investigation results the authors describe a number of key tasks, which should be solved by a specialized automated system. A decision-making support model for choosing the safest drugs has been developed. The model is based on the hierarchy analysis method. The developed model allows taking into account the individual characteristics of a particular patient. The article shows a generalized algorithm for choosing the safest drug. It is a basis for the developed software system.

**Keywords:** model, decision making, medicine, undesirable side reaction, drug choice.

### References

1. Bolshakov A.L., Butenko D.V. Use of the analytic hierarchy for decision support when selecting medications. *Programmnye produkty i sistemy* [Software and Systems]. 2010, no. 2, pp. 148–149 (in Russ.).
2. Edwards I.R., Lindquist M., Wiholm B. E. *Quality criteria for early signals of possible adverse drug reactions*. *Lancet Publ.*, 1990, vol. 336, pp. 156–158.
3. Astakhova A.V., Lepakhin V.K. *Neblagopriyatnye pobochnye reaktsii i kontrol bezopasnosti lekarstv* [Adverse Drug Reactions and Monitoring Drug Safety]. Moscow, 2004, 200 p.
4. Chomskaya E.D. *X76 Neyropsikhologiya* [X76 Neuropsychology]. 4th ed. St. Petersburg, Peter Publ., 2005, 496 p.
5. Larman C. *The application of UML 2.0 and design patterns: an introduction to object-oriented analysis and design and iterative development*. 3rd ed., Addison Wesley Pro. Publ., 2004, 736 p. (Russ.ed.: Moscow, Williams Publ., 2006, 736 p.)
6. Larichev O.I. *Teoriya i metody prinyatiya resheny, a takzhe khronika sobyty v volshebnykh stranakh* [Theory and methods of decision-making as well as a chronicle of events in magic countries]. Moscow, Logos Publ., 2000, 296 p.
7. Saaty T.L. *Prinyatie resheny pri zavisimostyakh i obratnykh svyazyakh: Analiticheskie seti* [Decision-Making with Dependencies and Feedbacks: Analytical Networks]. Moscow, LKI Publ., 2008, 360 p.
8. Bolshakov A.L., Butenko D.V. Information-analytical decision support system when selecting medications. *Vestn. novykh meditsinskikh tekhnologiy. Tematichesky vyp. "Izbrannye tekhnologii diagnostiki i lecheniya"* [Journ. of New Medical Technologies. Special Iss. "Selected Technologies for Diagnosis and Treatment"]. 2009, no. 1, pp. 60–61 (in Russ.).
9. Butenko D.V., Ananyev A.S., Popov K.V. A conceptual design of software information systems. *Programmnye produkty i sistemy* [Software and Systems]. 2012, no. 2, pp. 101–104 (in Russ.).



УДК 630.43

Дата подачи статьи: 14.04.16

DOI: 10.15827/0236-235X.115.101-108

## **АГЕНТНЫЙ ПОДХОД ПРИ МОДЕЛИРОВАНИИ ЛЕСОПОЖАРНЫХ СИТУАЦИЙ**

*С.В. Яровой, аспирант, ach\_bask@mail.ru*

*(Сибирский государственный технологический университет,  
просп. Мира, 82, г. Красноярск, 660049, Россия)*

В статье рассматривается задача моделирования лесопожарных ситуаций. Для ее решения предлагается использовать агентный подход совместно с технологией геоинформационных систем, что позволит объединить преимущества обоих направлений. Разработана мультиагентная модель, дающая возможность как прогнозировать распространение кромки лесного пожара, так и имитировать прямой метод тушения низового лесного пожара немеханизированными противопожарными отрядами.

В качестве среды для агентов может быть выбран любой участок карты реальной местности. В предлагаемой модели используются два типа агентов. Агенты первого типа (А) представляют собой участки фронта распространения пожара и используются для моделирования поведения лесного пожара. В совокупности они представляют контур пожара на карте. Модель допускает наличие одновременно нескольких таких контуров, распространяющихся независимо друг от друга в различных частях карты. Агенты второго типа (В) имитируют действия противопожарных сил и оказывают воздействие на агентов первого типа.

На основе предложенной модели был разработан программный комплекс, который предлагается использовать для решения широкого круга задач. Во-первых, система будет полезна для оперативного определения параметров распространения лесного низового пожара в лесных насаждениях, а также разработки комплекса мер по его предотвращению и по противодействию ему. Во-вторых, она может служить основой для учебно-тренажерных систем, предназначенных для обучения студентов вузов и техникумов лесохозяйственного профиля, работников лесохозяйственной отрасли и МЧС на курсах повышения квалификации в учебно-игровой форме основам тактики борьбы с лесными пожарами.

**Ключевые слова:** природный пожар, локализация природного пожара, имитационное моделирование, геоинформационная система, агентная модель, программный комплекс.

Природный пожар – неконтролируемый процесс горения, стихийно возникающий и распространяющийся в природной среде. Данный термин объединяет лесные, торфяные и степные пожары. Природные пожары наносят огромный экономический и экологический ущерб регионам, в которых они возникают. Для управления эффективной борьбой с природными пожарами разработано большое количество информационных систем, позволяющих с различной степенью точности моделировать их поведение (системы BehavePlus, FARSITE, WFDS [1–3] и др.). Данные системы основаны на математических моделях, имеющих различную точность, время расчета и требовательность к вычислительным ресурсам.

Однако, помимо моделирования непосредственно самой кромки природного пожара, существует необходимость моделирования различного рода искусственных воздействий на пожар для его локализации и ликвидации. Подобные модели могут активно использоваться при управлении лесопожарной ситуацией, в частности, для решения задачи оптимального управления противопожарными силами и средствами. На практике выбор средств и методов для локализации природного пожара зависит от большого количества факторов, которые практически не поддаются формализации. В таких условиях на данный момент наиболее эффективным методом является агентное моделирование.

Агентное моделирование (agent-based modeling, ABM) представляет собой новый подход к имита-

ционному моделированию, который направлен на моделирование сложных динамических систем с помощью исследования поведения и взаимодействия автономных и независимых субъектов (агентов) в некоторой среде [4]. Агентные модели используются как в естественных науках (физике, биологии), так и в общественных (социологии, экономике). Наибольшее применение данный тип моделей получил в случаях, когда необходимо продемонстрировать работу системы, состоящей из большого числа взаимосвязанных субъектов (от десятков до нескольких тысяч), зависящих от влияния большого количества различных случайных воздействий.

Для моделирования лесопожарных ситуаций особый интерес также представляют *геоинформационные системы* (ГИС), обладающие возможностями хранения и визуализации пространственной информации. Совместное использование агентного моделирования и ГИС позволяет объединить преимущества обоих направлений, применить перспективные средства для моделирования процессов, протекающих в условиях реального ландшафта.

Идея моделирования кромки природного пожара с помощью агентного подхода не нова [5, 6]. Агентные модели такого рода создавались с помощью специализированных систем для программного описания деятельности агентов и состояния среды (агентных платформ), таких как NetLogo, StarLogo, Repast Simphony и др. Данный подход позволяет упростить программную реализацию

имитационной модели, но значительно затрудняет ее интеграцию с ГИС. Кроме того, в работах [5, 6] моделирование свободно распространяющегося пожара производится без моделирования противопожарных сил и средств.

Таким образом, существует необходимость создания имитационной системы на основе агентного подхода, позволяющей моделировать лесопожарные ситуации на реальной местности, включающие как моделирование пожара, так и процесс его ликвидации. Рассмотренная далее агентная модель ориентирована на имитацию прямого метода тушения низового лесного пожара немеханизированными противопожарными отрядами.

### Описание среды моделирования

Среда моделирования представляет собой участок карты реальной местности с нанесенным на нее слоем, описывающим *растительные горючие материалы* (РГМ). Таким образом, имеется возможность совмещать данные ГИС и агентной модели. В модели используется информация о населенных пунктах, реках, озерах, дорогах, оврагах, лесных насаждениях и др. Позиционирование на карте производится заданием пространственных координат, различных для каждого объекта. В качестве модели горючего используется модель тонкого слоя [7]. В данной модели лесной покров рассматривался как однородный слой на поверхности земли с усредненными характеристиками РГМ, определяемыми на основании пирологического описания выделов карты лесов.

Таким образом, для каждого местоположения на карте заранее задается тип поверхности, определяющий наличие *основных проводников горения* (ОПГ) и их тип. В зависимости от типа ОПГ определяются базовая скорость распространения кромки пожара  $V_{\text{ОПГ}}$  (м/мин.) и теплота воспламенения слоя  $Q$  (Дж/м<sup>2</sup>).

В качестве основных показателей среды также принимаются следующие параметры:

- величина скорости ветра по данным ближайшей метеостанции  $V_{\omega}$ ;
- вектор направления ветра  $W$ ;
- класс пожарной опасности.

Данные параметры в значительной степени влияют на скорость и направление распространения пожара, а также на интенсивность горения.

В модели выделены два типа агентов: участок фронта распространения пожара (агенты типа А) и противопожарная команда (агенты типа В).

Агенты типа А используются для моделирования поведения лесного пожара и в совокупности представляют контур пожара на карте. Модель допускает наличие одновременно нескольких таких контуров, распространяющихся независимо друг от друга в различных частях карты.

Агенты типа В имитируют действия противопожарных сил и оказывают воздействие на агентов типа А.

### Свойства и характеристики агентов типа А

Агент, представляющий участок фронта распространения пожара, характеризуется следующими параметрами:

- текущие пространственные координаты  $(X, Y)$ ;
- вектор внешней нормали к контуру  $n = [n_x, n_y]^T$ ;
- вектор скорости движения  $v = [v_x, v_y]^T$ ;
- значение интенсивности горения  $I$ ;
- значение уровня задымления  $M$ ;
- ссылки  $L$  и  $R$  на соседних агентов для создания единого контура;
- текущее состояние горения  $onFire = [true, false]$ .

Основная цель, которой агенты А руководствуются при исполнении, – перемещение по среде моделирования. Расчет пространственных координат данных агентов на каждом шаге моделирования осуществляется по заранее заданному закону. Также во время исполнения модели агенты А могут порождать новых агентов или погибать при определенных условиях. Область видимости агентов А ограничена расстоянием до соседних по контуру агентов за счет ссылок  $L$  и  $R$ .

**Расчет пространственных координат.** Расчет новых координат  $A_x, A_y$  на каждой итерации моделирования выполняется посредством основанного на *методе подвижных сеток* (МПС) алгоритма [8, 9]. Согласно данному методу, расчетная сеточная область не задается заранее, а продвигается и развивается вместе с контуром пожара, представленным в виде набора точек:  $C_i(t) = \{x_i(t), y_i(t), p_i(t), L(i), R(i)\}$ ,  $i \in N(t)$ , где  $x_i(t), y_i(t)$  – координаты точки на плоскости;  $p_i(t)$  – вектор внешней нормали к контуру;  $L(i)$  и  $R(i)$  – номера точек контура, расположенных, соответственно, слева и справа от  $C_i$ ;  $N(t)$  – множество номеров агентов, образующих контур в момент  $t$ .

Для расчета следующего состояния контура используются скорость распространения пожара  $v_0$  и индикатриса нормальной скорости  $\xi_n(\varphi, \omega)$ .

Приращение координат в точке  $C_i(t)$  производится по формулам:

$$x_i(t+\Delta t) = x_i(t) + v_0 \cdot \xi_n(\varphi_i, \omega) \cdot \Delta t \cdot \cos(\alpha_i),$$

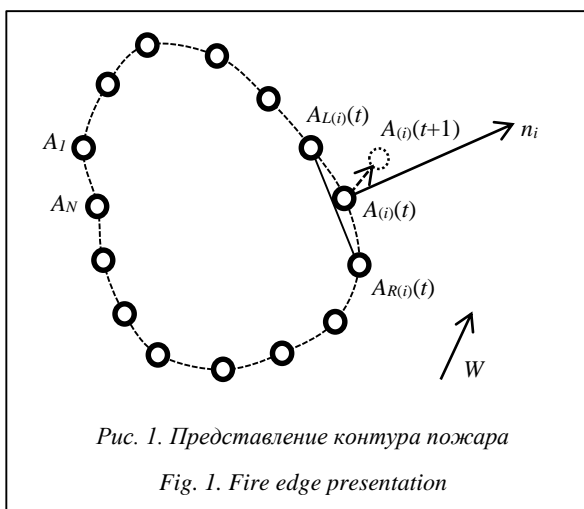
$$y_i(t+\Delta t) = y_i(t) + v_0 \cdot \xi_n(\varphi_i, \omega) \cdot \Delta t \cdot \sin(\alpha_i),$$

где  $\varphi_i$  – угол между нормалью к контуру в точке  $C_i(t)$  и направлением ветра  $\omega$ ;  $\alpha_i$  – угол между нормалью к контуру в точке  $C_i(t)$  и горизонтальной осью.

Для использования данного алгоритма сопоставили каждую точку  $C_i(t)$  с агентами типа А в нашей модели. В качестве входных данных для алгоритма используются характеристики агентов  $(X, Y, n, v, L, R)$ , а также параметры среды, такие как скорость и направление ветра, класс пожарной опасности, тип

ОПГ. Помимо расчета приращений координат для точек контура, алгоритм МПС предусматривает анализ пересечений и самопересечений контуров, а также упорядочение расстояний между соседними точками. Внедрение данных процессов в модель позволило эффективно управлять популяцией агентов  $A$  (добавлять и удалять агентов).

На рисунке 1 визуально представлен контур пожара при расчете, а также схематично указано положение агента  $A_i$  на разных итерациях процесса моделирования  $t$  и  $t+1$ . Стрелка  $n_i$  – это направление нормали к контуру для  $i$  агента,  $i \in N(t)$ . Вектор  $W$  – направление ветра.



Помимо расчета параметров агента  $A$ , характеризующих его пространственное положение в среде, во время моделирования производится расчет интенсивности горения и уровня задымления вокруг агента. Значения данных параметров имеют решающее значение при взаимодействии с агентами  $B$ .

**Расчет интенсивности горения РГМ.** Согласно [10], интенсивность  $I$  рассчитывается по формуле  $I = Q \cdot v_i \cdot \rho_{be}$ , где  $v_i$  – значение скорости распространения пламени, м/мин.;  $\rho_{be}$  – эффективная плотность горючего, кг/м<sup>3</sup>, то есть количество горючего материала (из общего запаса, содержащегося в слое), которое участвует в воспламенении;  $Q$  – теплота воспламенения (количество тепла, необходимое для доведения единицы веса горючего материала до воспламенения), Дж/кг.

Поскольку в предлагаемой модели используется модель тонкого слоя горючего, была использована эмпирическая формула, в которой в качестве  $Q$  используется теплота сгорания слоя, МДж/м<sup>2</sup> [11].

Таким образом, интенсивность горения агента  $A_i$ , на которого не оказывается воздействие агентом  $B$ , при моделировании рассчитывается следующим образом:  $I_i = (Q \cdot v_i + I_{R(i)} + I_{L(i)})/3$ , где  $I_i$  – интенсивность горения агента  $A_i$ ;  $Q$  – поверхностная теплота сгорания слоя;  $v_i$  – значение скорости движения

агента  $A_i$ ;  $I_{R(i)}$  – интенсивность горения агента  $A_{R(i)}$ ;  $I_{L(i)}$  – интенсивность горения агента  $A_{L(i)}$ .

То есть при расчете интенсивности горения текущего агента  $A_i$  учитывается аналогичный показатель соседних агентов.

Также на каждом шаге моделирования показатель интенсивности  $I_i$  может быть изменен случайным образом в положительную или отрицательную сторону с вероятностью  $P_{nf}$ . Это необходимо для имитации неоднородности таксационных выделов, ошибок в определении типов леса и других факторов, которые могут служить причиной ошибок прогноза [11].

**Расчет задымленности.** Задымление является одним из отрицательных факторов, увеличивающих ущерб от лесного пожара, а также значительно затрудняющих борьбу с ним. В настоящее время существует большое количество работ, посвященных моделированию дымового шлейфа [12, 13], однако реализация этих моделей требует высоких вычислительных затрат, а сами модели избыточны для данного исследования. Таким образом, была выбрана концепция локального пламени [8]. Согласно данной концепции, несколько огрубляя физическую картину распространения дымового облака, принимается, что каждое локальное пламя (в нашем случае каждый агент типа  $A$ ) порождает локальный дымовой поток, конфигурацию которого в приземном слое можно определить следующим образом:  $\lambda_s(x, y, t, \varphi_0) = v_{s0}(x, y, \omega) \cdot (t - t_0) \cdot \chi_s(\omega, \varphi_0 - \varphi)$ , где  $(x, y)$  – координаты точки в области активного горения;  $\varphi_0$  – направление ветра, рад;  $\varphi$  – текущий угол, рад,  $0 \leq \varphi \leq 2\pi$ ;  $t$  – текущее время, сек.,  $t_0 \leq t \leq t_0 + r$ ;  $t_0$  – время воспламенения горючего в точке  $(x, y)$ , сек.;  $r$  – продолжительность горения горючего в точке  $(x, y)$ ;  $\chi_s(\omega, \varphi_0 - \varphi)$  – индикатриса локального дымового потока, определяющая его форму;  $v_{s0}$  – горизонтальная скорость распространения дымового облака в приземном слое.

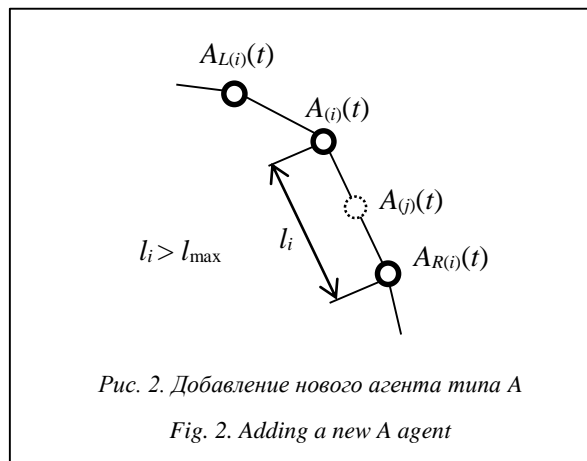
Согласно [13],  $v_{s0}$  вычисляется с помощью выражения  $v_{s0} = 0,28 + 0,607w - 0,0217w^2$ , где  $w$  – скорость ветра над пологом леса, м/с.

В предлагаемой модели на каждом шаге моделирования агенты типа  $A$  перемещаются в новую точку пространства, поэтому  $r$  всегда будет равна времени, эквивалентному одному шагу моделирования (при реализации модели использовался шаг = 5 мин.).

Сам уровень задымления  $M$  непосредственно около агента  $A$  принят в виде безразмерной величины, зависящей от интенсивности горения  $I$ , и может принимать значение от 0 (при  $I = 0$ ) до 1 (при  $I \rightarrow \max$ ).

**Добавление новых агентов в процессе моделирования.** В процессе моделирования агенты  $A$  перемещаются, увеличивая размер контура пожара в десятки и сотни раз. Расстояние между агентами также растет, что приводит к очень грубому представлению контура пожара. Таким образом, суще-

ствует необходимость упорядочения расстояния между соседними агентами  $A_i(t)$  и  $A_{R(i)}(t)$ ,  $i \in N(t)$ . Для этого рассчитывается расстояние  $l(i)$  между каждой парой агентов. Если  $l(i) > l_{\max}$ , порождается новый агент  $A_j(t)$ , который располагается между парой  $A_i(t)$  и  $A_{R(i)}(t)$ . В качестве характеристик нового агента  $A_j(t)$  записываются средние значения соответствующих характеристик агентов  $A_i(t)$  и  $A_{R(i)}(t)$ . Данный процесс схематично представлен на рисунке 2.



Значение  $l_{\max}$  является определяющим для точности расчета контура в обратно-пропорциональной зависимости. Однако слишком малое значение  $l_{\max}$  приведет к порождению большого числа агентов, что значительно увеличит требования к вычислительной мощности при моделировании. При внешней границе карты  $\approx 10$  км автором рекомендуется устанавливать значения  $l_{\max} \approx 10$  м.

**Удаление агентов в процессе моделирования.** Удаление (гибель) агентов типа A производится в двух случаях.

Во-первых, удаление агентов происходит, если расчетное расстояние  $l(i)$  между агентами  $A_i(t)$  и  $A_{R(i)}(t)$  меньше заданного значения  $l_{\min}$ . В этом случае агент  $A_i(t)$  удаляется, а ссылки R и L у агентов  $A_{R(i)}$  и  $A_{L(i)}$  корректируются следующим образом:  $R(L(i))=R(i)$ ,  $L(R(i))=L(i)$ .

Во-вторых, агенты удаляются при осуществлении обрезки контура, которую необходимо произвести при самопересечении или взаимном пересечении двух контуров пожаров. Для этого на каждой итерации в процессе моделирования производится анализ пересечений и самопересечений контуров.

Во время данного анализа вычисляются точки пересечения каждого из отрезков контура, образуемых агентами  $A_i$  и  $A_{R(i)}$ , со всеми остальными отрезками  $A_i$  и  $A_{R(i)}$ .

Пересечение контуров, в том числе и самопересечение, имеет место, если для некоторой пары отрезков  $A_i$ ,  $A_{R(i)}$  и  $A_j$ ,  $A_{R(j)}$  ( $i \neq j$ ,  $i \neq R(j)$ ,  $j \neq R(i)$ ) точка их пересечения  $\xi$  лежит внутри каждого из них. В этом случае изменяются ссылки у агентов  $A_i$  и  $A_{R(j)}$ :  $R(i)=R(j)$ ,  $L(R(j))=i$ . Агенты, оказавшиеся

внутри области получившегося контура, удаляются.

**Переход состояний агента.** Как уже было указано, для агентов A возможны два состояния, характеризующие текущее горение. Обозначим эти состояния как onF (агент «горит») и notF (агент «не горит»).

Переход агента из состояния onF в notF возможен в нескольких случаях.

Во-первых, агент может самостоятельно перейти в notF, когда попадает в область, в которой отсутствуют ОПГ. Во-вторых, он может быть «потушен» агентами типа B, оказывающими на них воздействие, снижающее интенсивность горения I.

Переход агента из состояния notF в onF в данной модели не предусмотрен: если агент однажды перешел в состояние notF, он продолжает в нем находиться до конца процесса моделирования. Таким образом, можно сформулировать два простых правила перехода состояния для агентов типа A.

**Правило агента A № 1.** Если  $V_{\text{опг}} = 0$ , переходить в состояние notF («не горит») и прекращать движение по среде, иначе – продолжать движение.

**Правило агента A № 2.** Если  $I = 0$ , переходить в состояние notF («не горит») и прекращать движение по среде, иначе – продолжать движение.

Исходя из данных правил, получим таблицу переходов состояний агентов типа A (см. таблицу).

#### Переходы состояний агентов типа A

##### State transitions of A agents

Состояние	$V_{\text{опг}}$	$I$	Переход состояния
onF	0	0	onF $\Rightarrow$ notF
onF	>0	0	onF $\Rightarrow$ notF
onF	0	>0	onF $\Rightarrow$ notF
onF	>0	>0	onF $\Rightarrow$ onF
notF	любое	любое	notF $\Rightarrow$ notF

#### Свойства и характеристики агентов типа B

Агент, моделирующий противопожарную команду, характеризуется следующими параметрами:

- текущие пространственные координаты (X, Y);
- значение скорости движения по местности v;
- направление обхода контура пожара  $\alpha = [R, L]$ ;
- значение скорости тушения кромки пожара  $V_T$  (мера воздействия агента B на интенсивность горения I агента A);
- номер  $fc$  ближайшего по расстоянию агента типа A;
- значение психологического состояния P;
- текущее состояние борьбы с пожаром *inFight* [true, false].

При исполнении агенты типа *B* руководствуются только одной целью – «потушить» все очаги возгорания. Для этого агент *B* перемещается по среде к ближайшему агенту  $A_i$  и воздействует на его показатель интенсивности горения  $I$ . Как только данный показатель становится равным 0, агент  $A_i$  меняет свое состояние на «не горит», а агент *B* перемещается к следующему агенту  $A_{R(i)}$ .

Перед запуском модели выбирается количество агентов типа *B*, затем задаются характеристики каждого агента ( $v$ ,  $\alpha$ ,  $V_T$ ,  $P$ ). Начальное расположение агентов на карте также задается заранее в зависимости от выбранной тактики тушения низового лесного пожара (схемы организации тушения). Схем тушения может быть четыре [14, 15]:

- по всему периметру (для слабых по интенсивности пожаров);
- охватом с фронта (для слабого и среднего по интенсивности лесного пожара);
- охватом с фланга (для пожаров средней интенсивности);
- сведением на клин (для сильного лесного пожара).

Область видимости агентов *B* ограничена только самой средой моделирования. Где бы ни находился данный агент на карте, он всегда будет передвигаться в сторону ближайшего очага пожара (агента *A*).

В процессе моделирования расчет новых координат  $B_X$ ,  $B_Y$  производится различными способами в зависимости от текущего состояния борьбы с пожаром. Если  $inFight = true$ , то по мере «тушения» в качестве координат  $X$ ,  $Y$  агента *B* принимаются текущие координаты агентов *A*. Если  $inFight = false$ , производится расчет расстояния  $S_{ji}$  от текущего агента  $B_j$  до каждого агента  $A_i$ . Затем выбирается наименьшее из этих расстояний и производится перезапись  $B_{jc}=i$ . Высчитывается нормаль направления движения  $N$  от  $B_j$  до  $A_i$ . Сами координаты  $B_X$ ,  $B_Y$  рассчитываются следующим образом:

$$B_X(t+1) = B_X(t) + N_X \cdot v,$$

$$B_Y(t+1) = B_Y(t) + N_Y \cdot v.$$

**Добавление новых агентов в процессе моделирования.** Автоматическое добавление новых агентов типа *B* в процессе моделирования не предусмотрено. Однако возможно добавление новых агентов вручную, если этого требуют цели и задачи исследования.

**Удаление агентов в процессе моделирования.** Удаление агентов типа *B* производится только в случае «гибели» противопожарной команды, которая может произойти в процессе «тушения» агентов *A*.

**Взаимодействие агентов типа *A* и *B*.** Рассмотрим более подробно взаимодействие агентов *A* и *B*, имитирующее процесс тушения участка кромки пожара противопожарной командой. Если агент *B* справится со своей задачей, то агент *A* перейдет в состояние  $notF$  («не горит»). В противном

случае агент *B* будет удален. Процесс тушения участка кромки может продолжаться одну или более итераций в процессе моделирования, а успех агента *B* напрямую зависит от показателей  $V_T$  и  $P$ .

Скорость тушения кромки пожара  $V_T$  определяется перед моделированием с учетом профессиональной подготовки, физического состояния, оснащенности и количества пожарных в команде. При тушении пожара эффективность труда во многом зависит также от психологического состояния пожарных  $P$ , которое обусловлено большим числом факторов, таких как напряженность работ, задымленность, высокая температура и др. [15]. Данный показатель может иметь значение от 0 до 1.

Если тушение одного и того же агента *A* занимает большое количество итераций моделирования, происходит значительное снижение психологического состояния пожарных  $P$ . На данный показатель влияет также показатель задымления  $M$  агента *A*. При высокой задымленности психологическое состояние  $P$  снижается значительно быстрее. Если данный показатель достигает 0, считается, что пожарная команда уже не в силах потушить данный участок кромки и агент погибает. В случае успеха  $P$  значительно увеличивается и пожарная команда может приступить к тушению следующего участка.

Помимо рассмотренных факторов, на работоспособность влияет и большое количество факторов вероятностного характера. Обозначим совокупность этих факторов как  $F$ . Они могут как снизить работоспособность (характер растительности, состояние погоды, неуверенные действия руководителя), так и повысить ее (кратковременный отдых, смена режима работы, уверенные действия руководителя). Значение  $F$  рассчитывается случайно в заданных пределах и напрямую увеличивает или уменьшает показатель  $P$ .

Таким образом, интенсивность горения агента  $A_i$ , находящегося под воздействием агента  $B_j$ , рассчитывается по формуле  $I_i(t+1) = I_i(t) - (V_{Tj} \cdot P_j)$ , где  $I_i(t+1)$  – интенсивность горения агента  $A_i$  в период времени  $t+1$ ;  $I_i(t)$  – интенсивность горения агента  $A_i$  в период времени  $t$ ;  $V_{Tj}$  – скорость тушения кромки пожара агентом  $B_j$ ;  $P_j$  – психологическое состояние пожарных агента  $B_j$ .

**Переход состояний агента.** Для агентов *B* возможны два состояния, характеризующие текущее состояние борьбы с пожаром. Обозначим эти состояния как  $fight$  (агент «тушит кромку») и  $move$  (агент «перемещается по карте»).

Переход агента из состояния  $move$  в  $fight$  возникает, когда пространственное расстояние от агента  $B$  до агента  $A_{jc}$  равно нулю, то есть противопожарная команда достигла кромки пожара.

Обратный переход возможен, если все агенты текущего контура пожара потушены и одновременно с этим на карте есть другой или другие непотушенные контуры.

Таким образом, сформулируем два простых правила перехода состояния для агентов типа *B*.

**Правило агента *B* № 1:** если твое состояние *move* и расстояние от агента *B* до агента *A<sub>ж</sub>* равно 0, переходи в состояние *fight* и начинай взаимодействие с агентом *A<sub>ж</sub>*, иначе – продолжай движение.

**Правило агента *B* № 2:** если твое состояние *fight* и количество непотушенных агентов *A* в контуре равно 0, то переходи в состояние *move* и двигайся к ближайшему горящему агенту *A*, иначе – продолжай борьбу.

### Реализация агентной модели

На основе предложенной мультиагентной модели автором был разработан программный комплекс, имеющий клиент-серверную архитектуру и веб-интерфейс. На сервере находится БД, в которой хранится информация о среде моделирования (реальной местности). Программный код серверной части реализован на языке PHP. Клиентская часть системы реализована на JavaScript (JS) с использованием библиотеки с открытым исходным кодом OpenLayers. Данная библиотека предназначена для создания карт на основе программного интерфейса (API) и позволяет создать веб-интерфейс для отображения картографических материалов. Библиотека может работать с данными различных картографических сервисов, таких как OpenStreetMap, Google, Yandex и др.

Представим два примера работы созданной системы.

#### Пример № 1.

Параметры среды, выставленные перед началом моделирования:

- тип ОПГ – лишайниковый;
- класс пожарной опасности – 2;
- скорость ветра под пологом леса – 1 м/с (Ю/З).

Согласно [14, 15], интенсивность представленного пожара определяется как слабая, поэтому выбрали схему тушения по всему периметру. Расположили 3 противопожарные команды по 6 человек вокруг пожара и задали направление обхода против часовой стрелки. Площадь очага пожара при обнаружении  $\approx 4$  га. На рисунке 3 представлены исходная позиция пожарных команд и исходный контур пожара, а на рисунке 4 – положение пожарных команд и контур пожара спустя 4 часа (48 итераций моделирования).

Как видно из представленного примера, после 4 часов локализации пожар продолжает развиваться и потушено менее 50 % текущей кромки пожара. Одновременно с этим общая длина горящей кромки пожара на 4-м часе тушения приблизительно равна длине исходной кромки пожара, что говорит о том, что 3 отряда пожарных по 6 человек не способны справиться с данным пожаром.



Рис. 3. Пример работы имитационной системы № 1 (стартовая позиция, площадь пожара – 4,3 га)

Fig. 3. Example of № 1 simulation system operation (starting position, fire area – 4,3 ha)

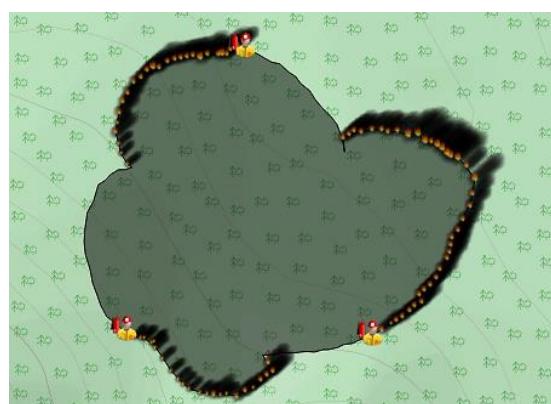


Рис. 4. Пример работы имитационной системы № 1 (время тушения пожара – 4 часа, площадь пожара – 7,5 га)

Fig. 4. Example of №1 simulation system operation (firefighting time – 4 hours, fire area – 7,5 ha)

#### Пример № 2.

Параметры среды, выставленные перед началом моделирования:

- тип ОПГ – лишайниковый;
- класс пожарной опасности – 3;
- скорость ветра – 1,5 м/с (С/В).

Согласно [14, 15], интенсивность представленного пожара определяется как средняя, поэтому выбрали схему тушения охватом с фронта. Расположили 2 противопожарные команды по 12 человек около фронта в одной точке и задали направление обхода в противоположных направлениях. Площадь очага пожара при обнаружении  $\approx 3$  га. На рисунке 5 представлены исходная позиция пожарных команд и исходный контур пожара, на рисунке 6 – положение пожарных команд и контур пожара спустя 3 часа (36 итераций моделирования). Как видим, спустя 3 часа локализации пожара фронт полностью потушен, а пожарные команды продолжают успешно продвигаться по флангам в сторону тыла пожара. Поскольку тушение фронта





Рис. 5. Пример работы имитационной системы № 2 (стартовая позиция, площадь пожара – 3,4 га)

Fig. 5. Example of № 2 simulation system operation (starting position, fire area – 3,4 ha)



Рис. 6. Пример работы имитационной системы № 2 (время тушения пожара – 3 часа, площадь пожара – 10,1 га)

Fig. 6. Example of № 2 simulation system operation (firefighting time – 3 hours, fire area – 10,1 ha)

пожара является наиболее опасным и трудоемким процессом относительно тушения других элементов пожара, то уже на этом этапе можно сказать, что для борьбы с данным пожаром была выбрана правильная схема тушения, а 2 отряда пожарных по 12 человек способны данную схему реализовать (см. [http://www.swsys.ru/uploaded/image/2016\\_3/2016-3-dop/5.jpg](http://www.swsys.ru/uploaded/image/2016_3/2016-3-dop/5.jpg)).

В результате проведенного исследования автором была разработана мультиагентная модель для имитации прямого метода тушения низового лесного пожара немеханизированными противопожарными отрядами. На основе данной модели со-

здан программный комплекс, который предлагается использовать для широкого круга задач.

Во-первых, система будет полезна для оперативного определения параметров распространения лесного низового пожара в лесных насаждениях, а также разработки комплекса мер по его предотвращению и противодействию ему. Во-вторых, она может служить основой для учебно-тренажерных систем, предназначенных для обучения студентов вузов и техникумов лесохозяйственного профиля, работников лесохозяйственной отрасли и МЧС на курсах повышения квалификации основам тактики борьбы с лесными пожарами.

#### Литература

1. Andrews P.L., Bevins C.D., Seli R.C. BehavePlus fire modeling system, version 2.0: User's Guide. USDA Forest Service Gen. Techn. Rep. RMRS-GTR-106WWW. Ogden, 2003, 45 p.
2. Finney M.A. FARSITE: Fire are simulator model, development and evaluation. USDA Forest Service, Res. Paper RMRS-RP-4. Ogden, 1998, 47 p.
3. Mell W., Jenkins M.A., Gould J., Cheney Ph. A physics-based approach to modeling grassland fires. Intern. Journ. of Wildland Fire. 2007, vol. 16, pp. 1–22.
4. Аксенов К.А., Гончарова Н.В. Динамическое моделирование мультиагентных процессов преобразования ресурсов. Екатеринбург: Изд-во УГТУ-УПИ, 2006. 311 с.
5. Кухта В.Б. Метод моделирования распространения низового пожара в лесных насаждениях с использованием агентного подхода // Вестн. Московского гос. ун-та леса – Лесной вестник. 2014. № 5. С. 92–97.
6. Dai D., Zhang Y. Simulating fire spread in a community using an agent-based model. Proc. 12th Intern. Conf. on GeoComputation. Wuhan, China, LIESMARS Wuhan Univ., 2013, pp. 130–132.
7. Shatalov P.S., Dorrer G.A. Parallel computation of forest fire and its interaction with infrastructure objects. Proc. 4th Fire Behave and Fuels Conf., St. Petersburg, 2013, p. 3.
8. Дорпер Г.А. Динамика лесных пожаров. Новосибирск: Изд-во СО РАН, 2008. 404 с.
9. Годунов С.К., Прокопов Г.П. Об использовании подвижных сеток в газодинамических расчетах // Журн. вычислит. матем. и математ. физики. 1972. Т. 12. № 2. С. 429–439.
10. Rothermel R.C. A mathematical model for fire spread predictions in wildland fuels. USDA Forest Service Research Paper INT-115, Ogden, 1972, 40 p.
11. Волокитина А.В. Защита населенных пунктов от чрезвычайных ситуаций, связанных с природными пожарами (практич. рекоменд.). Красноярск: Изд-во СО РАН, 2002. 63 с.
12. Brown D.F., Dunn W.E., Lazaro M.A., Policastro A.J. The smoke-fireplume model: tool for eventual application to prescribed burns and wildland fires. Joint Fire Sc. Program Conf. and Workshop, Boise, 1999, 12 p.
13. Мальбахов В.М., Шлычков В.А., Лежнин А.А., Дубровская О.А. Численная модель распространения дымового шлейфа при лесных пожарах с параметрическим учетом процессов горения // География и природные ресурсы. Спец. вып., 2004. С. 170–174.
14. Иванов В.А., Иванова Г.А., Москальченко С.А. Справочник по тушению природных пожаров. Красноярск: Изд-во ПРООН/МКИ, 2011. 130 с.
15. Щетинский Е.А. Тушение лесных пожаров: пособие для лесных пожарных. М.: Изд-во ВНИИЛМ, 2002. 104 с.

DOI: 10.15827/0236-235X.115.101-108

Received 14.04.16

#### FOREST FIRE SIMULATION USING AN AGENT-BASED MODEL

<sup>1</sup>Yarovoy S.V., Postgraduate Student, ach\_bask@mail.ru

<sup>1</sup>Siberian State Technological University, Mira Ave. 82, Krasnoyarsk, 660049, Russian Federation

**Abstract.** The article considers the task of modeling forest fire situation. To solve it the authors propose to use the agent-based approach with the geographic information systems (GIS) technology that allows combining the advantages

of both. As a result, an agent-based model has been developed. This model allows predicting the spread of a forest fire edge and simulating the direct method of fighting grass-roots forest fires.

Any portion of the real terrain maps can be used as an environment for agents. The proposed model uses two types of agents. Agents of the first type (A agents) are areas of fire front propagation and are used to model the behavior of forest fire. Taken together, they represent forest fire edge on a map. The model allows multiple edges, which independently spread in different parts of the map. Agents of the second type (B agents) are used for simulating firefighting actions and affect A agents.

Based on this model the authors developed a software package for a wide range of tasks. First, the system is useful for determining the operational parameters of grass-roots forest fires spread and development of measures to prevent and fight it. Second, this system can serve as a basis for forest firefighting training systems for forestry workers and the Ministry of Emergency Situations as a part of advanced professional training, as well as for teaching students of universities and colleges.

**Keywords:** wildfire, wildfire control, simulation modeling, geographic information system, agent-based model, software.

### References

1. Andrews P.L., Bevins C.D., Seli R.C. BehavePlus fire modeling system, ver. 2.0: User's Guide. *USDA Forest Service Gen. Techn. Rep. RMRS-GTR-106WWW*. Ogden, 2003, 45 p.
2. Finney M.A., FARSITE: Fire are simulator model, development and evaluation. *USDA Forest Service, Res. Paper RMRS-RP-4*. Ogden, 1998, 47 p.
3. Mell W., Jenkins M.A., Gould J., Cheney Ph. A physics-based approach to modeling grassland fires. *Int. Journ. of Wildland Fire*. 2007, vol. 16, pp. 1–22.
4. Aksenov K.A., Goncharova N.V. *Dinamicheskoe modelirovanie multiagentnykh protsessov preobrazovaniya resursov* [Dynamic Simulation of Multi-Agent Processes of Resource Conversion]. Ekaterinburg, UGTU-UI Publ., 2006, 311 p. (in Russ.).
5. Kukhta V.B. Method of modeling ground fire spread in forest plantations using an agent-based approach. *Vestnik Moskovskogo gos. universiteta lesa – Lesnoy vestnik* [Journ. of the Moscow State Forest Univ. Forestry Journ.]. 2014, no. 5, pp. 92–97 (in Russ.).
6. Dai D., Zhang Y. Simulating fire spread in a community using an agent-based model. *Proc. 12th Int. Conf. on Geo-Computation*. LIESMARS Wuhan Univ. Publ., Wuhan, China, 2013, pp. 130–132.
7. Shatalov P.S., Dorner G.A. Parallel computation of forest fire and its interaction with infrastructure objects. *Proc. 4th Fire Behave and Fuels Conf.* St. Petersburg, 2013, p. 3.
8. Dorner G.A. *Dinamika lesnykh pozharov* [Forest Fire Dynamics]. Novosibirsk, SB RAS Publ., 2008, 404 p. (in Russ.).
9. Godunov S.K., Prokopov G.P. On the use of moving grids in the gas-dynamic calculations. *Zhurn. vychislitel'noy matematiki i matematicheskoy fiziki* [Computational Mathematics and Mathematical Physics Journ.]. 1972, vol. 12, no. 2, pp. 429–439 (in Russ.).
10. Rothermel R.C. A mathematical model for fire spread predictions in wildland fuels. *USDA Forest Service Research Paper INT-115*. Ogden, 1972, 40 p.
11. Volokitina A.V. *Zashchita naselennykh punktov ot chrezvychaynykh situatsiy, svyazannykh s prirodnymi pozharemi (prakticheskie rekomendatsii)* [Protection of Inhabited Localities from Emergency Situations Connected with Wildfires (Practical Recommendations)]. Krasnoyarsk, SB RAS Publ., 2002, 63 p. (in Russ.).
12. Brown D.F., Dunn W.E., Lazaro M.A., Policastro A.J. The smoke-fireplume model: tool for eventual application to prescribed burns and wildland fires. *Joint Fire Science Program Conf. and Workshop*. Boise, 1999, 12 p.
13. Malbakhov V.M., Shlychkov V.A., Lezhnin A.A., Dubrovskaya O.A. Numerical model of smoke plume from forest fires with parametrical account of burning processes. *Geografiya i prirodnye resursy* [Geography and Natural Resources]. 2004, spec. iss., pp. 170–174 (in Russ.).
14. Ivanov V.A., Ivanova G.A., Moskalchenko S.A. *Spravochnik po tusheniyu prirodnikh pozharov* [Forest Fire Guide]. Krasnoyarsk, PROON/MKI Publ., 2011, 130 p. (in Russ.).
15. Shchetinsky E.A. *Tushenie lesnykh pozharov: posobie dlya lesnykh pozharnykh* [Forest Fire Fighting: A Guide for Forest Firefighters]. Moscow, VNIILM publ., 2002, 104 p.

### Примеры библиографического описания статьи

1. Яровой С.В. Агентный подход при моделировании лесопожарных ситуаций // Программные продукты и системы. 2016. Т. 29. № 3. С. 101–108.
2. Яровой С.В. Агентный подход при моделировании лесопожарных ситуаций // Программные продукты и системы. 2016; DOI: 10.15827/0236-235X.115.101-108.
3. Yarovoy S.V. Forest fire simulation using an agent-based model. *Programmnye produkty i sistemy* [Software & Systems]. 2016, vol. 29, no. 3, pp. 101–108 (in Russ.); DOI: 10.15827/0236-235X.115.101-108.



УДК 004.932

Дата подачи статьи: 05.05.16

DOI: 10.15827/0236-235X.115.109-120

## **УЛУЧШЕНИЕ ВИЗУАЛЬНОГО КАЧЕСТВА ИЗОБРАЖЕНИЙ, ПОЛУЧЕННЫХ В СЛОЖНЫХ УСЛОВИЯХ ОСВЕЩЕННОСТИ НА ОСНОВЕ ИНФРАКРАСНЫХ ДАННЫХ**

(Работа выполнена при поддержке РФФИ, грант № 16-07-00121 А)

А.Г. Зотин, к.т.н., доцент, zotinkrs@gmail.com;

А.И. Пахирка, к.т.н., доцент, pahirka@sibsau.ru;

М.В. Дамов, к.т.н., доцент, me@damov.pro;

Е.И. Савчина, магистрант, oreshkinaei@gmail.com

(Сибирский государственный аэрокосмический университет им. академика М.Ф. Решетнева,  
просп. им. газ. «Красноярский рабочий», 31, г. Красноярск, 660014, Россия)

В настоящее время широкое распространение получили системы видеонаблюдения на основе цифровых камер. Однако оператору довольно часто приходится сталкиваться с проблемой плохой видимости наблюдаемого объекта, обусловленной недостаточным освещением и погодными условиями. В статье рассматриваются методы повышения визуального качества изображений, полученных в сложных условиях освещенности на основе инфракрасных данных. Так, повышение визуального качества возможно за счет слияния данных визуального (RGB) и инфракрасного (IR) изображений. Методы слияния можно условно разделить на три основные категории в зависимости от принципа их работы: попиксельное слияние, слияние на основе характеристических особенностей и слияние на основе объектов. Поскольку в системах видеонаблюдения важна скорость обработки данных, в качестве основы разрабатываемого алгоритма был взят подход попиксельного слияния с установкой весовых коэффициентов.

Предлагаемый алгоритм условно разделен на два основных этапа – подготовка и обработка. В ходе подготовительного этапа формируются данные, необходимые для выполнения слияния и формирования карты хроматических компонентов цветовой модели YUV. Для повышения визуального качества предложено выполнять коррекцию инфракрасного изображения в зависимости от его типа (NIR SWIR или LWIR) и значения средней яркости. Для определения коэффициента коррекции предлагается использовать данные гистограмм и таблиц перекодировки, которые позволяют значительно сократить вычислительную сложность алгоритма.

На этапе обработки формируется улучшенное визуальное изображение с адаптивной подстройкой коэффициентов. При этом учитываются такие параметры, как яркость пикселей визуального и скорректированного инфракрасного изображений, а также сведения о насыщенности цвета.

Экспериментальные исследования показали целесообразность использования предлагаемого алгоритма.

**Ключевые слова:** обработка изображений, анализ изображений, слияние изображений, инфракрасные изображения, Retinex, LWIR, NIR, Image fusion.

Различные организации очень часто используют цифровые системы охранного видеонаблюдения. Однако в большинстве случаев применение таких систем ограничивается установкой обычных видеокамер для слежения оператором за происходящим. При этом данные с камер могут отображаться на множестве мониторов. При организации видеонаблюдения с использованием традиционных камер оператору приходится сталкиваться с проблемой плохой видимости наблюдаемого объекта в силу плохого освещения и погодных условий. Так, при недостаточном освещении в видеопоследовательности возникают шумы, характеризующиеся разноцветными точками, хаотично распределенными по экрану, и образуются засвеченные области, если в поле зрения камеры попадает источник освещения.

Для решения проблемы плохой видимости могут применяться различные методы цветовой коррекции [1–3] и шумоподавления [1, 4]. Однако при недостаточном уровне освещенности данные методы не всегда дают приемлемые результаты, в частности, возникающие шумы будут усилены алгоритмами коррекции освещения.

В последние годы растет популярность инфракрасных камер, позволяющих фиксировать различие тепловой энергии у наблюдаемого объекта и сцены. Это обусловлено тем, что они могут обеспечить достаточно четкое видео при неблагоприятных погодных условиях (дождь, туман), а также в условиях недостаточной освещенности (ночью, на рассвете и закате). Помимо этого, инфракрасные камеры полезны, когда имеются засвеченные участки, что может сделать традиционное видео непригодным для использования. Инфракрасные видеоданные в системах видеонаблюдения обычно представляются как оттенки серого, что создает хороший контраст между объектами и их фоном. Однако это не всегда удобно для оператора.

Объединение данных, полученных с помощью инфракрасной (IR) и традиционной (RGB) камер, позволит получить информацию, которую нельзя было бы получить путем просмотра данных по отдельности. Для этой цели могут применяться различные методы слияния инфракрасных и визуальных изображений. Характеристики инфракрасных изображений и область их применения определяются типом камер.

### Типы инфракрасных камер

Инфракрасное излучение (IR – infrared) – это электромагнитное излучение с длиной волны от 0,76 до 100 мкм, то есть больше, чем у видимого света. Этот диапазон условно разделяется на следующие составляющие [5]:

- ближний инфракрасный диапазон (NIR – Near Infrared Range) с длиной волны от 0,76 до 1,0 мкм;
- коротковолновый инфракрасный диапазон (SWIR – Short Wavelength Infrared Range) с длиной волны от 1 до 3 мкм;
- средний инфракрасный диапазон (MWIR – Medium Wavelength Infrared Range) с длиной волны от 3 до 6 мкм;
- длинноволновый инфракрасный диапазон (LWIR – Long Wavelength Infrared Range) с длиной волны от 6 до 15 мкм;
- сверхдлинноволновый инфракрасный диапазон (VLWIR – Very Long Wavelength Infrared Range) с длиной волны от 15 до 30 мкм;
- дальний инфракрасный диапазон (FIR – Far Infrared Range) с длиной волны от 30 до 100 мкм.

При использовании NIR- и SWIR-приемников на захватываемых изображениях у объектов есть тени и определенная контрастность, так как захватывается излучение, отражаемое от объектов. При этом объекты выглядят точно так же, как при видимом освещении с четко различимыми деталями, а разрешающая способность таких детекторов сопоставима с устройствами захвата, работающими в видимом диапазоне. Так как объекты легко различимы и узнаваемы, NIR и SWIR используют при организации систем видеонаблюдения с возможностью обнаружения и сопровождения объектов интереса. При этом при более длинной волне, чем видимый диапазон света, данные детекторы способны «видеть» сквозь туман, дым и пыль лучше, чем камеры, работающие в видимом диапазоне.

Стоит отметить, что большинство оптических материалов, используемых в обычных камерах, подходят и для NIR, и для SWIR (обычное стекло пропускает излучение до 2,5 мкм), что может существенно снизить стоимость оборудования.

При использовании MWIR и LWIR длина волны излучения больше 3,5 мкм, что влияет на захватываемое изображение, так как оно формируется на основе собственного теплового излучения объектов сцены, которое может не отличаться от теплового излучения фона. При этом уменьшается динамический диапазон изображения, ухудшаются контрастность и узнаваемость объектов.

Устройства, работающие в этом диапазоне длин волн (тепловизоры), имеют высокую стоимость, так как приходится применять специальные материалы для оптики приборов и при необходимости охлаждение детекторов, при этом разрешающая способность у дорогих устройств составляет на те-

кущий момент 1280×1024 пикселя. Данные детекторы применяются во многих областях: ночное видение, инспекция зданий, обеспечение безопасности, контроль качества в производстве товаров, авиация и т.д.

Примеры изображений [6] для различных спектральных диапазонов, полученных в различных условиях освещения, показаны на рисунке 1. Так, на рисунке приведены снимки, сделанные во время заката и в ночное время.

### Подходы к слиянию инфракрасных и видимых изображений

Методы слияния инфракрасных и видимых изображений можно разделить на три обширных класса, зависящих от уровня объекта, с которым они работают: пиксели изображения, характерные особенности изображения, конкретные объекты изображения [7–9].

Слияние на основе пикселей – это нижний уровень обработки изображений, на котором происходит слияние физических параметров исходных изображений. Наиболее известными подходами попиксельного слияния изображений являются алгоритмы простого усреднения [10], усреднение с весовыми коэффициентами ( $\alpha$ -blend), метод переноса хроматических характеристик (color transfer) [11], метод главных компонент (PCA – principal component analysis) [12], слияние с использованием пирамиды изображений [13], вейвлет-преобразование [14]. При использовании данного подхода требуется, чтобы пиксели входных изображений имели одинаковое положение на захватываемой сцене, то есть, чтобы захват инфракрасного и визуального изображений происходил в одно и то же время и в одной и той же точке сцены.

Следующий уровень обработки входных изображений – слияние на основе характерных особенностей. На данном уровне сопоставляются ключевые особенности каждого из входных изображений. Такой подход дает лучшие результаты слияния, чем обработка на пиксельном уровне, но требует предварительной подготовки используемых данных. Наиболее распространенными подходами слияния изображений на основе характерных особенностей [15] являются методы на основе выделения границ [16] и искусственные нейронные сети [17].

Слияние на основе выделения объектов на изображении – наивысший уровень обработки. На данном уровне происходит идентификация объектов индивидуально для каждого из входных изображений. Полученная информация комбинируется и используется для точного распознавания наблюдаемых объектов. Если объекты корректно распознаны со всех источников изображений, происходит их слияние для получения действительного представления сцены. Наиболее распространен-

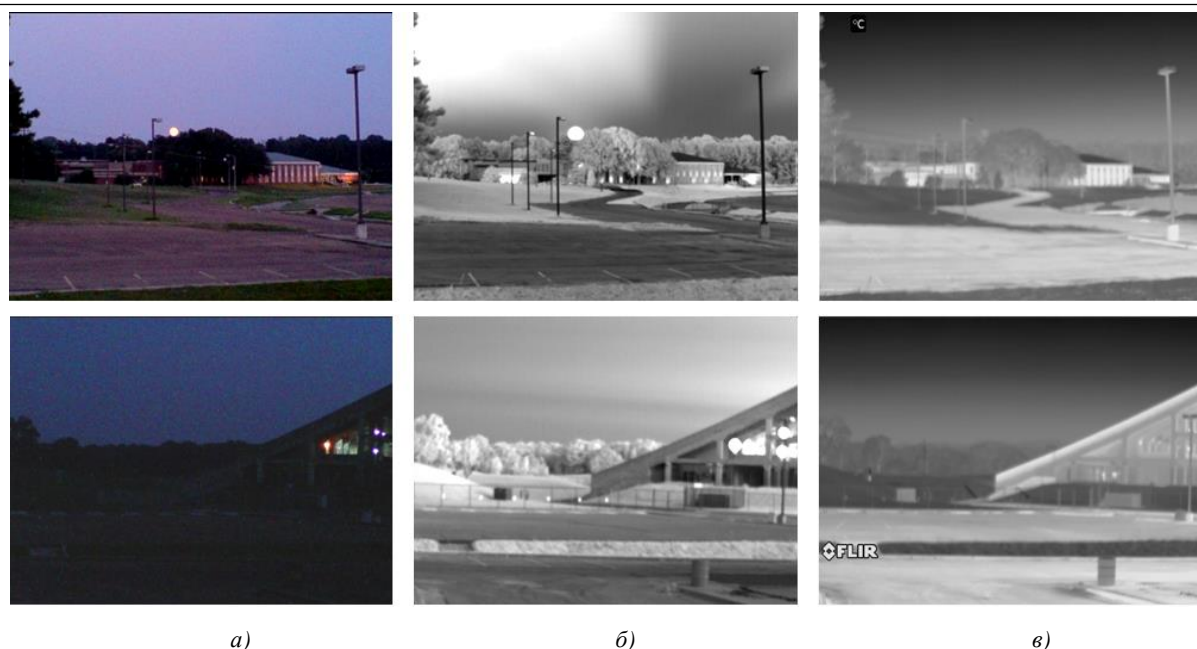


Рис. 1. Примеры изображений в различных спектральных диапазонах: а) видимый спектр – RGB-изображение; б) инфракрасное NIR-изображение; в) инфракрасное LWIR-изображение

Fig. 1. Examples of images in different spectral ranges: а) RGB-image in a visible spectrum; б) infrared NIR-image; в) infrared LWIR-image

ными подходами слияния изображений на основе выделения объектов являются алгоритмы нечеткой логики [18], байесовская сеть [19], алгоритмы машинного обучения [20].

В результате анализа принципов работы алгоритмов на каждом уровне можно прийти к выводу, что наибольшее быстродействие наблюдается у методов попиксельного слияния.

#### Методы попиксельного слияния

Наиболее простым методом слияния изображений является слияние на основе сложения данных. Так, яркость результирующего изображения может быть получена как среднеарифметическое значение инфракрасных (IR) и визуальных (RGB) данных. Помимо вычисления среднеарифметического значения, при расчете результирующей яркости может использоваться максимальное или минимальное значение.

Расширением этого метода является использование весовых коэффициентов, которые будут определять вклад исходных изображений при выполнении слияния. При этом весовые коэффициенты могут быть применены к изображению в целом или корректироваться для каждого пикселя на основе разумного правила. Такой метод иногда называют  $\alpha$ -смешиванием ( $\alpha$ -blend), поскольку весовой коэффициент представлен альфой:

$$I_{New}(x, y) = \alpha \times I_{RGB}(x, y) + (1 - \alpha) \times I_{IR}(x, y), \quad (1)$$

где  $x, y$  – координаты обрабатываемого пикселя изображений;  $\alpha$  – весовой коэффициент слияния;

$I_{New}$  – выходное изображение;  $I_{RGB}, I_{IR}$  – входные изображения RGB и IR соответственно.

Существенным недостатком этого метода является занижение контраста при выполнении слияния. Так, если одно из изображений имеет область с равномерной яркостью или низким контрастом, в результате слияния будет выравниваться контраст в этой области на результирующем изображении. Весовые коэффициенты могут помочь, но если задавать их глобально (для всего изображения), следует учитывать возможные варианты снижения контрастности.

Наиболее часто используемый алгоритм для объединения визуальных (RGB) и инфракрасных изображений типа Near Infrared (NIR) и Short Wavelength Infrared (SWIR) основывается на использовании цветовой модели HSI. К этой модели приводится RGB-изображение, затем его яркостная составляющая заменяется яркостью из NIR-изображения. Данный подход позволяет объединять изображения с большой скоростью, но обладает некоторыми недостатками. Темные пиксели NIR-изображения могут не являться таковыми на RGB-изображении, тогда при простой замене яркостей у результирующего изображения понижается интенсивность цветов пикселей или они просто становятся черными.

Эта проблема частично решается задействованием насыщенности и яркостной компоненты RGB-изображения:

$$I_{New}(x, y) = Saturation_{RGB} \times Intensity_{RGB}(x, y) + (1 - Saturation_{RGB}) \times Intensity_{IR}(x, y).$$

Однако сформированные таким способом изображения будут отображать растительность (кустарники, деревья и т.п.) в неестественных цветах.

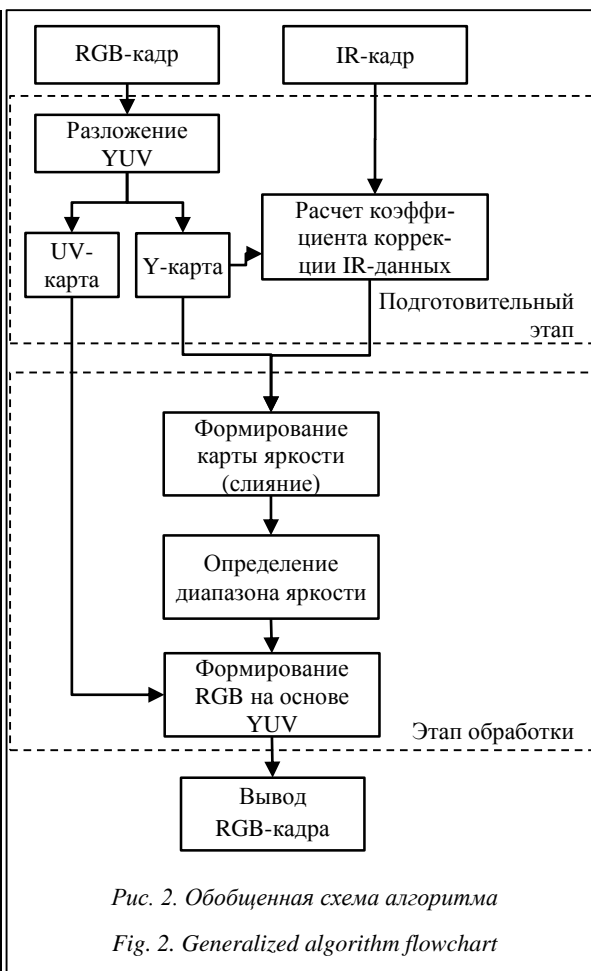
Метод главных компонент (PCA – principal component analysis) позволяет вычислить компактное описание набора данных. Он сводится к вычислению собственных векторов и собственных значений ковариационной матрицы входных данных. Слияние изображений при помощи метода главных компонент происходит следующим образом: входные изображения  $I_1(x, y)$  и  $I_2(x, y)$  преобразуются в два вектора-столбца, из которых вычитается их усредненный вектор. Далее они комбинируются в матрицу размером  $n \times 2$ , где  $n$  – длина вектора-столбца изображения, затем рассчитываются собственные векторы и собственные значения полученной матрицы. В дальнейшем используются собственные векторы ( $P_1, P_2$ ), соответствующие максимальным собственным значениям. Непосредственное слияние изображений происходит согласно выражению  $I_{New}(x, y) = P_1 I_1(x, y) + P_2 I_2(x, y)$ .

Используется также дискретное вейвлет-преобразование (DWT – discrete wavelet transform), которое во многих аспектах является аналогом преобразования Фурье. В преобразовании Фурье происходит декомпозиция сигнала на синусы и косинусы, а вейвлет-преобразование проецирует сигнал на базис, сформированный посредством масштабных изменений и переносов вейвлет-функций. Дискретное вейвлет-преобразование реализуется при помощи пирамидального алгоритма Малла [21]. При слиянии изображений с использованием DWT входные изображения раскладываются на сглаживающие и детализирующие коэффициенты с требуемым уровнем вейвлет-разложения. Далее полученные коэффициенты обоих изображений комбинируются в соответствии с выражением  $I_{New}(x, y) = IDWT[f\{DWT(I_1(x, y)), DWT(I_2(x, y))\}]$ , где  $I_1(x, y)$ ,  $I_2(x, y)$  – входные изображения;  $f$  – правило слияния изображений, в данном случае усреднение сглаживающих коэффициентов и выборка детализирующих коэффициентов в каждом диапазоне с наибольшей величиной;  $IDWT$  – обратное дискретное вейвлет-преобразование.

### Предлагаемый алгоритм

Поскольку в системах видеонаблюдения одним из значимых факторов является скорость обработки, в основу предлагаемого алгоритма взята попиксельная обработка. Так, для улучшения визуального качества видеопоследовательности с использованием инфракрасных данных предлагается алгоритм, обобщенная схема которого представлена на рисунке 2.

Условно данный алгоритм можно разделить на два основных этапа – подготовительный этап и этап обработки. На подготовительном этапе формируются данные, необходимые для выполнения



слияния, а на этапе обработки формируется улучшенное визуальное изображение.

В самом начале подготовительного этапа при считывании изображений кадров формируются гистограммы яркости  $Hist_{RGB}$  визуального ряда и  $Hist_{IR}$  для данных инфракрасной съемки и одновременно рассчитываются значения средней яркости  $AvgI_{RGB}$  и  $AvgI_{IR}$ . При этом параллельно с формированием гистограммы яркости визуального ряда происходит генерация карт хроматических компонент и карты яркости с использованием цветовой модели YUV.

Попиксельное слияние с учетом весовых коэффициентов позволяет улучшить визуальное качество, однако, если используются изображения, сильно отличающиеся по яркостному представлению, можно наблюдать затемнение или же, наоборот, осветление участков изображения, а также снижение уровня контрастности. Важным фактором является и тип инфракрасного изображения. Так, в ходе анализа исходных данных было выяснено, что в условиях недостаточной освещенности средняя яркость инфракрасного изображения LWIR зачастую меньше, чем у визуального ряда, а у изображений NIR и SWIR больше. Принимая это во внимание, было решено выполнить корректировку инфракрасных данных перед слиянием.

Поскольку интенсивность пикселей находится в диапазоне от 0 до 255, для ускорения расчетов предлагается проводить обработку с использованием таблиц перекодировки (LookUp Table – LUT).

Разные типы инфракрасных изображений обладают различными характеристиками, поэтому необходимо генерировать таблицы перекодировки с учетом этих особенностей. Так, формирование таблицы перекодировки для LWIR-изображений предлагается осуществлять по формуле

$$LUT_{kf}(i, K_{step}) = \left\lfloor \left( \frac{i}{255} \right)^{K_{step}} \times 255 \right\rfloor,$$

где  $K_{step}$  – коэффициент преобразования.

Эта же формула может быть использована для NIR- или SWIR-изображений, средняя яркость которых меньше, чем у RGB-изображения. Однако при использовании инфракрасных изображений, средняя яркость которых  $AvgI_{IR}$  превышает  $AvgI_{RGB}$ , таблица перекодировки задается по следующей формуле:

$$LUT_{kf}(i, K_{step}) = \begin{cases} \left\lfloor \left( \frac{i}{AvgI_B} \right)^{K_{step}} \times AvgI_B \right\rfloor, & \text{если } i < AvgI_B, \\ \left\lfloor \left( \frac{i - AvgI_B}{AvgI_T} \right)^{K_{step}} \times \right. \\ \quad \times (AvgI_T - \frac{AvgI_{IR} - AvgI_{RGB}}{K_Y}) + \\ \quad \left. + AvgI_B \right\rfloor, & \text{если } i \geq AvgI_B, \end{cases}$$

$$AvgI_B = \frac{(AvgI_{RGB})^2}{AvgI_{IR}}, \quad AvgI_T = 255 - AvgI_B,$$

$$K_Y = \frac{AvgI_{IR}}{AvgI_{RGB}}.$$

Вычисление среднего значения яркости на основе оригинальной гистограммы и таблицы перекодировки  $AvgI_{kf}(k)$  будет выполняться согласно

$$\text{выражению } AvgI_{kf}(k) = \sum_{i=0}^{255} LUT_{kf}(i, k) \times Hist_{IR}(i).$$

Для выполнения коррекции визуального (RGB) изображения необходимо определить значение коэффициента коррекции яркости инфракрасного изображения ( $K_{step}$ ), при котором отклонение средней яркости  $AvgI_{kf}(K_{step})$  минимально отличалось бы от значения эталонной яркости  $AvgI_{ET}$ . Значение эталонной яркости зависит от яркости визуального изображения и вычисляется согласно условию

$$AvgI_{ET} = \begin{cases} AvgI_{RGB}, & \text{если } AvgI_{RGB} > AvgI_{IR}, \\ \frac{AvgI_{RGB} + AvgI_{IR}}{2}, & \text{если } AvgI_{RGB} < AvgI_{IR}. \end{cases}$$

Расчет коэффициента коррекции  $K_{step}$  осуществляется по формуле

$$K_{step} = \arg \min_{k \in R | k_{min} < k < k_{max}} (|AvgI_{ET} - AvgI_{kf}(k)|),$$

где  $k_{min}$ ,  $k_{max}$  – минимально и максимально возможные значения для коэффициента коррекции, определяемые следующим образом:

$$k_{min} = \begin{cases} \frac{AvgI_{RGB}}{AvgI_{IR}}, & \text{если } AvgI_{RGB} \leq AvgI_{IR}, \\ 1, & \text{если } AvgI_{RGB} > AvgI_{IR}, \end{cases}$$

$$k_{max} = \begin{cases} \frac{AvgI_{RGB}}{AvgI_{IR}}, & \text{если } AvgI_{RGB} \geq AvgI_{IR}, \\ 1, & \text{если } AvgI_{RGB} < AvgI_{IR}. \end{cases}$$

В связи с этим формирование скорректированного значения яркости инфракрасного изображения ( $Y_{IR}$ ) будет осуществляться следующим образом:  $Y_{IR} = LUT_{IR}[I_{IR}] = LUT_{kf}(I_{IR}, K_{step})$ , где  $I_{IR}$  – оригинальное значение интенсивности (яркости) инфракрасного изображения;  $LUT_{IR}$  – таблица перекодировки для инфракрасного изображения.

Во время экспериментального исследования выяснено, что в случае использования LWIR-изображений для повышения визуального качества видеопоследовательностей с низким уровнем освещения (вечер/ночь/утро) целесообразно использовать скорректированное значение яркости инфракрасного изображения в инвертированной форме. Это связано с тем, что инвертированное инфракрасное изображение более приемлемо для восприятия человеком (рис. 3). Итоговая таблица перекодировки для инфракрасного LWIR-изображения с учетом инвертирования значения будет генерироваться следующим образом:

$$LUT_{IR}[I] = 255 - LUT_{kf}(I, K_{step}).$$

После завершения подготовительного этапа формируется карта яркости путем попиксельного слияния на основе весовых коэффициентов. В предлагаемом алгоритме слияние осуществляется в зависимости от яркости пикселя RGB-изображения  $Y_{RGB}$  и скорректированного значения яркости инфракрасного изображения. При этом весовые коэффициенты для каждого изображения в каждом пикселе вычисляются динамически и учитывают такие показатели, как разница яркости пикселя у изображений и насыщенность пикселя RGB-изображения. В общем виде расчет нового значения яркости можно представить следующим образом:

$$Y_{New} = \begin{cases} \frac{(K_{RGB} - K_{sat}) \times Y_{RGB} + (K_{IR} + K_{sat} - K_d) \times Y_{IR}}{K_{RGB} + K_{IR} - K_d}, & \text{если } Y_{RGB} < Y_{IR}; \\ \frac{(K_{RGB} + K_{mc} - K_{sat}) \times Y_{RGB} + (K_{IR} - K_{mc} + K_{sat}) \times Y_{IR}}{K_{RGB} + K_{IR}}, & \text{если } Y_{RGB} \geq Y_{IR}. \end{cases}$$



Рис. 3. Пример инфракрасных данных: а) оригинальное LWIR-изображение;  
б) скорректированное LWIR-изображение с учетом инверсии

Fig. 3. Example of infrared data: a) original LWIR-image; б) adjusted LWIR-image considering the inversion

$$K_d = \begin{cases} \frac{|Y_{RGB} - Y_{IR}|}{255} \times 0,5, & \text{если } Y_{RGB} < Ts_D \text{ и } |Y_{RGB} - Y_{IR}| > Ts_D, \\ 0 & \text{в противном случае;} \end{cases}$$

$$K_{mc} = \begin{cases} \frac{Y_{RGB} - Y_{IR} - 0,5 \times Ts_D}{Ts_D}, & \text{если } Y_{RGB} \leq Ts_D, \\ \frac{Y_{RGB} - Y_{IR} - Ts_D}{Ts_L - Ts_D}, & \text{если } Ts_D < Y_{RGB} \leq Ts_L, \\ \frac{Y_{RGB} - Y_{IR} - Ts_D}{Ts_L}, & \text{если } Y_{RGB} > Ts_L, \end{cases}$$

где  $Ts_D$  – порог определения темной области спектра;  $Ts_L$  – порог определения светлой области спектра;  $K_{RGB}$ ,  $K_{IR}$  – значения базовых весовых коэффициентов для RGB- и IR-изображений соответственно.

Значение насыщенности, используемое для подстройки коэффициентов, вычисляется согласно цветовой модели HSV по формуле

$$S = \begin{cases} 0, & \text{если } \max(R, G, B) = 0, \\ \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)}, & \text{если } \max(R, G, B) \neq 0, \end{cases}$$

где  $R$ ,  $G$ ,  $B$  – значения цветовых каналов модели RGB.

Величина коэффициента насыщенности  $K_{sat}$  определяется по формуле

$$K_{sat} = \begin{cases} 0, & \text{если } S > T_{sat}, \\ (1 - \frac{S}{T_{sat}}) \times Kf_{BaseS}, & \text{если } S \leq T_{sat}, \end{cases}$$

где  $T_{sat}$  – пороговое значение насыщенности цвета;  $S$  – насыщенность цвета;  $Kf_{BaseS}$  – базовое значение коэффициента влияния насыщенности цвета.

Одновременно с вычислением нового значения яркости пересчитываются минимальное ( $Y_{min}$ ) и

максимальное ( $Y_{max}$ ) значения для определения динамического диапазона.

Если динамический диапазон яркости будет меньше предусмотренного в компьютерной графике (0–255), перед формированием RGB-изображения производится дополнительная корректировка значения яркости пикселя для расширения диапазона по формуле  $Y_{New} = \frac{(Y_{New} - Y_{min}) \times 255}{Y_{max} - Y_{min}}$ .

Непосредственное формирование улучшенного изображения осуществляется на основе перевода значений цветовой модели YUV в RGB. При этом используется новое значение яркостной компоненты и от оригинального изображения берется карта хроматических данных.

### Сравнение с существующими решениями

Предлагаемый адаптивный алгоритм сравнивался с такими алгоритмами цветокоррекции Retinex [2], для которых были взяты реализация графического редактора Gimp и модифицированный Multiscale Retinex [1]. Дополнительно проведено сравнение с широко распространенными методами попиксельного слияния на основе вычисления среднего и максимального значений, а также использования статически заданных взвешенных коэффициентов ( $\alpha$ -blend). Для инфракрасных изображений NIR осуществлено сравнение с методом слияния, использующим цветовую модель HSI. Подробные результаты работы алгоритмов приведены в работе [22].

Для выполнения экспериментального исследования с инфракрасными изображениями LWIR использовались тестовые базы KAIST Multispectral Pedestrian Dataset [23] и Litiv Dataset [24]. В базе KAIST Multispectral Pedestrian Dataset содержатся



12 наборов пар инфракрасных и визуальных изображений с разрешением  $640 \times 480$ , полученных в различных условиях освещения. При этом захват данных происходит из автомобиля.

Изображение на рисунке 4 взято из набора KAIST set07 V002.seq, данные которого получены в вечернее время суток, при этом в последовательности наблюдается относительно равномерное освещение. На рисунке 4а отражаются вечернее

небо и городская улица с автомобилями и небоскребами. При этом наблюдаются как чрезмерно темные, так и яркие области в оптическом диапазоне.

Как можно заметить, предлагаемый алгоритм формирует достаточно качественное изображение (рис. 4д), на котором можно различать детали объектов, при этом не искажаются светлая область (небо) и естественные цвета сцены, как в случае с

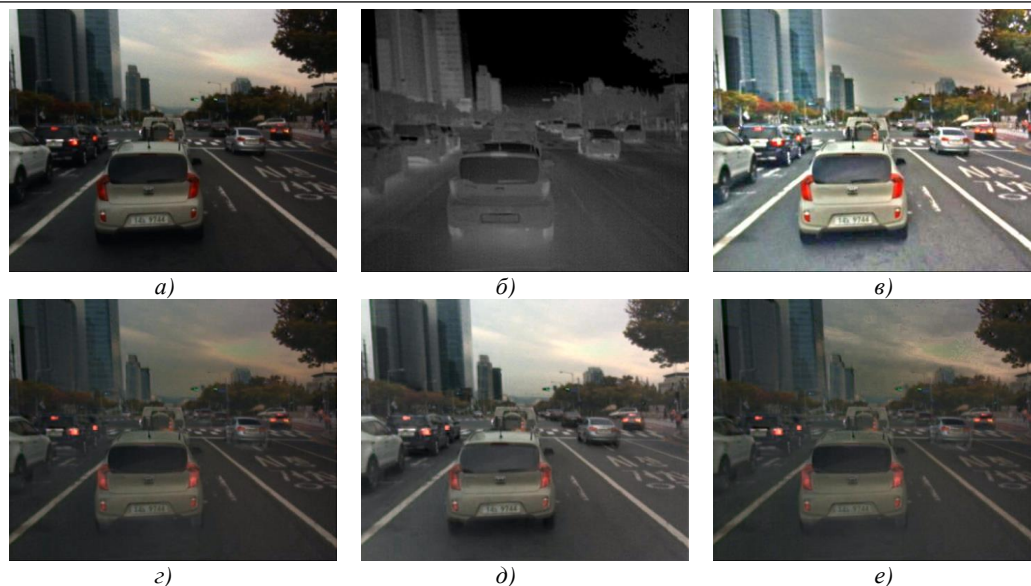


Рис. 4. Пример обработки данных, полученных в вечернее время

Fig. 4. Examples of processing data obtained in the evening

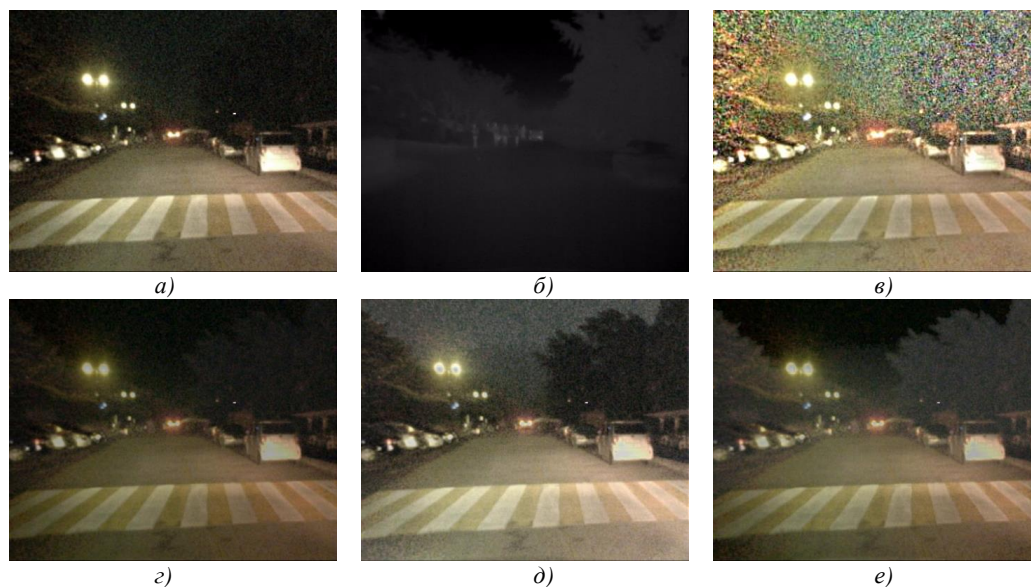


Рис. 5. Пример обработки данных, полученных в ночное время

Примечание: на рисунках 4 и 5 а) исходное RGB-изображение; б) исходное инфракрасное LWIR-изображение; в) модифицированный алгоритм MSR; г)  $\alpha$ -blend,  $\alpha=0,6$ ; д) предлагаемый алгоритм; е) предлагаемый алгоритм без инвертирования инфракрасных данных.

Fig. 5. Examples of processing data obtained at night

Comment for fig. 4 and fig. 5: а) original RGB image; б) original infrared LWIR image; в) modified algorithm MSR; г)  $\alpha$ -blend  $\alpha = 0.6$ ; д) proposed algorithm; е) proposed algorithm without inverting the infrared data.

применением алгоритмов на основе технологии Retinex. Алгоритм  $\alpha$ -blend снижает среднюю яркость изображения, но при этом позволяет лучше разглядеть часть объектов сцены (деревья, фрагменты автомобилей). В случае применения предлагаемого алгоритма без инвертирования инфракрасных данных можно наблюдать ряд артефактов в области неба, это связано с тем, что небо на LWIR-изображении (рис. 4б) имеет почти черный цвет.

На рисунке 5 показаны примеры обработки ночной сцены, полученной из набора KAIST set03 V000.seq. В данной последовательности наблюдаются значительная неравномерность освещения и наличие шумов. На изображении в оптическом спектре видна дорога с автомобилями, подсвеченная фарами. Инфракрасное изображение в основном темное, с небольшими яркими пятнами на месте впереди идущих пешеходов и автомобиля по ходу движения.

На основе анализа результатов обработки ночных сцен можно сделать заключение, что алгоритмы на основе технологии Retinex не только делают сцену более светлой, но и усиливают шум, возникающий в камерах при недостаточном уровне освещения. Изображение после обработки предлагаемым алгоритмом имеет некоторые шумовые искажения в области неба, но в целом комфортно для восприятия, к тому же только на нем четко видны деревья на обочине. Изображение после обработки предлагаемым алгоритмом без инвертирования не имеет видимых шумовых артефактов, более комфортно для восприятия, но на нем менее четко видны деревья на обочине. Результаты алгоритмов попиксельного слияния на основе среднеарифметического и максимального значений, а также  $\alpha$ -blend показывают наихудшие результаты с точки зрения визуального восприятия [22]. Это связано с очень низкой интенсивностью инфракрасного LWIR-изображения.

Улучшение результатов работы этих алгоритмов возможно в случае корректировки исходного LWIR-изображения, которая может быть основана на нормализации интенсивности или же расширении диапазона путем линейного контрастирования. Однако даже в таком случае наблюдаемая ночная сцена по визуальному качеству уступает предложенному алгоритму (рис. 6) [22]. Аналогичная ситуация характерна и для вечерней сцены, что можно наблюдать на рисунке 7 в работе [22]. Слияние по среднеарифметическому значению и алгоритм  $\alpha$ -blend с  $\alpha=0,6$  для ночной сцены с учетом предварительной коррекции инфракрасного изображения дают результат, схожий с тем, что и предлагаемый алгоритм без выполнения инвертирования скорректированной яркости инфракрасного изображения.

Тестирование алгоритма при работе с инфракрасными изображениями NIR происходило с ис-

пользованием тестовой базы IVRL – RGB-NIR Scene Dataset [25], в которой представлены 477 пар изображений, разделенных на 9 категорий.

В качестве демонстрационных примеров выбраны изображения из категории Street. Так, на рисунке 6 представлена естественная сцена с деревьями, снятая во время восхода, где наблюдается сильный засвет. В связи с особенностью съемки дерева, попадающее на линию света, выглядит очень темным, кусты на клумбе в тени почти не различимы. Инфракрасные данные NIR этой сцены (рис. 6б) позволяют оператору увидеть то, что было засвечено, а также разглядеть растительность (кусты на клумбе и деревья на заднем плане). Средняя яркость инфракрасного изображения в данном тестовом наборе ниже, чем у визуального RGB-изображения.

Работа предлагаемого алгоритма в условиях, когда средняя яркость инфракрасного изображения выше, чем у визуального, продемонстрирована на рисунке 7 (набор Street № 25). На нем показана сцена, снятая в дневное время, при этом живая изгородь, расположенная в левой части, попадает в тень и плохо различима на визуальном изображении, но хорошо видна на инфракрасном. Также на визуальном изображении есть засвеченная область, фрагменты объектов в которой хорошо различимы в инфракрасном диапазоне.

Можно заметить (см рис. 6в, 7в, а также рис. 8 и 9 в [22]), что алгоритмы на основе технологии Retinex позволяют разглядеть все детали объектов, находящихся в тени, но они не способны исправить засвеченную область изображения, и при этом происходит искажение цветов. Так, алгоритм Retinex редактора GIMP нарушает естественность изображения в сторону снижения насыщенности цвета, изображение при этом выглядит блеклым. Модифицированный алгоритм MSR, наоборот, чрезмерно увеличивает насыщенность цветов.

Алгоритм, выполняющий слияние на основе максимального значения, значительно повышает визуальное качество в затененных областях, однако не может устранить засвет. Слияние с использованием модели HSI делает различимыми детали и при этом показывает, что было в засвеченной области. Однако у него проявляется недостаток в виде снижения уровня яркости в области с низкой насыщенностью (например, фрагмент засвеченного участка неба на рисунке 6) и растительность приобретает неестественный цветовой оттенок.

Слияние как на основе среднеарифметического значения, так и алгоритмом  $\alpha$ -blend дает более естественную сцену, но объекты, попавшие в тень дерева (кустарник на клумбе), выглядят темнее по сравнению со слиянием на основе максимального значения. Предложенный алгоритм демонстрирует лучшее визуальное качество, он дает естественные цвета сцены и при этом в большей степени позволяет разглядеть объекты, попавшие в тень.



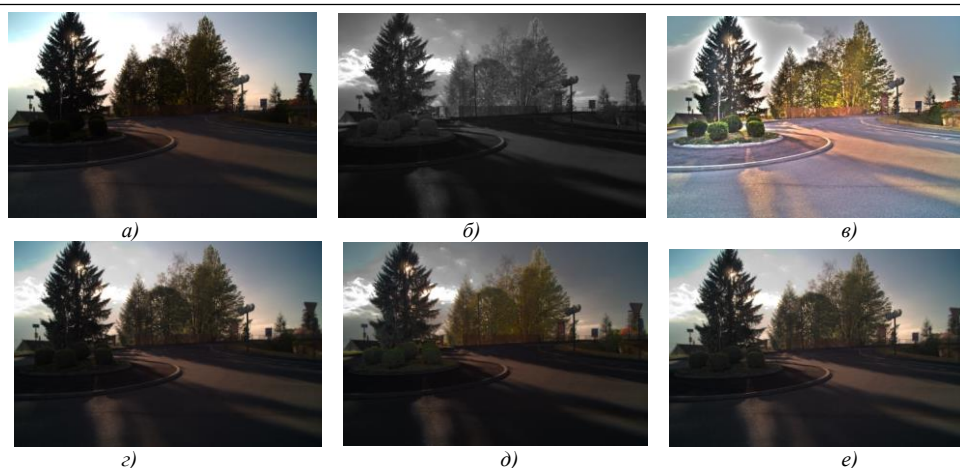


Рис. 6. Пример обработки набора Street 57  
Fig. 6. Examples of processing of data set Street 57

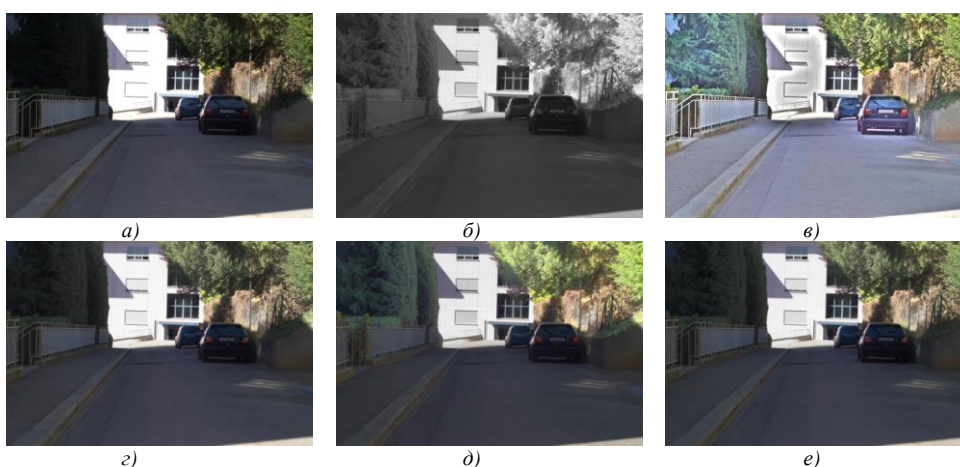


Рис. 7. Пример обработки набора Street 25  
Fig. 7. Examples of processing of data set Street 25

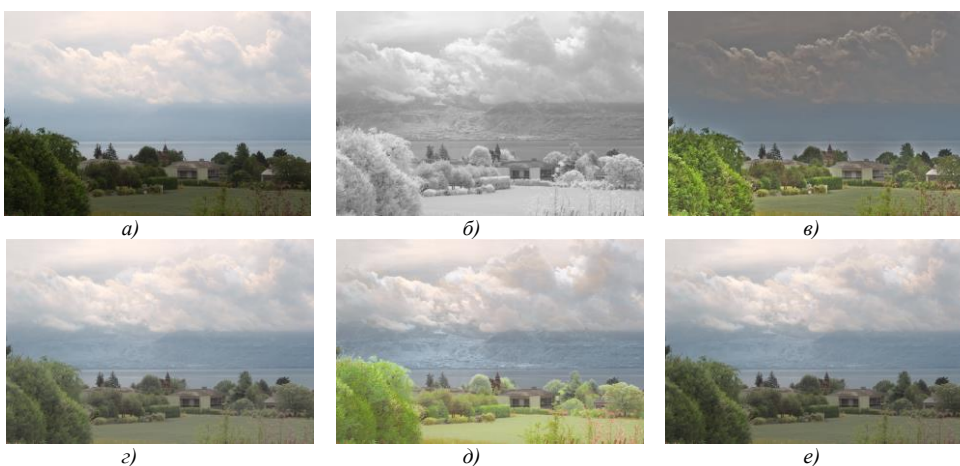


Рис. 8. Пример обработки набора Country 01

Примечание: на рисунках 6, 7, 8 а) исходное RGB-изображение; б) исходное инфракрасное NIR-изображение; в) модифицированный алгоритм MSR; г)  $\alpha$ -blend,  $\alpha=0,6$ ; д) слияние с использованием модели HSI; е) предлагаемый алгоритм

Fig. 8. Examples of processing of data set Country 01

Comment for fig. 6, 7, 8: а) original RGB image; б) original infrared NIR image; в) modified algorithm MSR; г)  $\alpha$ -blend,  $\alpha = 0,6$ ; д) HSI image fusion; е) proposed algorithm

Принимая во внимание особенности инфракрасных изображений типа NIR, было решено про-

верить возможности предлагаемого алгоритма по устранению эффекта дымки (Haze). Данный эф-

фekt возникает при наблюдении объектов на значительном расстоянии или же при наличии таких погодных явлений, как туман, мелкий дождь и т.п. В связи с этим данный эффект иногда называют эффектом тумана. Рассмотрим на примере набора Country № 01, представленного на рисунке 8 (рис. 12 в [22]), результаты работы различных алгоритмов.

В нижней части рисунка 8 изображены деревенские дома на фоне зеленой растительности, в верхней части – небо с облаками, на заднем плане расположен горный рельеф, фактически неразличимый в видимом спектре, но отчетливо просматриваемый на инфракрасном изображении. Наиболее корректно это изображение улучшают предложенный алгоритм и алгоритм  $\alpha$ -blend. Остальные результаты обработки обладают различными искажениями: тусклостью или перенасыщенностью цвета, размытостью или излишней резкостью границ объектов, что некомфортно для оператора. При этом алгоритмы на основе технологии Retinex не в состоянии сделать различимым рельеф гор на заднем плане.

В ходе экспериментального исследования было выявлено, что вне зависимости от типа инфракрасного изображения наибольшее повышение визуального качества достигается при значениях базовых коэффициентов  $K_{RGB}=3$ ,  $K_{IR}=2$  и пороговых значениях, определяющих темные и светлые участки изображения:  $T_{SD}=80$ ,  $T_{SL}=210$ . Базовое значение коэффициента влияния насыщенности цвета и пороговое значение насыщенности цвета для LWIR-изображений:  $T_{sat}=0,3$ ,  $K_{fBaseS}=0,2$ . В случае использования инфракрасных изображений NIR пороговое значение  $T_{sat}$  рекомендуется брать в диапазоне от 0,4 до 0,7, а значение коэффициента  $K_{fBaseS}$  – в диапазоне от 0,2 до 0,5.

Для выяснения, какой же тип инфракрасного изображения (LWIR или NIR) дает наилучший результат с точки зрения визуального качества, было выполнено сравнение на примере ночной сцены, отраженной на рисунке 1. Проанализировав результат (рис. 11 [22]), можно сделать вывод, что инфракрасные NIR-изображения дают более высокое визуальное качество. Это непосредственно связано с характеристиками инфракрасных изображений и особенностями их формирования.

На основании изложенного можно сделать следующие выводы. Инфракрасные камеры относительно недавно стали активно появляться на потребительском рынке. Поэтому возникла возможность задействовать их преимущества (ночная съемка, независимость от погодных условий, таких как туман и дым) наряду с преимуществами камер, работающих в визуальном спектре (высокие разрешения, позволяющие задействовать в анализе текстурные и цветовые характеристики объектов). Использование мультиспектральных изображений позволяет скомбинировать преимущества обоих

типов камер, сделать захват изображения частично независимым от погодных условий и условий освещения. Представленный в данной работе адаптивный алгоритм попиксельного слияния входных визуальных и инфракрасных изображений продемонстрировал свою состоятельность. Алгоритм может работать как с инфракрасными снимками типа NIR, так и LWIR. Определенный в ходе экспериментального исследования набор параметров позволяет получать качественный результат вне зависимости от типа инфракрасного изображения. При обработке материала, снятого в ночное время, удалось получить лучшие результаты по сравнению с алгоритмами на основе технологии Retinex и классическими алгоритмами слияния изображений.

### Литература

1. Зотин А.Г., Пахирка А.И., Буряченко В.В. Система видеонаблюдения с возможностью улучшения визуального качества видеопотока // Программные продукты и системы. 2013. № 2. С. 191–197.
2. Petro A., Sbert C., Morel J.-M. Multiscale Retinex. Image Processing On Line. 2014, pp. 71–88.
3. Chao An, Mei Yu. Fast color image enhancement based on fuzzy multiple-scale Retinex. Strategic Technology (IFOST). 6th Intern. Forum, Harbin, Heilongjiang. 2011, pp. 1065–1069.
4. Зотин А.Г., Пахирка А.И. Адаптивный алгоритм шумоподавления в системах видеонаблюдения при низком уровне освещения сцены // Цифровая обработка сигналов и ее применение – DSPA 2016: сб. докл. 18-й Междунар. конф. М., 2016. Т. 2. С. 838–843.
5. Rogalski A. Infrared Detectors. 2nd Edn. CRC Press, Taylor & Francis Group, Boca Raton, Florida, USA, 2011, 898 p.
6. Zheng Y. An exploration of color fusion with multispectral images for night vision enhancement. 2011. URL: <http://www.intechopen.com/books/image-fusion-and-its-applications/an-exploration-of-color-fusion-with-multispectral-images-for-night-vision-enhancement> (дата обращения: 29.03.2016).
7. Zin T., Takahashi H., Hiromitsu T. Fusion of Infrared and Visible Images for Robust Person Detection. Image Fusion, Osamu Ukimura (Ed.). 2011, pp. 239–264.
8. Malviya A., Bhirud S.G. Image Fusion of Digital Images. Intern. Journ. of Recent Trends in Engineering. 2009, vol. 2, no. 3, pp. 146–148.
9. Nikolakopoulos G.K. Comparison of nine fusion techniques for very high resolution data. Photogrammetric Engineering & Remote Sensing. 2008, vol. 74, no. 5, pp. 647–659.
10. Khaleghi B., Khamis A., Karray F.O., Razavi S.N. Multi-sensor data fusion: A review of the state-of-the-art. Information Fusion. 2013, vol. 14, pp. 28–44.
11. Al-Wassai F.A., Kalyankar N.V., Al-Zuky A.A. The IHS transformations based image fusion. Journ. of Global Research in Comp. Sc. 2011, vol. 2, no. 5, pp. 70–77.
12. Wang Z., Deller J.R., Jr., Fleet B.D. Pixel-level multisensor image fusion based on matrix completion and robust principal component analysis. Journ. Electron. Imaging. 2015, vol. 25, no. 1. URL: <http://electronicimaging.spiedigitallibrary.org/article.aspx?articleid=2482741> (дата обращения: 29.03.2016).
13. Wang W., Chang F. A multi-focus image fusion method based on laplacian pyramid. Journ. of Computers. 2011, vol. 6, no. 12, pp. 2559–2566.
14. Pajares G., Cruz J.M. A wavelet-based image fusion tutorial. Pattern Recognition. 2004, vol. 37, no. 9, pp. 1855–1872.
15. Liu Y., Liu S., Wang Z. Multi-focus image fusion with dense SIFT. Information Fusion. 2015, vol. 23, pp. 139–155.
16. Ehlers M., Tomowski D. On segment based image fusion. Springer Berlin Heidelberg. 2008, pp. 735–754.
17. Gao L.-Q., Wang R., Yang S. An image fusion algorithm based on RBF neural networks. Machine Learning and Cybernetics. 2005, vol. 8, pp. 5194–5199.

18. Zheng Y., Zheng P. Multisensor image fusion using fuzzy logic for surveillance systems. *Fuzzy Systems and Knowledge Discovery (FSKD)*. 2010, pp. 588–592.
19. Smaili C.C., Najjar M.E., Charpillat F. Multi-sensor fusion method using dynamic bayesian network for precise vehicle localization and road matching. *Tools with Artificial Intelligence*. 2007, pp. 146–151.
20. Tomowski D., Ehlers M., Michel U., Bohmann G. Decision based data fusion techniques for settlement area detection from multisensor remote sensing data. 1st EARSeL Workshop of the SIG Urban Remote Sensing Humboldt-Universität zu Berlin. 2006, pp. 1–8.
21. Малла С. Вейвлеты в обработке сигналов; [пер. с англ.]. М.: Мир, 2005. 672 с.
22. Зотин А.Г., Пахирка А.И., Дамов М.В., Савчина Е.И. Улучшение визуального качества изображений на основе инфракрасных данных в системах видеонаблюдения // Программные продукты, системы и алгоритмы. 2016. № 2. URL: <http://swsys-web.ru/improved-visual-quality-of-images-in-video-surveillance-systems.html> (дата обращения: 30.04.2016).
23. KAIST Multispectral Pedestrian Dataset. URL: <https://sites.google.com/site/pedestrianbenchmark/home> (дата обращения: 29.03.2016).
24. Litiv Dataset. URL: <http://www.polymtl.ca/litiv/en/vid/> (дата обращения: 29.03.2016).
25. IVRL – RGB-NIR Scene Dataset. URL: [http://ivrl.epfl.ch/supplementary\\_material/cvpr11/](http://ivrl.epfl.ch/supplementary_material/cvpr11/) (дата обращения: 29.03.2016).

DOI: 10.15827/0236-235X.115.109-120

Received 05.05.16

# VISUAL QUALITY ENHANCEMENT OF IMAGES OBTAINED IN ADVERSE ILLUMINATION CONDITIONS BASED ON INFRARED DATA

(Acknowledgements. The work has been done with support from the Russian Foundation for Basic Research, project no. 16-07-00121 A)

<sup>1</sup>Zotin A.G., Ph.D. (Engineering), Associate Professor, [zotinkrs@gmail.com](mailto:zotinkrs@gmail.com)

<sup>1</sup>Pakhirka A.I., Ph.D. (Engineering), Assistant Professor, [pahirka@sibsau.ru](mailto:pahirka@sibsau.ru)

<sup>1</sup>Damov M.V., Ph.D. (Engineering), Assistant Professor, [me@damov.pro](mailto:me@damov.pro)

<sup>1</sup>Savchina E.I., Undergraduate, [oreshkinaei@gmail.com](mailto:oreshkinaei@gmail.com)

<sup>1</sup>Siberian State Aerospace University named after academician M.F. Reshetnev, Krasnoyarsky Rabochy Av. 31, Krasnoyarsk, 660014, Russian Federation

**Abstract.** Nowadays video surveillance systems based on digital cameras are extremely popular. However, a video surveillance operator is often faced the problem of poor visibility of the observed object. This occurs in low light and adverse weather conditions.

The article considers the methods of visual quality enhancement of the images obtained in adverse lighting conditions. Thus, visual quality enhancement is possible to obtain by fusing visual data (RGB) and infrared (IR) images.

Fusion methods can be divided into three main categories according to the principles of their operation: pixel fusion, features fusion and object-based fusion. The speed of data processing is very important in video surveillance systems. The images are fused in the transform domain using novel pixel-based rules.

The proposed algorithm is divided into two main stages, which are preparation and processing. The preparation stage forms the data necessary to carry out the fusion and to form the chromatic components map of YUV color model. To enhance the visual quality the authors offer to perform the adjustment of the infrared image depending on its type (NIR SWIR and LWIR) and average brightness value. To determine the correction factor they propose to use histograms and lookup tables, which can significantly reduce the computational complexity of the algorithm.

At the processing stage forms the fused image with enhanced visual quality. During fusion the adaptation of fusion coefficients is conducted for each pixel. It is based on such parameters as brightness values of pixels of visual and corrected infrared images, as well as information of color saturation.

Experimental results show that images fused by the proposed algorithm demonstrated better spectral features than the original ones.

**Keywords:** image processing, image analysis, infrared image, Retinex, LWIR, NIR, image fusion.

## References

1. Zotin A.G., Pakhirka A.I., Buryachenko V.V. Development of video surveillance system with visual quality enhancement. *Programmnye produkty i sistemy* [Software & Systems]. 2013, no. 2, pp. 191–197 (in Russ.).
2. Petro A., Sbert C., Morel J.-M. Multiscale Retinex. *Image Processing on Line*. 2014, pp. 71–88.
3. Chao An, Mei Yu. Fast color image enhancement based on fuzzy multiple-scale Retinex. *Strategic Technology (IFOST). Proc. of 6th Int. Forum on*. Harbin, Heilongjiang, 2011, pp. 1065–1069.
4. Zotin A.G., Pakhirka A.I. The adaptive noise reduction algorithm in video surveillance systems at low light levels scene. *Doklady 18 Mezhdunar. konf. "Tsifrovaya obrabotka signalov i ee primeneniye – DSPA 2016"* [Proc. 18th Int. Conf. "Digital signal processing and its applications. DSPA 2016"]. Moscow, BRIS-M Publ., 2016, pp. 838–843 (in Russ.).
5. Rogalski A. *Infrared Detectors*. 2nd ed. CRC Press, Taylor & Francis Group, Boca Raton, Florida, USA, 2011, 898 p.
6. Zheng Y. An Exploration of Color Fusion with Multispectral Images for Night Vision Enhancement. *Image Fusion and its Applications*. 2011. Available at: <http://www.intechopen.com/books/image-fusion-and-its-applications/an-exploration-of-color-fusion-with-multispectral-images-for-night-vision-enhancement> (accessed March 29, 2016).
7. Zin T., Takahashi H., Hiromitsu T. Fusion of Infrared and Visible Images for Robust Person Detection. *Image Fusion*. Ukimura O. (Ed.). 2011, pp. 239–264.
8. Malviya A., Bhurud S.G. Image Fusion of Digital Images. *Int. Journal of Recent Trends in Engineering*. 2009, vol. 2, no. 3, pp. 146–148.
9. Nikolakopoulos G.K. Comparison of Nine Fusion Techniques for Very High Resolution Data. *Photogrammetric Engineering & Remote Sensing*. 2008, vol. 74, no. 5, pp. 647–659.

10. Khaleghi B., Khamis A., Karray F.O., Razavi S.N. Multisensor Data Fusion: A Review of the State-of-the-Art. *Information Fusion*. 2013, vol. 14, pp. 28–44.
11. Al-Wassai F.A., Kalyankar N.V., Al-Zuky A.A. The IHS Transformations Based Image Fusion. *Journal of Global Research in Computer Science*. 2011, vol. 2, no. 5, pp. 70–77.
12. Wang Z., Deller J.R. Jr., Fleet B.D. Pixel-level multisensor image fusion based on matrix completion and robust principal component analysis. *Journ. Electron. Imaging*. 2015, vol. 25, no. 1. Available at: <http://electronicimaging.spiedigitallibrary.org/article.aspx?articleid=2482741> (accessed March 29, 2013).
13. Wang W., Chang F. A Multi-focus Image Fusion Method Based on Laplacian Pyramid. *Journal of Computers*. 2011, vol. 6, no. 12, pp. 2559–2566.
14. Pajares G., Cruz J.M. A wavelet-based image fusion tutorial. *Pattern Recognition*. 2004, vol. 37, no. 9, pp. 1855–1872.
15. Liu Y., Liu S., Wang Z. Multi-focus image fusion with dense SIFT. *Information Fusion*. 2015, vol. 23, pp. 139–155.
16. Ehlers M., Tomowski D. On segment based image fusion. *Springer Berlin Heidelberg*. 2008, pp. 735–754.
17. Gao L.-Q., Wang R., Yang S. An image fusion algorithm based on RBF neural networks. *Machine Learning and Cybernetics*. 2005, vol. 8, pp. 5194–5199.
18. Zheng Y., Zheng P. Multisensor image fusion using fuzzy logic for surveillance systems. *Fuzzy Systems and Knowledge Discovery (FSKD)*. 2010, pp. 588–592.
19. Smaili C.C., Najjar M.E., Charpillat F. Multi-sensor Fusion Method Using Dynamic Bayesian Network for Precise Vehicle Localization and Road Matching. *Tools with Artificial Intelligence*. 2007, pp. 146–151.
20. Tomowski D., Ehlers M., Michel U., Bohmann G. Decision Based Data Fusion Techniques For Settlement Area Detection From Multisensor Remote Sensing Data. *1st EARSeL Workshop of the SIG Urban Remote Sensing Humboldt-Universität zu Berlin*. 2006, pp. 1–8.
21. Mallat S. *A wavelet tour of signal processing*. Academic press, 1999 (Russ. ed.: Zhileykina Ya.M. Moscow, Mir Publ., 2005, 672 p).
22. Zotin A.G., Pakhirka A.I., Damov M.V., Savchina E.I. Visual quality enhancement of the image based on infrared data in video surveillance systems. *Programmnye produkty, sistemy i algoritmy*. 2016, vol. 2. Available at: <http://swsys-web.ru/improved-visual-quality-of-images-in-video-surveillance-systems.html> (accessed April 30, 2016).
23. *KAIST Multispectral Pedestrian Dataset*. Available at: <https://sites.google.com/site/pedestrianbenchmark/home> (accessed March 29, 2016).
24. *Litiv Dataset*. Available at: <http://www.polymtl.ca/litiv/en/vid/> (accessed March 29, 2016).
25. *IVRL – RGB-NIR Scene Dataset*. Available at: [http://ivrl.epfl.ch/supplementary\\_material/cvpr11/](http://ivrl.epfl.ch/supplementary_material/cvpr11/) (accessed March 29, 2016).

---

---

## Фонд «Сколково» и Национальный исследовательский ядерный университет МИФИ объявляют конкурс проектов по информационной безопасности Skolkovo Cybersecurity Challenge 2016

Конкурс посвящен поиску лучших инновационных решений, направленных на защиту бизнеса и частных лиц от киберугроз, и проводится при поддержке Банка России.

Партнерами выступили Kaspersky Lab, InfoWatch, Almaz Capital, НТИ SafeNet, Eset и Московский политехнический университет.

Стратегический партнер конкурса – компания Cisco, мировой лидер в области разработки решений по обеспечению информационной безопасности и ключевой партнер Фонда «Сколково».

**Заявки принимаются с 1 августа по 15 ноября 2016 года.**

Итоги конкурса станут известны в финале, 8 декабря 2016 года, на ежегодной Международной конференции **Skolkovo Cyberday**, посвященной актуальным вопросам развития инновационных технологий в области информационной безопасности.

Три призовых проекта получают финансирование по 5 млн. рублей на развитие технологии. Особым решением организаторов количество призовых мест может быть увеличено до 5.

Подробная информация на сайте  
**<http://cyber2016.sk.ru/>**  
Оргкомитет

УДК 621.391

Дата подачи статьи: 08.04.16

DOI: 10.15827/0236-235X.115.121-128

## **РЕАЛИЗАЦИЯ СУБЪЕКТИВНЫХ МЕТОДОВ ВЕРОЯТНОСТНОЙ ОЦЕНКИ АДЕКВАТНОСТИ ТРЕНАЖЕРОВ НА ОСНОВЕ КОМПЛЕКСА ПРОГРАММ С ПРИВЛЕЧЕНИЕМ ЭКСПЕРТОВ И АНКЕТИРОВАНИЕМ ОБУЧАЕМЫХ**

*А.Н. Потанов, к.т.н., доцент, зам. начальника кафедры, potanov\_il@mail.ru  
(Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия  
им. проф. Н.Е. Жуковского и Ю.А. Гагарина»,  
ул. Старых Большевиков, 54а, г. Воронеж, 394064, Россия);*

*Ф.Н. Абу-Абед, к.т.н., доцент, aafares@mail.ru;*

*Д.В. Мартынов, к.т.н., доцент, idpo@tstu.tver.ru  
(Тверской государственный технический университет,  
наб. Аф. Никитина, 22, г. Тверь, 170026, Россия);*

*Е.П. Угловский, начальник лаборатории, vavko@mail.ru  
(Военная академия воздушно-космической обороны им. Маршала Советского Союза  
Г.К. Жукова, ул. Жигарева, 50, г. Тверь, 170100, Россия)*

Рассмотрено применение субъективных методов вероятностной оценки степени адекватности имитационного моделирования в тренажерах операторов эрготехнических радиоэлектронных средств. Предложено алгоритмическое обеспечение оценки адекватности имитационного моделирования с привлечением экспертов, основанной на использовании значений весовых коэффициентов компетентности  $\omega_i$  каждого  $i$ -го эксперта для определения степени согласованности мнений экспертов при независимой экспертизе, то есть когда все эксперты изолированы. На основании компетентности каждого эксперта данный метод позволяет определить действительные значения параметров оцениваемой системы и тем самым разрешить проблему ограниченного проведения для этих целей исследовательских испытаний. Так как число экспертов ограничено и это может не обеспечить необходимое количество экспериментальных данных, предлагается выполнять оценку адекватности тренажеров с привлечением самих обучаемых, которые прошли полный курс тренажной подготовки с использованием как штатной техники, так и тренажеров и могут сравнивать их. Метод оценки адекватности имитационного моделирования анкетированием обучаемых позволяет избежать дополнительных расходов на проведение испытаний тренажеров и штатной техники.

Предложено выполнять реализацию субъективных методов вероятностной оценки степени адекватности имитационного моделирования в тренажерах операторов эрготехнических радиоэлектронных средств на основе комплекса программ с привлечением экспертов и анкетированием обучаемых.

Особенностью применения программ с привлечением экспертов и анкетированием обучаемых является то, что все они обеспечивают не подтверждение гипотез об адекватности или неадекватности имитационного моделирования, а собственно оценку степени его адекватности, которая является индикатором конфликта применения тренажеров из-за их ограниченной адекватности.

**Ключевые слова:** вероятностные оценки, имитационное моделирование, оценка степени адекватности, информационные системы, эксперт, анкетирование обучаемых, комплекс программ.

Адекватность имитационного моделирования в тренажерах операторов эрготехнических радиоэлектронных средств (РЭС) (например, средств управления воздушным движением (ВД)) может быть определена, если имеются или могут быть получены необходимые и достоверные экспериментальные данные функционирования и условий применения штатной техники (ШТ).

Однако статистические экспериментальные данные, представляющие собой выборку  $\{W_i(x)\}_n$  из  $n$  элементов некоторого параметра (отклика)  $W(x)$  РЭС и необходимые для оценки вероятности  $\gamma$  адекватности имитационного моделирования в тренажере, как правило, отсутствуют [1]. Единственно возможный путь получения необходимых экспериментальных данных  $\{W_i(x)\}_n$  – это проведение натурных испытаний РЭС.

Известно, что заданной точности  $\beta$  и уровню значимости  $\alpha$  оценки адекватности имитации пара-

метра  $W(x)$  соответствует необходимое (минимальное) число испытаний  $n_{\min}$  РЭС [2]. Поэтому для оценки адекватности имитационного моделирования в тренажерах необходима выборка значений  $\{W_i(x)\}_n$  параметра  $W(x)$ , состоящая из  $n = n_{\min}$  элементов. А это означает, что необходимо проведение  $n_{\min}$  количества натурных испытаний РЭС. Однако порой из-за сложности этих испытаний в организационно-техническом плане и по экономическим соображениям их число ограничивается  $k$ , то есть возможная выборка значений  $\{W_i(x)\}_k$  параметра  $W(x)$  составляет  $k$ , причем  $k < n_{\min}$ .

Это приводит к увеличению ошибки  $\beta$  оценки адекватности имитации параметра  $W(x)$ . А с учетом значимости этого параметра  $W(x)$  увеличение ошибки  $\beta$  может привести к возрастанию конфликтов планирования тренажной подготовки (ТП) операторов с использованием тренажера [3].

**Алгоритмическое обеспечение  
оценки адекватности  
имитационного моделирования  
с привлечением экспертов**

Для получения недостающего  $n-k$  количества экспериментальных данных  $\{W_i(x)\}_{n-k}$  параметра  $W(x)$  до  $\{W_i(x)\}_n$ , необходимых для оценки адекватности  $\gamma$  имитационного моделирования в тренажере РЭС, предлагается метод экспертной оценки  $W_{\Sigma}(x)$  параметра  $W(x)$  [4]. Суть его заключается в следующем: для получения недостающей статистической информации  $\{W_i(x)\}_{n-k}$  о значениях параметра  $W(x)$  привлекаются лица (эксперты), обладающие специальными знаниями и опытом работы с данным РЭС (или близким к нему аналогом) [4]. Каждый  $i$ -й эксперт дает свою оценку  $W_{i\Sigma}(x)$  параметра  $W(x)$ . Для получения выборки  $\{W_i(x)\}_{n-k}$  из  $n-k$  элементов необходимо  $N$  экспертов, число которых определяется количеством  $n_{\min}$  элементов выборки  $\{W_i(x)\}_n$  и  $k$  проведенных испытаний РЭС, то есть  $N = n_{\min} - k$ .

Однако необходимо учитывать, что число экспертов, как и экспериментальных данных, также ограничено. Реальное количество экспертов может быть определено по усовершенствованному методу [5], суть которого в следующем: некоторой начальной группе, состоящей из  $M_0$  известных исследователю экспертов ( $M_0 < N$ ), предлагается независимо друг от друга назвать  $g$  известных им кандидатов, которые могут принять участие в данной экспертизе. Если на основании этого опроса получится, что реальное количество кандидатов, определяемое как

$$N_p = \sum_{m=1}^{M_0} \mu(m) + M_0, \quad (1)$$

примерно будет равно

$$N^* = \frac{(g \cdot M_0 + 1)(M_0 - 1)}{(g \cdot M_0 + 1) - \sum_{m=1}^{M_0} \mu(m)} + 1, \quad (2)$$

где

$$\mu(m) = \begin{cases} g, & \text{если } m\text{-й кандидат из } M_0 \text{ назвал} \\ & g \text{ кандидатов, не входящих в } M_0, \\ g-1, & \text{если } m\text{-й кандидат из } M_0 \text{ назвал} \\ & g-1 \text{ кандидатов, не входящих в } M_0, \\ \dots & \dots \\ 0, & \text{если } m\text{-й кандидат из } M_0 \text{ назвал} \\ & 0 \text{ кандидатов, не входящих в } M_0, \end{cases}$$

то достаточно одного тура опроса. В противном случае будет проводиться следующий тур с привлечением выявленных в предыдущем туре компетентных экспертов.

Для повышения достоверности экспертной оценки  $i$ -го эксперта  $W_{i\Sigma}(x)$  параметра системы  $W(x)$  используют числовую оценку, характеризующую эксперта, – степень его компетенции (вес  $\omega_i$

$i$ -го эксперта) в определении параметра  $W(x)$  [5]. А с учетом веса  $\omega_i$   $i$ -го эксперта по его оценке  $W_{i\Sigma}(x)$  можно определить соответствующее ей  $i$ -е значение выборки  $\{W_i(x)\}_{n-k}$ .

Для этого определяется выборочное среднее

$$M(W_{\Sigma}(x)) = \frac{1}{N} \sum_{i=1}^N W_{i\Sigma}(x)$$

и средневзвешенное значение

$$W(x) \approx m(W_{\Sigma}(x)) = \left( \sum_{i=1}^N \omega_i \cdot W_{i\Sigma}(x) \right) / \sum_{i=1}^N \omega_i.$$

Величина

$$\Delta W_{\Sigma}(x) = \frac{1}{N} \sum_{i=1}^N W_{i\Sigma}(x) - \left( \sum_{i=1}^N \omega_i \cdot W_{i\Sigma}(x) \right) / \sum_{i=1}^N \omega_i \quad (3)$$

показывает, насколько выборочное среднее значение экспертной оценки (без учета компетентности экспертов) отличается от средневзвешенного (с учетом компетентности экспертов).

На основании соотношения (3) имеем:

$$W_i(x) = W_{i\Sigma}(x) - \frac{1}{N} \sum_{i=1}^N W_{i\Sigma}(x) + \left( \sum_{i=1}^N \omega_i \cdot W_{i\Sigma}(x) \right) / \sum_{i=1}^N \omega_i. \quad (4)$$

Таким образом, если  $\omega_i=1$  (для всех  $i=1, 2, \dots, N$ ), то  $\Delta W_{\Sigma}(x)=0$ , следовательно,  $W_i(x)=W_{i\Sigma}(x)$ .

Вес  $\omega_i$   $i$ -го эксперта можно определить на основании его тестирования по экспериментально подтвержденным заданиям. Так, если имеются  $k$  фактических статистических данных параметра  $W(x)$ , полученных в ходе  $s$ -го испытания системы [5], то

$$\omega_i = \frac{1}{k} \sum_{s=1}^k \exp \left[ -(W_{i\Sigma}(x) - W_s(x))^2 / 2S_s^2(W_{\Sigma}(x)) \right],$$

где  $W_s(x)$  – фактическое значение параметра  $W(x)$ ;  $S_s^2(W_{\Sigma}(x))$  – выборочная дисперсия, определяемая

$$\text{как } S_s^2(W_{\Sigma}(x)) = \frac{1}{N-1} \sum_{i=1}^k (W_{i\Sigma}(x) - W_s(x))^2.$$

На основании полученных значений весовых коэффициентов компетентности  $\omega_i$  каждого  $i$ -го эксперта можно определить степень согласованности мнений экспертов в независимой экспертизе (то есть когда все эксперты изолированы):

$$\sigma^2(W_{\Sigma}(x)) = \left( \sum_{i=1}^N (W_{i\Sigma}(x) - m(W_{\Sigma}(x)))^2 \cdot \omega_i \right) / \sum_{i=1}^N \omega_i.$$

Чем меньше величина  $\sigma^2(W_{\Sigma}(x))$ , тем с большей уверенностью можно доверять найденному значению  $m(W_{\Sigma}(x))$  и, следовательно,  $W_i(x)$  в соответствии с соотношением (4).

Очевидно, что реальное число экспертов  $N=N_p$  может оказаться как больше, так и меньше  $N_{nom}$ . Поэтому при  $N > N_{nom}$  рекомендуется по результатам тестирования из  $N$  произвести выбор  $N_{nom}$  – наиболее компетентных экспертов. Если  $N < N_{nom}$ , то вероятностно-метрическая оценка адекватности имитационного моделирования в тренажере РЭС может производиться лишь по  $N+k$  статистическим данным, следовательно, ее точность не будет соот-

ветствовать заданной. Однако экспертиза значений параметров позволяет определить дополнительное количество экспериментальных данных, что повышает точность индикации сопутствующего признака конфликта применения тренажеров.

Таким образом, можно констатировать следующее: разработанное алгоритмическое обеспечение экспертной оценки значений параметров, необходимых для оценки адекватности моделирующего устройства тренажера, позволяет на основании компетентности каждого эксперта определить действительные значения параметров оцениваемой системы и тем самым разрешить проблему ограниченного проведения для этих целей исследовательских испытаний.

### Оценка адекватности имитационного моделирования анкетированием обучаемых

Известно, что оценку адекватности тренажеров можно произвести на основании сопоставления их имитационного моделирования с функционированием ШТ. В силу того, что число экспертов ограничено и это может не обеспечить необходимое количество экспериментальных данных, предлагается выполнять оценку адекватности тренажеров силами учебного заведения с привлечением самих обучаемых, которые прошли полный курс ТП с использованием как ШТ, так и тренажеров и могут сравнить их между собой. Этот метод позволяет избежать дополнительных расходов, связанных с проведением испытаний тренажеров и ШТ. При малых группах обучаемых необходимую статистику для требуемой точности оценки адекватности можно получить методом ее накопления.

Суть данного метода состоит в проведении анкетирования обучаемых. Для этого должны быть сформированы анкеты, включающие вопросы в соответствии с информационными признаками, по которым должно происходить сравнение тренажеров с ШТ. Чтобы получить общее представление о методе оценки адекватности имитационного моделирования анкетированием обучаемых, рассмотрим отдельно особенности, связанные с его разработкой.

Постановка вопроса в анкете обучаемых по конкретному информационному признаку тренажера должна иметь следующий смысл: насколько верно то, что при воздействиях на органы управления (...) расхождение между информационным признаком (...) тренажера и ШТ не превышает 10...15 %? При ответе на поставленный вопрос обучаемый выставляет свою балльную оценку  $\omega \in [0, \omega_{\max}]$ , где  $\omega_{\max}$  – максимально допустимая.

Необходимо обобщенное уточнение: в соответствии с балльной оценкой адекватность тренажера принимает значение

$$\gamma = \frac{\omega}{\omega_{\max}}. \quad (5)$$

Будем считать, что каких-либо эталонных данных, характеризующих функционирование ШТ, которые могут быть использованы для определения компетентности обучаемых, нет, а значит, в анкете отсутствуют вопросы, связанные с ее определением. Не будем заострять внимание ни на принципах формирования анкеты, ни на информационных признаках, по которым должны сравниваться тренажеры с ШТ.

Предположим, что анкета уже сформирована. Она состоит из  $M$  вопросов. Причем каждый  $j$ -й вопрос характеризует конкретный  $j$ -й информационный признак тренажера. В анкете отсутствуют дополнительные вопросы, направленные на выявление объективности ответов анкетированных, а значит, и их компетентности в оценках адекватности тренажера. Необходимо определить методику обработки результатов анкет с учетом введенных ограничений. Очевидно, что каждый вопрос в анкете имеет свою значимость. Экспертным путем [6] можно предварительно проранжировать эти вопросы и определить вес каждого  $j$ -го вопроса. С учетом этого результирующая оценка адекватности тренажера, данная  $i$ -м обучаемым, будет определяться как

$$\omega_i^* = \sum_{j=1}^M \delta_{\omega_j} \omega_{ij}, \quad (6)$$

где  $\delta_{\omega_j}$  – эталонный вес  $j$ -го информационного признака;  $\omega_{ij}$  – оценка  $i$ -м обучаемым  $j$ -го информационного признака ( $\omega_{ij} \in [0, \omega_{\max}]$ ) [7].

В свою очередь, общая оценка адекватности тренажера при обработке  $N$  анкет может определяться по следующему соотношению:

$$\omega^* = \frac{\sum_{i=1}^N \alpha_i \omega_i^*}{\sum_{i=1}^N \alpha_i}, \quad (7)$$

где  $\alpha_i$  – компетентность  $i$ -го обучаемого [8].

Так как по условию в анкете отсутствуют дополнительные вопросы, по ответам на которые можно было бы определить компетентность анкетированных, в прямой постановке, используя традиционные методы, определить ее не представляется возможным [7]. Остановимся подробнее на определении компетентности на примере использования одного традиционного метода. Пусть одни и те же анкетированные многократно используются для оценивания числовых величин. Относительная ошибка  $i$ -го анкетированного при  $k$ -м опросе

$$\varepsilon_{ik} = \frac{|T_{\omega k} - T_{ik}|}{T_{\omega k}}, \quad (8)$$

где  $T_{\omega k}$  – эталонное значение;  $T_{ik}$  – оценка, данная  $i$ -м анкетированным.



$$\text{Тогда } \alpha_i = 1 - \frac{\sum_{k=1}^H \left( \frac{1}{m_k} \sum_{s=1}^{m_k} \varepsilon_{ks} \right)}{\frac{1}{m_i} \sum_{s=1}^{m_i} \varepsilon_{is}}, \quad (9)$$

где  $m_n$  ( $n=i, k$ ) – количество оценок, которые дал  $n$ -й анкетиремый [7].

Примем, что оценка компетентности анкетиремых возможна только в рамках текущего опроса ( $H=1$ ). Следовательно, количество оценок, которое может дать анкетиремый, ограничивается количеством в анкете вопросов ( $m_n=M$ ). С учетом этого соотношения (8) и (9) будут иметь вид

$$\varepsilon_{ij} = \frac{|\delta_{\varepsilon j} - \delta_{ij}|}{\delta_{\varepsilon j}}, \quad (10)$$

$$\alpha_i = 1 - \frac{1}{M} \sum_{j=1}^M \varepsilon_{ij}, \quad (11)$$

где  $\delta_{\varepsilon j}$  – эталонное значение веса  $j$ -го вопроса анкеты;  $\delta_{ij}$  – оценка веса  $j$ -го вопроса анкеты, данная  $i$ -м анкетиремым.

Так как оценки  $\delta_{ij}$  в явном виде отсутствуют, то использование (10) и (11) для определения компетентности в прямой постановке не представляется возможным. Оценку компетентности ответа  $i$ -го обучаемого на  $j$ -й вопрос анкеты можно осуществить и без наличия эталонных данных по следующему соотношению:

$$\alpha_{ij} = \exp \left\{ -\frac{(\omega_{ij} - \bar{\omega}_j)^2}{2S_j} \right\}, \quad (12)$$

$$\text{где } \bar{\omega}_j = \frac{\sum_{i=1}^N \alpha_{ij} \omega_{ij}}{\sum_{i=1}^N \alpha_{ij}}, \quad (13)$$

$$S_j = \frac{1}{N-1} \sum_{i=1}^N (\omega_{ij} - \bar{\omega}_j)^2 \quad (14)$$

– соответственно средневзвешенная оценка и дисперсия ошибки  $j$ -го информационного признака тренажера [7].

Процедура определения  $\alpha_{ij}$  основана на методе итерации и состоит в следующем:

- а) первоначально задают  $\alpha_{ij}=1$ ;
- б) по соотношениям (13) и (14) на основании личных оценок  $\omega_{ij}$  определяются компоненты соотношения (12);
- в) по соотношению (12) осуществляется уточненное вычисление  $\alpha_{ij}$ ;
- г) в соответствии с новым значением  $\alpha_{ij}$  происходит повторное выполнение пунктов б) и в) до тех пор, пока не будет достигнуто стационарное состояние значений  $\alpha_{ij}$ ,  $\bar{\omega}_j$  и  $S_j$ .

Очевидно, что возможна ситуация, когда по каждому оцениваемому  $j$ -му информационному признаку тренажера анкетиремый будет иметь различные степени компетентности, что не совсем

корректно. Для устранения этого недостатка предлагается определять средневзвешенную оценку  $\bar{\alpha}_i$  компетентности с учетом ранее полученных  $\alpha_{ij}$ . Нахождение средневзвешенной оценки  $\bar{\alpha}_i$  происходит аналогичным образом, как и  $\bar{\omega}_j$ , по следую-

$$\text{щей формуле: } \bar{\alpha}_i = \frac{\sum_{j=1}^M \beta_{ij} \alpha_{ij}}{\sum_{j=1}^M \beta_{ij}}, \quad \text{где}$$

$$\beta_{ij} = \exp \left\{ -\frac{(\alpha_{ij} - \bar{\alpha}_i)^2}{2D_i} \right\}, \quad D_i = \frac{1}{M-1} \sum_{j=1}^M (\alpha_{ij} - \bar{\alpha}_i)^2$$

– степень значимости  $j$ -го информационного признака тренажера и дисперсия ошибки компетентности  $i$ -го анкетиремого соответственно.

Имея значения  $\beta_{ij}$ , можно определить вес каждого  $j$ -го вопроса, соответствующий  $i$ -му анкетиремому:  $\delta_{ij} = \frac{\beta_{ij}}{\sum_{j=1}^M \beta_{ij}}$ .

На основании сравнения  $\delta_{\varepsilon j}$  с  $\delta_{ij}$  в соответствии с (10) можно определить относительную ошибку оценки веса  $j$ -го вопроса  $i$ -го анкетиремого, а значит, используя соотношение (11), и его компетентность  $\alpha_i$ .

Необходимо отметить, что для обеспечения организованного управления ТП важно знать не только результирующую оценку адекватности тренажера, определяемую по соотношениям (6) и (7), но и оценки его  $j$ -х информационных признаков. Это позволит выявлять причинно-следственные факторы, характеризующие адекватность тренажера, и определять управляющие воздействия, направленные на корректировку конкретных аспектов его использования в интересах повышения качества обучения. С учетом этого, не нарушая общности рассуждений, (6) и (7) можно представить иначе:

$$\omega_j^* = \frac{\sum_{i=1}^N \alpha_i^* \omega_{ij}}{\sum_{i=1}^N \alpha_i^*}, \quad (15)$$

$$\omega^* = \sum_{j=1}^M \delta_j^* \omega_j, \quad (16)$$

где  $\alpha_i^* = \alpha_i \bar{\alpha}_i$  – уточненная оценка компетентности

$i$ -го анкетиремого;  $\delta_j^* = \frac{\sum_{i=1}^N \beta_{ij}}{\sum_{j=1}^M \sum_{i=1}^N \beta_{ij}}$  – вес  $j$ -го вопроса

анкеты, определенный по результатам анкетирования.

Окончательную степень согласованности результатов анкетирования можно оценить по следующим соотношениям:



$$\sigma^2(\omega_j) = \frac{\sum_{i=1}^N \alpha_i^* (\omega_{ij} - \omega_j^*)^2}{\sum_{i=1}^N \alpha_i^*}, \quad \sigma^2(\omega) = \frac{\sum_{j=1}^M \beta_j (\omega_j^* - \omega^*)^2}{\sum_{j=1}^M \beta_j},$$

где  $\beta_j = \sum_{i=1}^N \beta_{ij}$  – суммарная по результатам анкетирования оценка степени значимости  $j$ -го вопроса анкеты.

Чем меньше величины  $\sigma^2(\omega_j)$  и  $\sigma^2(\omega)$ , тем с большей уверенностью можно доверять найденным значениям  $\omega_j^*$  и  $\omega^*$ .

На основании (15) и (16) с учетом (5) получаем соответствующие оценки адекватности тренажера, которые будут иметь вид

$$\gamma_j^* = \frac{\omega_j^*}{\omega_{\max}^*}, \quad (17)$$

$$\gamma^* = \frac{\omega^*}{\omega_{\max}^*}. \quad (18)$$

Определение  $\delta_{\varepsilon j}$  возможно на основании предварительной экспертизы вопросов анкеты обучаемого. По аналогии с вышеизложенным суть такой экспертизы может заключаться в привлечении  $N_{\varepsilon}$  экспертов, которые по каждому  $j$ -му вопросу анкеты обучаемого выставляют оценку  $\beta_{\varepsilon jm}$  его значимости,  $m = \overline{1, N_{\varepsilon}}$ . Причем эта оценка может быть дана в баллах. Зная компетентность  $\alpha_m$   $m$ -го эксперта, можно определить средневзвешенную экспертную оценку  $j$ -го вопроса и дисперсию ее ошибки:

$$\bar{\beta}_{\varepsilon j} = \frac{\sum_{m=1}^{N_{\varepsilon}} \alpha_m \beta_{\varepsilon jm}}{\sum_{m=1}^{N_{\varepsilon}} \beta_{\varepsilon jm}}, \quad \sigma^2(\beta_{\varepsilon j}) = \frac{\sum_{m=1}^{N_{\varepsilon}} \alpha_m (\beta_{\varepsilon jm} - \bar{\beta}_{\varepsilon j})^2}{\sum_{m=1}^{N_{\varepsilon}} \beta_{\varepsilon jm}}.$$

Следовательно, эталонное значение веса  $j$ -го вопроса, используемого в дальнейшем для обработки результатов анкетирования, может быть определено как  $\delta_{\varepsilon j} = \frac{\bar{\beta}_{\varepsilon j}}{\sum_{j=1}^M \bar{\beta}_{\varepsilon j}}$ .

$$\delta_{\varepsilon j} = \frac{\bar{\beta}_{\varepsilon j}}{\sum_{j=1}^M \bar{\beta}_{\varepsilon j}}.$$

Чем выше  $\alpha_m$  и больше  $N_{\varepsilon}$ , тем меньше дисперсия  $\sigma^2(\beta_{\varepsilon j})$  и с тем большей уверенностью можно доверять значениям  $\bar{\beta}_{\varepsilon j}$ , а значит, и  $\delta_{\varepsilon j}$ .

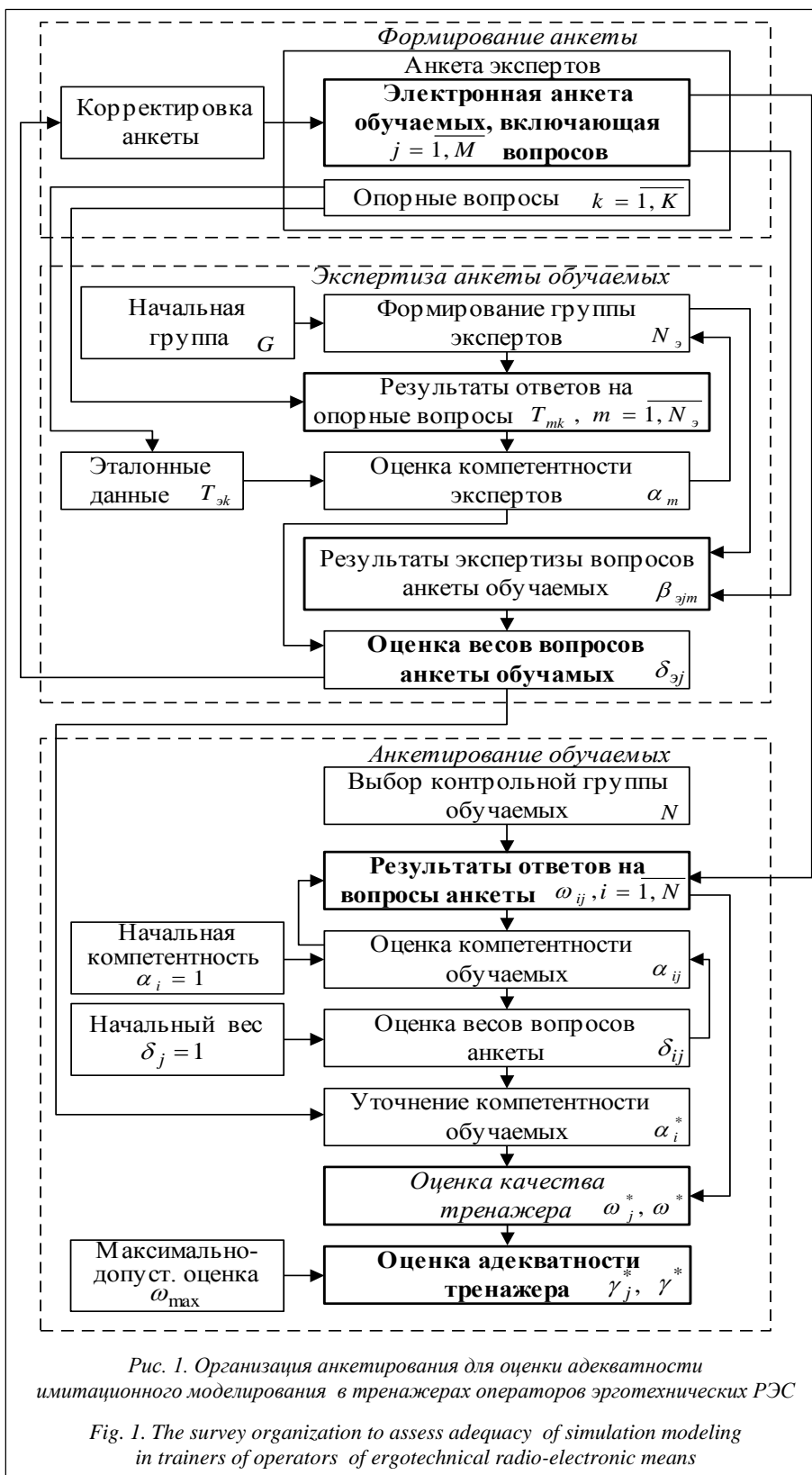
С учетом этого для проведения экспертизы необходимо первоначально сформировать группу, состоящую из  $N_{\varepsilon}$  компетентных экспертов. Для этого можно воспользоваться несколько модернизированной методикой [8].

Пусть организатору экспертизы анкеты обучаемых известно  $G$  кандидатов, которые могут принять в ней участие. Сначала для этих кандидатов по контрольному тесту определяется оценка их компетентности  $\alpha_m$ . Это возможно, если в основу теста заложить некоторые эталонные значения, которые

сравниваются со значениями ответов кандидатов в соответствии с соотношениями (8) и (9). Далее, задав граничное значение компетентности  $\alpha_{\text{зад}}$ , производится выбраковка некомпетентных кандидатов, у которых  $\alpha_m < \alpha_{\text{зад}}$ . Таким образом, формируется группа из  $M_0 < G$  компетентных экспертов ( $\alpha_m \geq \alpha_{\text{зад}}$ ). Затем производится независимый опрос  $M_0$  экспертов, суть которого заключается в следующем: каждый  $m$ -й эксперт из  $M_0$  называет  $p$  известных ему кандидатов, которые могут принять участие в данной экспертизе и не входят в  $M_0$ . Далее для каждого из  $p$  известных кандидатов по аналогии с формированием  $M_0$  происходит выбраковка некомпетентных. С учетом этого определяются  $g$  компетентных кандидатов, не известных руководителю экспертизы. Если на основании этого опроса получится, что реальное количество компетентных кандидатов, определяемое по (1), будет примерно равно (2), то достаточно одного тура опроса. В противном случае будет проводиться следующий тур с привлечением выявленных в предыдущем туре компетентных экспертов. Далее производится экспертиза по оценке весовых показателей вопросов анкеты. Каждому эксперту должна предлагаться анкета. Отличительной особенностью анкеты эксперта от анкеты обучаемого может быть то, что, помимо вопросов, включенных в последнюю для проведения экспертизы, могут быть включены дополнительные вопросы, необходимые для определения компетентности  $\alpha_m$  эксперта и  $p$  известных ему кандидатов на участие в экспертизе. Помимо этого, проведение такой экспертизы позволит формировать корректную анкету обучаемых с учетом весовых оценок  $\delta_{\varepsilon j}$ , например, исключая те вопросы, которые имеют в анкете наименьшую значимость.

Процесс организации анкетирования и обработки его результатов можно представить в виде структуры [8], изображенной на рисунке 1.

Таким образом, разработан метод, позволяющий обеспечить вероятностную оценку степени адекватности имитационного моделирования в тренажерах анкетированием обучаемых, когда необходимые экспериментальные данные от объектов моделирования отсутствуют или не могут быть получены. Этот метод базируется на применении анкет обучаемых и анкет экспертов и вполне может быть реализован в рамках учебного заведения, где в учебном процессе используются тренажеры. Оценка адекватности тренажера (18) является индикатором конфликта применения тренажеров. В свою очередь, оценка адекватности  $j$ -го информационного признака (17) с учетом его значимости может быть использована как для совершенствования методического обеспечения и деятельности инструкторского состава, направленных на сглаживание выявленных недостатков тренажеров, так и для предъявления требований к разработчикам тренажеров по их модернизации.



### Реализация методов

Реализацию субъективных методов вероятностной оценки степени адекватности имитационного моделирования в тренажерах операторов эрготех-

нических РЭС предлагается выполнять на основе комплекса программ оценки адекватности тренажеров с привлечением экспертов и анкетированием обучаемых. Особенностью применения этих программ с привлечением экспертов и анкетированием обучаемых является то, что они обеспечивают не подтверждение гипотез об адекватности или неадекватности имитационного моделирования, а собственно оценку степени его адекватности, которая является индикатором конфликта применения тренажеров из-за их ограниченной адекватности.

На рисунке 2 представлена структура комплекса программ, реализующих субъективные методы вероятностной оценки степени адекватности имитационного моделирования в тренажерах операторов эрготехнических РЭС.

Оценка адекватности имитационного моделирования в тренажерах с привлечением экспертов основана на использовании значений весовых коэффициентов компетентности каждого  $i$ -го эксперта для определения степени согласованности мнений экспертов в независимой экспертизе, то есть когда все эксперты изолированы. Компетентность каждого эксперта позволяет определить действительные значения параметров оцениваемой системы и, таким образом,

решить проблему по ограниченному проведению для этих целей исследовательских испытаний. На рисунке 3а представлен интерфейс программы оценки адекватности тренажеров с привлечением экспертов [9].



Программа оценки адекватности тренажеров анкетированием обучаемых [10] создана для автоматизации оценки адекватности тренажеров обучаемыми, прошедшими курс тренажерной подготовки с использованием тренажеров и ШТ и способными сравнить их между собой по информационным признакам. Интерфейс этой программы изображен на рисунке 3б. Перечень информационных признаков в виде анкеты обучаемых формируется на основании обработки данных экспертизы. Для этого в программе по новым алгоритмам определяется группа экспертов с учетом их компетентности и производится корректировка вопросов анкеты обучаемых. Помимо этого, в программе производится обработка данных анкетирования обучаемых, на основании которой определяется оценка адекватности информационных признаков и тренажера в целом.

### Литература

1. Дудоров А.Д. Методика оценки тренажеров лиц группы руководства полетами в интересах повышения качества их подготовки // Современные проблемы и перспективные направления развития авиационных комплексов и систем военного назначения, формы и способы их боевого применения: сб. стат. Всерос. науч.-практич. конф. Воронеж: Изд-во ВАИУ, 2011. С. 92–98.
2. Сысоев Е.С. Оптимизация тренажерной подготовки операторов сложных информационных радиоэлектронных систем управления воздушным движением // Информатика: проблемы, методология, технологии: матер. 11-й Междунар. науч.-методич. конф. Воронеж: Изд-во ВГУ. Т. 2. С. 216–222.
3. Лебедев В.В., Романов А.В. Теоретико-множественный подход к описанию функционирования эрготехнических радиоэлектронных средств // Проблемы эффективности и безопасности функционирования сложных технических и информационных систем: тр. 32-й Всерос. науч.-технич. конф. Серпухов: Изд-во филиала ВА РВСН, 2013. С. 118–122.
4. Помазуев О.Н., Платонов А.Ю., Миронов А.М., Абу-Абед Ф.Н., Рюмшин А.Р. Методика определения показателей качества обнаружения радиолокационных станций // Программные продукты и системы. 2014. № 2. С. 118–124.
5. Агапов Е.А. Проектирование автоматизированных информационных систем освоения эрготехнических комплексов //

а)

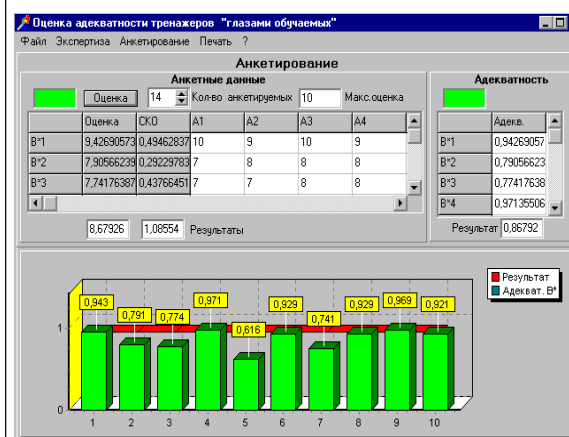


Рис. 3. Программа оценки адекватности тренажеров: а) с привлечением экспертов, б) анкетированием обучаемых

Fig. 3. The program of trainer adequacy evaluation: а) with experts involved, б) a questionnaire survey of trainees

Эффективность автоматизированных систем управления авиацией, систем связи и РТО ВВС: науч.-технич. сборник. Ногинск: Изд-во 4 ЦНИИ МО РФ, 2012. С. 112–118.

6. Дикарев В.А. Особенности профессиональной подготовки операторов при использовании компьютерных систем тренажа // XXXIV научные чтения, посвященные разработке творческого наследия К.Э. Циолковского: сб. стат. Всерос. науч.-технич. конф. Калуга: Калужский дом знаний, 1999. С. 181–184.
7. Пономаренко В.А. Психология жизни и труда летчика. М.: Воениздат, 1992. 58 с.
8. Сысоев В.В., Крутских П.П., Свинцов А.А. Математическая модель информационного конфликта // Радиотехника (журнал в журнале). 1999. № 3. С. 77–80.
9. Потапов А.Н., Пасмурнов С.М. Математическая модель количественного критерия оценки адекватности имитационного моделирования в тренажерах операторов эрготехнических систем // Вестн. Воронеж. гос. технич. ун-та. 2012. Т. 8. № 3. С. 4–8.
10. Губсков Ю.А., Прокофьев С.Н. Системная модель информационной конфликтно-устойчивой автоматизированной системы освоения эрготехнических комплексов // Информатика: проблемы, методология, технологии: матер. 12-й Междунар. науч.-методич. конф. Воронеж: Изд-во ВГУ. 2012. Т. 1. С. 109–112.

**IMPLEMENTATION OF SUBJECTIVE METHODS  
OF SIMULATOR ADEQUACY PROBABILISTIC ASSESSMENT BASED  
ON A SOFTWARE SYSTEM INVOLVING EXPERTS AND QUESTIONING TRAINEES**

<sup>1</sup>Potapov A.N., Ph.D. (Engineering), Associate Professor, Deputy Head of Chair, potapov\_il@mail.ru

<sup>2</sup>Abu-Abed F.N., Ph.D. (Engineering), Associate Professor, aafares@mail.ru

<sup>2</sup>Martynov D.V., Ph.D. (Engineering), Associate Professor, idpo@tstu.ver.ru

<sup>3</sup>Uglovsky E.P., Head of Laboratory, vavko@mil.ru

<sup>1</sup>Military Scholastic-Scientific Centre of the Air Forces "Zhukovsky and Gagarin Air Forces Academy", Starykh Bolshevikov St. 54a, Voronezh, 394064, Russian Federation

<sup>2</sup>Tver State Technical University, Nikitin Quay 22, Tver, 170026, Russian Federation

<sup>3</sup>Military Academy of the Aerospace Defence, Zhigarev St. 50, Tver, 170022, Russian Federation

**Abstract.** The article considers the issues of applying subjective methods of probable assessment of a simulation modeling level of adequacy in trainers of operators of ergotechnical radio-electronic means. The authors offer a method of simulation modeling adequacy assessment involving experts. The method is based on using values of weight factors of competence  $\omega_i$  for each  $i$ -th expert to determine a level of coherence of experts' judgements in independent examination, i.e. when all experts are isolated. This method allows defining real values of the evaluated system's parameters based on each expert's competence and, thereby, to resolve a problem of limited investigation tests for these purposes of. As the number of experts is limited and it can not provide enough quantity of experimental data, it is offered to assess trainers' adequacy involving trainees who completed a full course of simulation training using both standard equipment and simulators, and can compare them. The proposed method allows avoiding the additional expenditures connected with carrying out tests of trainers and standard equipment.

The authors propose to implement subjective probabilistic methods to assess the degree of simulation adequacy in training equipment of operators of ergotechnical radio-electronic means on the basis of the software complex with experts and a questionnaire survey among trainees. A special feature of application programs with experts and a questionnaire survey among trainees is the fact that they all provide no verification of the hypothesis about the adequacy or inadequacy of simulation, but its adequacy evaluation, which is the indicator of a conflict of using simulators due to their limited adequacy.

**Keywords:** probable estimates, simulation modeling, adequacy level assessment, information systems, expert questioning.

**References**

1. Dudorov A.D. A trainer assessment method for of a flight management group to improve their preparation quality. *Sovremennye problemy i perspektivnye napravleniya razvitiya aviatsionnykh kompleksov i sistem voennogo naznacheniya, formy i sposoby ikh boevogo primeneniya: sb. stst. Vseros. nauch.-praktich. konf.* [Proc. All-Russian Science and Practice Conf. "Modern Problems and Prospects of Aviation Complexes and Military Systems, Forms and Methods of Their Fighting Application"]. 2011, Voronezh, VAIU Publ., pp. 92–98 (in Russ.).
2. Sysoev E.S. Optimization of simulation training of operators of difficult information radio-electronic traffic air management systems. *Informatika: problemy, metodologiya, tekhnologii: mater. 11 Mezhdunar. nauch.-metodich. konf.* [Proc. 11th int. Science and Methodological Conf. "Informatics: Problems, Methodology, Technologies"]. Voronezh, VSU Publ., 2011, vol. 2, pp. 216–222 (in Russ.).
3. Lebedev V.V., Romanov A.V. The set-theoretical approach to the description of ergotechnical radio-electronic means functioning. *Problemy effektivnosti i bezopasnosti funktsionirovaniya slozhnykh tekhnicheskikh i informatsionnykh sistem: tr. 32 Vseros. nauch.-tekhnich. konf.* [Proc. 32nd All-Russian Science and Technical Conf. "Problems of Efficiency and Safety of Difficult Technical and Information System Functioning"]. Serpukhov, Peter the Great Military Academy of Strategic Missile Forces Publ., 2013, pp. 118–122 (in Russ.).
4. Pomazuev O.N., Platonov A.Yu., Mironov A.M., Abu-Abed F.N., Ryumshin A.R. The method of quality factors estimation for radar detection. *Programmye produkty i sistemy* [Software and Systems]. 2014, no. 2, pp. 118–124 (in Russ.).
5. Agapov E.A. Design of automated information systems for ergotechnical complexes acquisition. *Effektivnost avtomatizirovannykh sistem upravleniya aviatsiey, sistem svyazi i RTO VVS: nauch.-tekhnich. sb.* [Efficiency of Automated Control Systems for Aircraft, Communication Systems and Air Force RTO]. Noginsk, 2012, pp. 112–118 (in Russ.).
6. Dikarev V.A. Features of operators' vocational training using computer training systems. *34 nauchnye chteniya, posvyashchennye razrabotke tvorcheskogo naslediya K.E. Tsiolkovskogo: sb. stat. Vseros. nauch.-tekhnich. konf.* [Proc. All-Russian Science and Technical Conf. 34th Scientific Readings Devoted to K.E. Tsiolkovsky's Creative Heritage Development]. Kaluga, Kaluzhsky dom znany Publ., 1999, pp. 181–184 (in Russ.).
7. Ponomarenko V.A. *Psikhologiya zhizni i truda letchika* [Psychology of Pilot's Life and Work]. Moscow, Voenizdat Publ., 1992, 58 p.
8. Sysoev V.V., Krutskikh P.P., Svintsov A.A. A mathematical model of an information conflict. *Radiotekhnika* [Radio Engineering]. 1999, no. 3, pp. 77–80 (in Russ.).
9. Potapov A.N., Pasmurnov S.M. A mathematical model of the quantitative criterion of simulation modeling adequacy assessment in trainers of operators of ergotechnical systems. *Vestn. Voronezh. gos. tekhnich. univ.* [Bulletin of the Voronezh State Technical Univ.]. 2012, Voronezh, VGTU Publ., vol. 8, no. 3 pp. 4–8 (in Russ.).
10. Gubskov Yu.A., Prokofev S.N. System model of an information conflict stable automated system of ergotechnical complexes acquisition. *Informatika: problemy, metodologiya, tekhnologii: mater. 12 Mezhdunar. nauch.-metodich. konf.* [Proc. 12th int. Science and Methodological Conf. "Informatics: Problems, Methodology, Technologies"]. Voronezh, VSU Publ., 2012, vol. 1, pp. 109–112 (in Russ.).

УДК 621.391

Дата подачи статьи: 08.04.16

DOI: 10.15827/0236-235X.115.129-135

**КОМПЛЕКС ПРОГРАММНЫХ СРЕДСТВ  
ДЛЯ АНАЛИЗА ПРИЧИННО-СЛЕДСТВЕННЫХ ФАКТОРОВ  
КОНФЛИКТА ПРИМЕНЕНИЯ ТРЕНАЖЕРОВ  
ЭРГОТЕХНИЧЕСКИХ РАДИОЭЛЕКТРОННЫХ СРЕДСТВ**

А.Н. Потанов, к.т.н., доцент, зам. начальника кафедры, [potanov\\_il@mail.ru](mailto:potanov_il@mail.ru)  
(Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия  
им. проф. Н.Е. Жуковского и Ю.А. Гагарина»,  
ул. Старых Большевиков, 54а, г. Воронеж, 394064, Россия);

Ф.Н. Абу-Абед, к.т.н., доцент, [aafares@mail.ru](mailto:aafares@mail.ru);

Д.В. Мартынов, к.т.н., доцент, [ido@tstu.tver.ru](mailto:ido@tstu.tver.ru)  
(Тверской государственный технический университет,  
наб. Аф. Никитина, 22, г. Тверь, 170026, Россия);

А.Б. Талалаев, д.т.н., профессор, генеральный директор, [hotline@rtisvko.ru](mailto:hotline@rtisvko.ru)  
(Компания «Радиотехнические и информационные системы воздушно-космической обороны»,  
ул. Зинаиды Коноплянниковой, 89, корп. 1, г. Тверь, 170041, Россия)

В статье рассмотрены причинно-следственные факторы возникновения конфликта применения тренажеров эрготехнических радиоэлектронных средств. Предложена обобщенная классификация возможных навыков, приобретаемых операторами на средствах тренажа. На основании введенной обобщенной классификации навыков конкретизированы факторы возникновения конфликта оценок уровней обученности операторов на тренажерах и штатной технике. Указана необходимость определения количественной формы адекватности тренажеров эрготехнических радиоэлектронных комплексов, которая позволила бы устранить конфликт оценок обученности. Сформулирована вероятностно-метрическая форма оценки адекватности тренажеров эрготехнических радиоэлектронных комплексов в виде математической задачи оценки вероятности предположения о том, что значения выходных имитируемых информационных параметров  $\Xi_m(x)$  отличаются от значений реальных выходных информационных параметров  $\Xi(x)$  при решении единичных задач не более, чем на  $\{0, 1\}$   $\Xi(x)$ .

Для анализа причинно-следственных факторов конфликта применения тренажеров эрготехнических радиоэлектронных средств (на примере радиотехнических средств управления воздушным движением) предложен комплекс программных средств, включающий программно-ориентированный комплекс оценки адекватности автоматизированных информационных систем освоения эрготехнических объектов управления воздушным движением, программу оценки адекватности автоматизированной информационной системы освоения эрготехнического комплекса, имеющего иерархическую структуру построения, и систему оперативного объективного контроля действий операторов информационных комплексов управления воздушным движением.

На основе применения комплекса программных средств, обеспечивающего выполнение анализа причинно-следственных факторов конфликта применения тренажеров эрготехнических радиоэлектронных средств, определены рекомендации по совершенствованию тренажной подготовки операторов эрготехнических радиоэлектронных средств.

**Ключевые слова:** тренажер, математическое моделирование, оценка степени адекватности, информационные системы, навык, радиоэлектронное средство, программное средство.

Использование в тренажере имитационных моделей с ограниченной адекватностью является одной из основных причин возникновения конфликта применения тренажеров, заключающегося в некорректном определении содержательного компонента учебных планов *тренажной подготовки* (ТП) – количества тренировок на средствах тренажа. Вследствие ограниченной адекватности тренажеров наблюдаются взаимозависимые события:

- происходит приобретение операторами неверных навыков по эффективным способам и тактическим приемам применения эрготехнических радиоэлектронных средств (РЭС);

- осуществляется искаженное оценивание обрабатываемых операторами учебных упражнений.

Искаженное оценивание происходит из-за несоответствия между оценками обученности операторов на тренажерах и на *штатной технике* (ШТ) по

единым упражнениям. Причина несоответствия оценок обученности  $\mathfrak{R}(K)$  операторов на средствах тренажа в наличии неучтенных навыков, которые порой являются вредными, хотя воспринимаются как необходимые [1].

Из теории психологии и педагогики [2] известно, что навыки – это умение выполнять целенаправленные действия, доведенное до автоматизма в результате сознательного многократного повторения одних и тех же движений или решения типовых задач.

Применительно к ТП операторов сознательное многократное повторение ими одних и тех же движений или решение типовых задач осуществляется в процессе тренировок на средствах тренажа.

Однако также известно, что действия можно классифицировать как правильные (положительные) и неправильные (отрицательные, ошибочные) [3].

При этом под положительными и отрицательными понимают такие действия операторов, которые обеспечивают соответственно положительный (повышение функции полезности) и отрицательный (снижение функции полезности) эффект применения эргатических объектов.

Поэтому целесообразно ввести понятия навыков согласно классификации их действий [4]:

- положительные навыки  $C$  (доведенные во время тренировок до автоматизма положительные действия);
- отрицательные навыки  $E$  (доведенные во время тренировок до автоматизма отрицательные действия).

С учетом этого конфликт оценки уровней обученности  $\mathcal{R}(K)$  заключается в наличии ошибок идентификации навыков, так как часть положительных навыков  $C$  воспринимается как отрицательные навыки  $E$  и, наоборот, часть отрицательных навыков  $E$  – как положительные навыки  $C$ .

Для формализации этих ошибок уместно ввести следующие понятия:

- ложные действия (ошибка первого рода) – это действия, которые идентифицируются как положительные, хотя на самом деле они отрицательные;
- пропущенные положительные действия (ошибка второго рода) – это действия, которые идентифицируются как отрицательные, хотя на самом деле они положительные.

Согласно такой классификации ошибок идентификации действий, которые, безусловно, преобразуются при доведении действий до автоматизма в ошибки идентификации навыков, уместно ввести определения ошибочных навыков:

- ложные навыки  $F$  – это доведенные во время тренировок до автоматизма отрицательные действия, которые идентифицируются как положительные;
- пропущенные положительные навыки  $D$  – это доведенные во время тренировок до автоматизма положительные действия, которые идентифицируются как отрицательные.

Обобщенная классификация возможных навыков, приобретаемых операторами на средствах тренажа, представлена на рисунке 1.

По своей сути эта классификация навыков образует полную группу событий. Ввод такой обобщенной классификации навыков позволил в соответствии с [4] конкретизировать факторы существования  $\mathcal{R}(K)$ , которые представлены на рисунке 2.

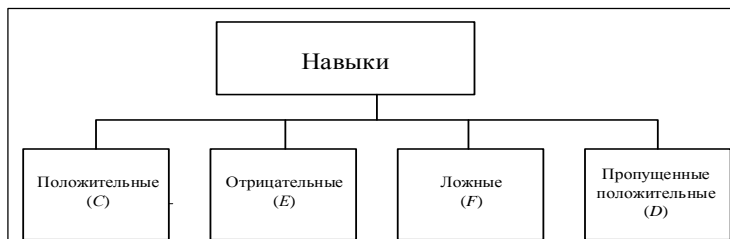


Рис. 1. Классификация навыков, приобретаемых операторами на тренажерах

Fig. 1. Classification of skills acquired by simulator operators

Очевидным является то, что если первые три фактора, присущие как тренажерам, так и ШТ, в силу однородности их происхождения при оценке уровня навыков могут быть скомпенсированы (они не влияют на расхождение оценок навыков, приобретаемых на тренажерах и реальной технике при выполнении условия их идентичности), то последний проявляется только на тренажерах и не может быть скомпенсирован как первые три.

Необходимо отметить, что ложные  $F$  и пропущенные правильные  $D$  являются объективной реальностью независимо от отсутствия возможности их наблюдаемости и следствием ограниченной адекватности тренажеров.

Так как наличие конфликта обученности  $\mathcal{R}(K)$  приводит к возникновению конфликта планирования ТП с использованием тренажеров, актуальной является разработка метода формирования содержательного компонента учебных планов, защищенного от данных конфликтов, существующих из-за их ограниченной адекватности.

Сама адекватность является сопутствующим признаком конфликта применения тренажеров. Но так как существующая понятийная база по оценке адекватности тренажеров не позволяет представить ее в виде индикатора сопутствующего признака конфликта применения тренажеров, первоочередной задачей является определение ее количественной формы, которая позволила бы устранить конфликт оценок обученности.

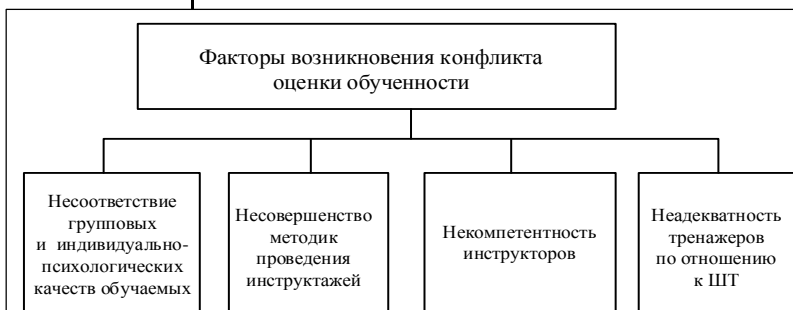


Рис. 2. Факторы возникновения конфликта оценок обученности операторов на тренажерах и ШТ

Fig. 2. Factors of the conflict of operators' training estimates on simulators and equipment

### Понятие адекватности тренажеров

В настоящее время имеется ряд работ, в которых делаются попытки проверки адекватности как имитационного моделирования, так и самих тренажеров [5]. Однако эти попытки ограничиваются определением факта адекватности или неадекватности проверяемых систем, что не может быть использовано для выявления приобретаемых на них навыков в соответствии с введенной обобщенной классификацией.

Для ясности в понимании предлагаемой количественной формы оценки адекватности тренажеров, позволяющей разрешить существующую проблему, обоснуем ее на примере задачи по оценке адекватности тренажеров операторов управления воздушным движением (ВД).

Известно, что тренажеры операторов авиационных комплексов управления ВД предназначены для привития и поддержки у операторов навыков по оценке информации  $\Xi$  и принятию решений  $x$  на использование средств управления ВД. Основой построения данных систем является имитационное моделирование на ЭВМ процессов, происходящих в реальных полетах летательного аппарата при решении соответствующих задач управления ВД. Однако известно, что любая имитационная модель является некоторым прообразом реального процесса и полученные с ее помощью результаты необходимо проверять на предмет адекватности.

Адекватность означает идентичность, неразличимость двух предметов, явлений, процессов по определенным признакам, показателям [5].

Следует заметить, что адекватность можно представить как случай близости модели к объекту. Если имеется возможность ввести метрическое пространство математических образов, описывающих модель и объект, то в качестве меры близости модели к объекту целесообразно использовать метрическое расстояние  $\rho(M_0, M)$  между точками метрического пространства  $M_0$  и  $M$ , соответствующими объекту и модели [5]. Расстояние в метрическом пространстве должно удовлетворять трем условиям:

$$\rho(M_0, M) = 0 \quad (1)$$

тогда и только тогда, когда  $M_0$  и  $M$  – аксиома тождества;

$$\rho(M_0, M) = \rho(M, M_0) \quad (2)$$

тогда и только тогда, когда  $M_0$  и  $M$  – аксиома симметрии;

$$\rho(M_0, M_1) + \rho(M_1, M_2) \geq \rho(M_0, M_2) \quad (3)$$

тогда и только тогда, когда  $M_0$  и  $M$  – аксиома треугольника.

Исходной информацией для оценки адекватности имитационной модели являются входные и выходные сигналы моделируемого объекта и функционал модели. По известным функционалу модели и входным сигналам объекта могут быть получены выходные сигналы модели.

Очевидно, что в данной ситуации адекватность имитационной модели объекту (действительности) может быть оценена только путем сравнения совокупностей входных-выходных сигналов объекта и модели. Расстояние, по которому оценивается идентичность имитационной модели и объекта, должно определяться по данным входных и выходных сигналов.

Если предположить, что входные сигналы объекта  $x$  и имитационной модели  $x_m$  одинаковы ( $x = x_m$ , в детерминированном случае они совпадают, в стохастическом имеют одинаковые вероятностные характеристики), то основная исходная информация для оценки адекватности задается результатами измерений выходных сигналов объекта  $W(x)$  и его модели  $W_m(x)$ .

Обычно оценка адекватности результатов имитационного моделирования объекта сводится к определению рассогласования (невязки) между выходными сигналами объекта  $W(x)$  и его модели  $W_m(x)$  [5], то есть

$$e(x) = W(x) - W_m(x). \quad (4)$$

Невязка  $e(x)$  не удовлетворяет условиям (1)–(3), но, если от нее взять модуль, она будет соответствовать метрическому расстоянию. Поэтому в дальнейшем  $e(x)$  будем называть условной мерой близости модели и объекта.

Очевидно, что при проведении серии испытаний модели и объекта значение математического ожидания условной меры близости модели и объекта  $M\{e(x)\}$  может оказаться равным нулю, хотя на самом деле все ее значения отличны от нуля, что может привести к неверным выводам об адекватности модели. Поэтому при оценке адекватности имитационной модели пользоваться соотношением (4) можно ограниченно. Оно вполне может быть использовано в процессе формализации (представления в математической форме) метрического расстояния между результатами моделирования и откликами реального объекта [6].

Таким образом, под адекватностью результатов имитационного моделирования реального объекта следует понимать метрическую меру близости между выходными сигналами объекта и соответствующими им выходными сигналами ее модели:  $\varepsilon(x) = |W(x) - W_m(x)|$ , где  $W(x) = \{W_i(x)\}_n$  и  $W_m(x) = \{W_{mi}(x)\}_n$  – отклики объекта и имитационной модели (получены в ходе  $i$ -го физического и имитационного экспериментов,  $i = 1, 2, \dots, n$ ) [6].

Математическая модель  $M$  абсолютно адекватна объекту  $M_0$ , если метрическая мера близости равна нулю [6], то есть  $\varepsilon(x) = 0$ . В силу ряда причин (принципиального различия модели и объекта, неопределенности исходных данных, погрешности реализации модели на ПЭВМ, измерительных и вычислительных погрешностей и т.д.) вероятность выполнения условия  $\varepsilon(x) = 0$  является, как правило, нулевой.

Поэтому пользуются условием адекватности [7] следующего вида:

$$|W(x) - W_m(x)| \leq \varepsilon^*(W(x)), \quad (5)$$

где  $\varepsilon^*(W(x))$  – требуемая точность имитационной модели объекта (при решении, например, задач управления ВД может выбираться по тактико-техническим соображениям).

При выполнении условия (5) считается, что математическая модель адекватна объекту с точностью  $\varepsilon^*(W(x))$ . Пользоваться имитационной моделью рекомендуется, если вероятность выполнения условия (5) больше 0,5 [7].

Следовательно, для допуска имитационной модели к практическому использованию необходимо производить оценку вероятности выполнения этого условия. С учетом этого введем такое понятие, как вероятность адекватности имитационной модели объекту: вероятностью адекватности  $\gamma(W(x))$  модели объекту называется вероятность  $P$  выполнения того, что  $\varepsilon(x) \leq \varepsilon^*(W(x))$  [7], то есть

$$\gamma(W(x)) = P[|W(x) - W_m(x)| \leq \varepsilon^*(W(x))]. \quad (6)$$

Другими словами, вероятность адекватности  $\gamma(W(x))$  модели объекту есть не что иное, как достоверная вероятность выполнения условия (5).

Для характеристики степени (вероятности) неадекватности имитационной модели объекту определим понятие вероятности  $\bar{\gamma}(W(x))$  неадекватности: вероятностью неадекватности  $\bar{\gamma}(W(x))$  имитационной модели объекту называется вероятность  $P$  выполнения того, что  $\varepsilon(x) > \varepsilon^*(W(x))$ , то есть

$$\begin{aligned} \bar{\gamma}(W(x)) &= P[|W(x) - W_m(x)| > \varepsilon^*(W(x))] = \\ &= 1 - P[|W(x) - W_m(x)| \leq \varepsilon^*(W(x))]. \end{aligned}$$

Математически задача оценки адекватности имитационной модели действительности формулируется как задача оценки вероятности того, что значение отклика исследуемого параметра  $W_m(x)$  модели отличается от реального отклика  $W(x)$  объекта не более, чем на заданную величину  $\varepsilon^*(W(x))$ .

Известно, что для обеспечения адекватности имитационной модели в процессе ее создания возникает необходимость в идентификации объекта моделирования. Идентификация объекта, как правило, проводится с целью выявления функциональных связей между элементами его структуры и определения (уточнения) его характеристик. На основании этого производится корректировка имитационной модели. Точность идентификации оценивается некоторой функцией от меры близости системы и модели  $Q(\varepsilon(x))$ , обычно называемой функцией потерь. Качество решения задачи оценки идентификации в среднем для заданной реализации входного сигнала  $x$  удобно оценивать условным математическим ожиданием от функции потерь для данного входного сигнала  $r(Q/x) = M\{Q(\varepsilon(x))\}$ , обычно называемым условным риском. Он определяется близостью характеристик

объекта и модели. Среднее качество решения задачи оценки адекватности при всех возможных входных сигналах характеризуется математическим ожиданием условного риска, равным безусловному ожиданию функции потерь (средний риск):  $R(Q) = M\{r(Q/x)\} = M\{Q(\varepsilon(x))\}$ .

Если принять, что функция потерь имеет вид

$$r(Q/x) = \begin{cases} 1, & \text{если } |W(x) - W_m(x)| \leq \varepsilon^*(W(x)), \\ 0, & \text{если } |W(x) - W_m(x)| > \varepsilon^*(W(x)), \end{cases}$$

то значение среднего риска

$$\begin{aligned} R(Q) &= 1 \cdot P[|W(x) - W_m(x)| \leq \varepsilon^*(W(x))] + \\ &+ 0 \cdot P[|W(x) - W_m(x)| > \varepsilon^*(W(x))] = \\ &= P[|W(x) - W_m(x)| \leq \varepsilon^*(W(x))], \end{aligned}$$

представляет собой не что иное, как вероятность адекватности (6).

Достоверность выявления функциональных связей между элементами структуры объекта и определения ее характеристик зависит от точности идентификации – вероятности адекватности моделирования объекта. Если в качестве входных  $x$  и выходных  $\Xi(x)$  параметров для полета на ЛА принять соответствующие параметры  $x_m$  и  $\Xi_m(x_m)$  тренажеров РЭБ, причем  $x = x_m$ , то количественную оценку адекватности модулирующего устройства [6] следует производить в форме (4)

$$\gamma(\Xi(x)) = P[|\Xi(x) - \Xi_m(x)| \leq \varepsilon^*(\Xi(x))], \quad (7)$$

где  $\varepsilon^*(\Xi(x)) = (0, 1 \dots 0, 15) \cdot \Xi(x)$  – требуемая точность имитационного моделирования в тренажере, при которой оператор не чувствителен к различию информационных параметров  $\Xi(x)$  и  $\Xi_m(x)$ , определенная нормой годности тренажеров [7]:

$$\frac{|\Xi(x) - \Xi_m(x)|}{\Xi(x)} \cdot 100 \% \leq \{10 \div 15\} \% . \quad (8)$$

Возможна такая ситуация, когда при определенных действиях  $x$  оператора условие годности тренажера (8) не выполняется. Однако в процессе тренировок на тренажере оператор может осуществлять такие действия  $x'$ , которые по содержательной компоненте отличаются от требуемых действий  $x$  для сложившейся обстановки, определяющейся  $\Xi$ , но и для них

$$\frac{|\Xi(x) - \Xi_m(x')|}{\Xi(x)} \cdot 100 \% \leq \{10 \div 15\} \% .$$

Если действия  $x'$  по отношению к требуемым действиям  $x$  для сложившейся обстановки  $\Xi$  приводят к снижению функции полезности ШТ ( $q_{\Xi}(x') < q_{\Xi}(x)$ ), то эти действия  $x'$  являются отрицательными. А так как в процессе тренировок на тренажере отрицательные действия  $x'$  доводятся до автоматизма, по ним будут сформированы отрицательные навыки. С учетом этого при боевом применении ШТ по информации  $\Xi$  оператор будет выполнять действия  $x'$ , а не адекватные ей действия  $x$ .



Оценка адекватности  $\gamma$  характеризует вероятность того, что оператор не будет чувствителен к различию между  $\Xi(x)$  и  $\Xi_m(x)$ . А это не приводит к выполнению на тренажере действий, отличных от действий на ШТ, которые для последней являются правильными. Другими словами, величина  $\gamma$  определяет потенциальный уровень положительных  $S$  навыков, приобретаемых оператором на тренажере.

Согласно соотношению (7), математически задача оценки адекватности моделирующего устройства тренажера формулируется как задача оценки вероятности предположения о том, что значения выходных имитируемых информационных параметров  $\Xi_m(x)$  отличаются от значений реальных выходных информационных параметров  $\Xi(x)$  при решении единых задач не более, чем на  $\{0,1-0,15\} \Xi(x)$ .

Таким образом, вероятность адекватности тренажера – это степень идентичности (неразличимости) информационных параметров и психолого-физиологических процессов тренажера по отношению к ШТ, при которой оператор совершает действия, соответствующие реальной обстановке.

#### **Комплекс программных средств для анализа причинно-следственных факторов конфликта применения тренажеров эрготехнических РЭС**

В рамках реализации требований к структурной конфигурации элементов конфликтно-устойчивой АСУ ТП операторов РЭС и для анализа причинно-следственных факторов конфликта применения тренажеров эрготехнических РЭС разработан комплекс программ для ЭВМ, часть которых имеют официальную регистрацию (рис. 3) [8–10].

Программно-ориентированный комплекс оценки адекватности автоматизированных информационных систем освоения эрготехнических объектов управления ВД [8] позволяет выполнять оценку адекватности имитационного моделирования в тренажерах операторов управления ВД (лиц группы руководства полетами) по информационному полю.

Программа оценки адекватности автоматизированной информационной системы освоения эрготехнического комплекса, имеющего иерархическую структуру построения, «Адекватность АИС и ЭТК-иерархия» [8], предназначена для оценки адекватности автоматизированных информационных систем подготовки операторов РЭС, имеющих иерархическую структуру.

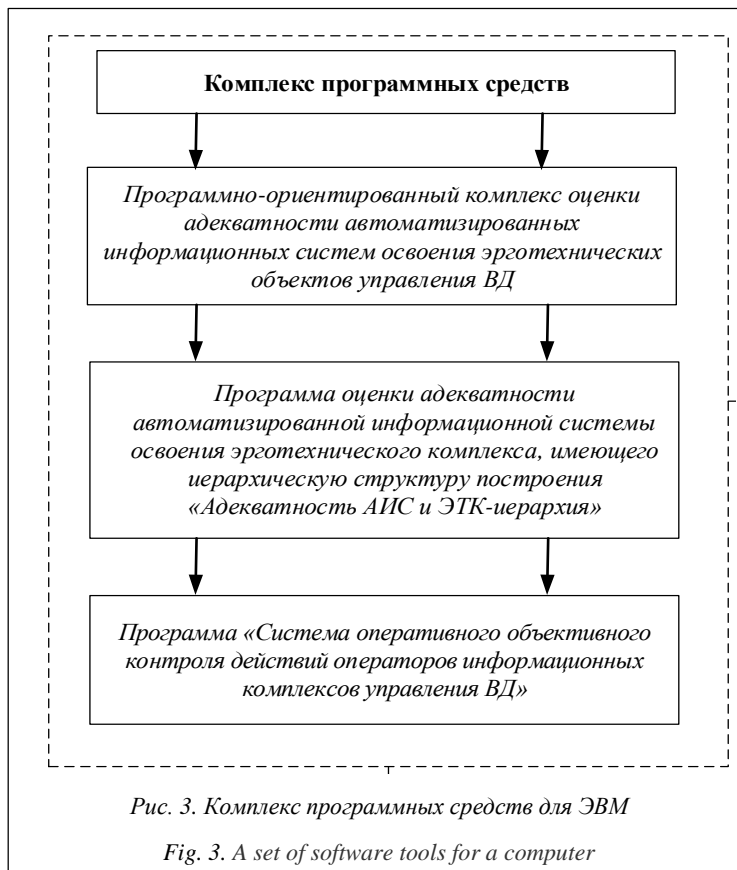


Рис. 3. Комплекс программных средств для ЭВМ

Fig. 3. A set of software tools for a computer

Программа «Система оперативного объективного контроля действий операторов информационных комплексов управления ВД» [10] предназначена для оперативного объективного контроля действий операторов управления ВД на основе учета оценки адекватности автоматизированных информационных систем подготовки операторов управления ВД (лиц ГРП) в целях разрешения конфликтов использования ШТ и тренажеров в процессе практической подготовки операторов управления ВД.

На рисунке 4 показан интерфейс системы оперативного объективного контроля действий операторов информационных комплексов управления ВД.

На основе разработанного комплекса программных средств для анализа причинно-следственных факторов конфликта применения тренажеров эрготехнических РЭС в целях повышения оперативности планирования тренажной подготовки, заключающейся в гибкой адаптации учебных планов с учетом индивидуальных особенностей обучаемых, целесообразно сформулировать рекомендации по управлению планированием практической подготовки операторов эрготехнических РЭС:

- определение всех возможных вариантов подготовки операторов с использованием тренажеров;
- оценка адекватности тренажеров операторов эрготехнических систем;

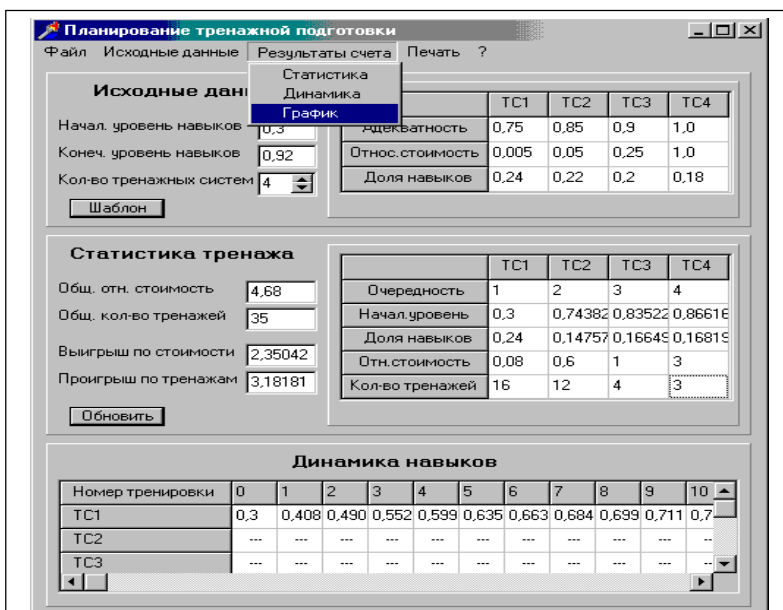


Рис. 4. Интерфейс программы «Система оперативного объективного контроля действий операторов информационных комплексов управления ВД»

Fig. 4. The interface of the program "System for operational objective control of actions of operators of information systems in the Department of Internal Affairs"

- учет ограничений в использовании тренажных комплексов;
- оценка экономической эффективности тренажных средств обучения;
- оценка доли навыков, приобретаемых на различных типах тренажеров.

Таким образом, для разрешения существующего конфликта применения тренажеров необходимо производить оценку как уровня обученности, так и доли навыков, приобретаемых операторами на тренажных системах с учетом их адекватности по отношению к ШТ.

#### Литература

1. Дудоров А.Д. Методика оценки тренажеров лиц группы руководства полетами в интересах повышения качества их под-

готовки // Современные проблемы и перспективные направления развития авиационных комплексов и систем военного назначения, формы и способы их боевого применения: сб. стат. Всерос. науч.-практич. конф. Воронеж: Изд-во ВАИУ, 2011. С. 92–98.

2. Сысоев Е.С. Оптимизация тренажерной подготовки операторов сложных информационных радиоэлектронных систем управления воздушным движением // Информатика: проблемы, методология, технологии: матер. 11 Междунар. науч.-методич. конф. Т. 2. 2011. Воронеж: Изд-во ВГУ. С. 216–222.

3. Лебедев В.В., Романов А.В. Теоретико-множественный подход к описанию функционирования эрготехнических радиоэлектронных средств // Проблемы эффективности и безопасности функционирования сложных технических и информационных систем: тр. 32 Всерос. науч.-технич. конф. Серпухов: Изд-во ВАРВСН им. Петра Великого, 2013. С. 118–122.

4. Помазуев О.Н., Платонов А.Ю., Мионов А.М., Абу-Абед Ф.Н., Рюмшин А.Р. Методика определения показателей качества обнаружения радиолокационных станций // Программные продукты и системы. 2014. № 2. С. 118–124.

5. Агапов Е.А. Проектирование автоматизированных информационных систем освоения эрготехнических комплексов // Эффективность автоматизированных систем управления авиацией, систем связи и РТО ВВС. Ногинск, 2012. С. 112–118.

6. Дикарев В.А. Особенности профессиональной подготовки операторов при использовании компьютерных систем тренажа // XXXIV научные чтения, посвященные разработке творческого наследия К.Э. Циолковского: сб. стат. Всерос. науч.-технич. конф. 1999. Калуга: Калужский дом знаний, 1999. С. 181–184.

7. Пономаренко В.А. Психология жизни и труда летчика. М.: Воениздат, 1992. 58 с.

8. Сысоев В.В., Крутских П.П., Свинцов А.А. Математическая модель информационного конфликта // Радиотехника (журнал в журнале). 1999. № 3. С. 77–80.

9. Губсков Ю.А., Прокофьев С.Н. Системная модель информационной конфликтно-устойчивой автоматизированной системы освоения эрготехнических комплексов // Информатика: проблемы, методология, технологии: матер. 12 Междунар. науч.-методич. конф. Воронеж: Изд-во ВГУ. 2012. Т. 1. С. 109–112.

10. Потапов А.Н., Пасмурнов С.М. Математическая модель количественного критерия оценки адекватности имитационного моделирования в тренажерах операторов эрготехнических систем // Вестн. Воронеж. гос. технич. ун-та. 2012. Т. 8. № 3. С. 4–8.

DOI: 10.15827/0236-235X.115.129-135

Received 08.04.16

### A SOFTWARE SUITE TO ANALYZE CAUSE-EFFECT FACTORS OF A CONFLICT OF ERGOTECHNICAL RADIO-ELECTRONIC EQUIPMENT SIMULATORS

<sup>1</sup>Potapov A.N., Ph.D. (Engineering), Associate Professor, Deputy Head of Chair, potapov\_il@mail.ru

<sup>2</sup>Abu-Abed F.N., Ph.D. (Engineering), Associate Professor, aafares@mail.ru

<sup>3</sup>Martynov D.V., Ph.D. (Engineering), Associate Professor, idpo@tstu.tver.ru

<sup>4</sup>Talalayev A.B., Dr.Sc. (Engineering), Professor, Director General, hotline@rtisvko.ru

<sup>1</sup>Military Scholastic-Scientific Centre of the Air Forces "Zhukovsky and Gagarin Air Forces Academy", Starykh Bolshevikov St. 54a, Voronezh, 394064, Russian Federation

<sup>2</sup>Tver State Technical University, Nikitin Quay 22, Tver, 170026, Russian Federation

<sup>3</sup>Radiotechnical and Information Systems Aerospace Defence, Z. Konoplyannikova St. 89/1, Tver, 170041, Russian Federation

**Abstract.** The article considers cause-effect factors of conflict in using ergotechnical radio-electronic equipment (REE) simulators. It proposes a generalized classification of possible skills, which can be acquired by operators on

simulators. Based on the generalized classification of skills the authors define specific factors causing the conflict of evaluating operators' training levels on simulators and organic equipment. The paper indicates the need to the adequacy of ergotechnical electronic system simulators, which would eliminate the conflict of training evaluations. It also determines a probabilistic-metric evaluation of ergotechnical electronic system simulators' adequacy in the form of a mathematical problem of estimating the probability of the assumption that the values of simulated output information parameters  $\Xi_m(x)$  differ from the actual output values of information parameters  $\Xi_m(x)$  when solving a single task, but no more than for  $\{0,1\}$ .

To analyze cause-effect factors of a conflict when applying ergotechnical radio-electronic equipment simulators (for example, air traffic control equipment) the authors propose a software suit that includes: a program-oriented system of evaluating the adequacy of automated information systems for training ergotechnical air traffic control objects; a program for assessing the adequacy of an automated information system for developing ergotechnical complex with a hierarchical structure; an operational system for objective controlling the actions of operators of air traffic control information systems.

Based on applying the software suite, which performs the analysis of cause-effect factors of a ergotechnical radio-electronic equipment conflict, the article specifies recommendations for improving ergotechnical REE operators' training.

**Keywords:** simulator, mathematical modeling, adequacy assessment, information systems, radio-electronic equipment, software means.

### References

1. Dudorov A.D. An evaluation method of management operations group equipment in order to improve their preparation quality. *Sovremennye problemy i perspektivnye napravleniya razvitiya aviatsionnykh kompleksov i sistem voennogo naznacheniya, formy i sposoby ikh boevogo primeneniya: sb. stat. Vseros. nauch.-praktich. konf.* [Proc. All-Russian Science and Practice Conf. "Modern problems and promising directions of development of aircraft systems and military systems, forms and methods of their combat use"]. Voronezh, VAIU Publ., 2011, pp. 92–98 (in Russ.).
2. Sysoev E.S. Optimization of simulator training of operators of complex electronic information systems air traffic control. *Informatika: problemy, metodologiya, tekhnologii: mater. 11 Mezhdunar. nauch.-metodich. konf.* [Proc. 11 Int. Science Conf. "Informatics: Problems, Methodology, Technologies"]. 2011, vol. 2, Voronezh, VSU Publ., pp. 216–222 (in Russ.).
3. Lebedev V.V., Romanov V.A. A set-theoretic approach to the description of ergo-technical operation of radio-electronic equipment. *Problemy effektivnosti i bezopasnosti funktsionirovaniya slozhnykh tekhnicheskikh i informatsionnykh sistem: tr. 32 Vseros. nauch.-tekhnich. konf.* [Proc. 32nd All-Russian Science and Technical Conf. "Problems of Efficiency and Safety of Difficult Technical and Information System Functioning"]. Serpukhov, Peter the Great Military Academy of Strategic Missile Forces Publ., 2013, pp. 118–122 (in Russ.).
4. Pomazuev O.N., Platonov A.Yu., Mironov A.M., Abu-Abed F.N., Ryumshin A.R. The method of quality factors estimation for radar detection. *Programmnye produkty i sistemy* [Software & Systems]. 2014, no. 2, pp. 118–124 (in Russ.).
5. Agapov E.A. Design of automated information systems for ergo technical complexes development. *Effektivnost avtomatizirovannykh sistem upravleniya aviatsiey, sistem svyazi i RTO VVS* [The Efficiency of Aircraft Automated Control Systems, Communications Systems and Radio Equipment of the Air Forces]. Noginsk, 2012, pp. 112–118 (in Russ.).
6. Dikarev V.A. Peculiarities of training operators when using computer training systems. *34 nauchnye chteniya, posvyashchennye razrabotke tvorcheskogo naslediya K.E. Tsiolkovskogo: sb. stat. Vseros. nauch.-tekhnich. konf.* [Proc. All-Russian scientific-technical conference "XXXIV scientific readings, devoted to the development of the creative heritage of K.E. Tsiolkovsky"]. 1999, Kaluga, Kaluzhsky dom znany Publ., 1999, pp. 181–184 (in Russ.).
7. Ponomarenko V.A. *Psikhologiya zhizni i truda letchika* [Psychology of Pilot's Life and Work]. Moscow, Voenizdat Publ., 1992, 58 p.
8. Sysoev V.V., Krutskikh P.P., Svintsov A.A. A mathematical model of an information conflict. *Radiotekhnika* [Radio Engineering]. 1999, no. 3, pp. 77–80 (in Russ.).
9. Gubskov Yu.A., Prokofev S.N. System model of an information conflict stable automated system of ergotechnical complexes acquisition. *Informatika: problemy, metodologiya, tekhnologii: mater. 12 Mezhdunar. nauch.-metodich. konf.* [Proc. 12th int. Science and Methodological Conf. "Informatics: Problems, Methodology, Technologies"]. Voronezh, VSU Publ., 2012, vol. 1, pp. 109–112 (in Russ.).
10. Potapov A.N., Pasmurnov S.M. A mathematical model of the quantitative criterion of simulation modeling adequacy assessment in trainers of operators of ergotechnical systems. *Vestn. Voronezh. gos. tekhnich. univ.* [Bulletin of the Voronezh State Technical Univ.]. 2012, Voronezh, VGTU Publ., vol. 8, no. 3, pp. 4–8 (in Russ.).

УДК 004.4'2

Дата подачи статьи: 16.11.15

DOI: 10.15827/0236-235X.115.136-141

## **СПЕЦИАЛИЗИРОВАННЫЙ ПРОГРАММНЫЙ КОМПЛЕКС «АСТ» ДЛЯ СОЗДАНИЯ ТРЕНАЖЕРОВ**

*П.Ю. Вильвер, программист, wilwer@icc.ru;*

*А.Ю. Юрин, к.т.н., доцент, j80@yandex.ru*

*(Институт динамики систем и теории управления им. В.М. Матросова СО РАН,  
ул. Лермонтова, 134, г. Иркутск, 664033, Россия)*

В статье описан специализированный программный комплекс «АСТ» (Автоматизированное создание тренажеров) для создания компьютерных тренажеров с функциями поддержки принятия решений. Он основан на использовании формализма цветных сетей Петри для моделирования технических систем.

Комплекс обеспечивает следующее: создание иерархической статической модели технических систем в виде мнемосхем, схем и чертежей при помощи как встроенного графического редактора, так и системы автоматизированного проектирования AutoCad; создание динамической модели технических систем в виде сети Петри; визуальное моделирование статических и динамических свойств технических систем в результате взаимодействия (или имитации взаимодействия) с ОРС-сервером; прогнозирование параметров технических систем и построение деревьев событий.

Приведено описание функций программного инструментального комплекса, его архитектуры и методики создания тренажерных комплексов. Рассмотрены примеры применения комплекса для создания программных тренажеров по обучению специалистов послеремонтного испытания компрессоров и контроля ректизол-процесса при производстве сжиженного природного газа.

**Ключевые слова:** цветные сети Петри, программный тренажер, автоматизированное создание, тепломеханические, химико-технологические установки, ОРС-сервер, обучение.

Одной из причин нарушения безопасной эксплуатации технологических установок и оборудования предприятий и потерь от простоя и аварий являются ошибки персонала, управляющего технологическим процессом. Аварии приводят к гибели людей, нарушению технологических регламентов, загрязнению окружающей среды. В связи с этим необходимо постоянно поддерживать и повышать уровень подготовки операторов, для чего применяются либо сложные системы тестов и экзаменов, либо аппаратно-программные тренажеры. Представляется, что применение тренажеров является более эффективным методом обучения, способствующим выработке и закреплению моторных навыков при работе со SCADA-системами и АСУ технологическим процессом.

В настоящий момент на рынке тренажеров много предложений готовых решений для различных задач, например [1]. Однако большинство из них при довольно высокой стоимости (так, цена статического тренажера автомобильного крана – 750 000 руб.) требуют доработки и не учитывают специфику решаемых задач, например, в области тепломеханических, химико-технологических установок. Поэтому целесообразно сотрудникам предприятий самим создавать тренажеры, что трудновыполнимо из-за отсутствия предложений программных систем для создания тренажерных комплексов, так как разработчики тренажеров не распространяют используемый инструментарий (за исключением научных разработок), и отсутствия на предприятиях квалифицированных специалистов, имеющих знания и практические навыки в области моделирования, программирования, надежности и технологических процессов.

Таким образом, актуальна разработка инструментальных программных систем, обеспечивающих создание тренажерных комплексов непрограммирующими пользователями и включающих подсистемы моделирования технологических процессов, в частности, для тепломеханических, химико-технологических установок.

### **Программный комплекс «АСТ»**

**Функции и назначение.** Для обеспечения сотрудников предприятий возможностью самостоятельно создавать тренажеры для персонала тепломеханических, химико-технологических установок был разработан специализированный программный комплекс (СПК).

Особенностью СПК является использование обобщенной модели *технической системы* (ТС) [2], являющейся основой для описания моделируемых установок с помощью модифицированных цветных сетей Петри [3]. Реализованная модификация сетей Петри позволяет создавать иерархические модели, получать в качестве атрибутов меток данные с ОРС-сервера и прогнозировать состояния моделируемой ТС [4].

Основные функции СПК:

- возможность создания моделей сложных ТС непрограммирующим пользователем, ограниченная только вычислительной мощностью используемых компьютеров, так как применяемый для описания модели язык цветных сетей Петри с приоритетами равен по мощности машинам Тьюринга [5, 6];
- визуальное моделирование технологического процесса как в форме сетевого представле-

ния в виде цветных сетей Петри, так и в форме мнемосхем;

- прогнозирование значений параметров технологического процесса на основе статистического прогнозирования, методов последовательных уступок и взвешивания критериев [7–9];

- поддержка оператора при принятии решений путем информационной и цветовой индикации отклонения значений параметров технологического процесса от допустимых.

**Архитектура СПК**, реализующая данные функции, представлена на рисунке 1.

Рассмотрим основные модули и их назначение.

**Подсистема создания модели** обеспечивает создание статической и динамической моделей ТС. Для создания статической модели, представляющей собой статическую схему (мнемосхему или чертеж), может быть применен как встроенный графический редактор, так и внешний, например AutoCad. В частности, использование объектной модели AutoCad обеспечивает доступ ко всем элементам как чертежа, так и самой оболочки AutoCad, что позволяет использовать уже готовые чертежи и схемы. Динамическая модель создается с помощью редактора сетей Петри, обеспечивающего наложение вершин и переходов на статическую модель, а также задание источника данных для модели (внешний OPC-сервер или его имитация).

**БД моделей** обеспечивает хранение и повторное использование разработанных ранее моделей. В состав каждой модели входят сетевая модель ТС на языке сетей Петри (динамическая модель ТС), графические схемы элементов ТС (статическая модель), список подключаемых к сетевой модели параметров OPC-сервера (или имитатора данных), последние значения параметров ТС (количество определяется настройками модели, от 1 до 1000).

Также БД содержит стандартизированные согласно ГОСТ 21.101-2013 и ГОСТ 27833-88 изображения элементов ТС, используемые для создания графических схем.

**Подсистема имитационного моделирования** обеспечивает интерпретацию модели с учетом данных от OPC-сервера. При получении данных от OPC-сервера интерпретатор модифицированной сети Петри обеспечивает перемещение меток и вычисление их параметров, а также производит «вызывание» меток модели и параметров OPC-клиента, определение выполнения условий перехода меток и вычисление атрибутов меток в соответствии с функциями, определенными в сетевой модели ТС.

**Библиотека математических модулей** содержит алгоритмы для работы с матрицами, метод наименьших объемов и др.

**Модуль визуализации результатов моделирования** обеспечивает обновление мнемосхемы в соответствии с изменением состояния модели, а также построение графиков изменения параметров.

**Модуль поддержки принятия решений** обеспечивает построение *деревьев сценариев* (ДС), *частичных деревьев отказов* (ДО), *деревьев достижимых маркировок* (ДДМ) и экспертную оценку сценариев по уровню безопасности [8].

Построение ДС и ДДМ может автоматически производиться при выходе параметра за допустимые пределы.

**Модуль имитации OPC-сервера** является источником данных для модели, обеспечивающим генерацию набора аналоговых и дискретных сигналов от датчиков оборудования.

**OPC-сервер** – реальный источник данных, при реализации OPC-клиента для этого сервера в СПК использовались контроллеры NL. Синхронизация состояния модели осуществляется только в соответствии с параметрами датчиков

и контроллеров, время в модели не используется. OPC-клиент СПК может получать данные с нескольких OPC-серверов.

**Модуль прогнозирования** обеспечивает использование статистических методов (взвешивания критериев и метод уступок) и нейронной сети для прогнозирования состояния модели ТС.

**Методика создания компьютерного тренажера на основе СПК** представляет собой следующую последовательность шагов.

1. Создание статической модели ТС с помощью графического редактора

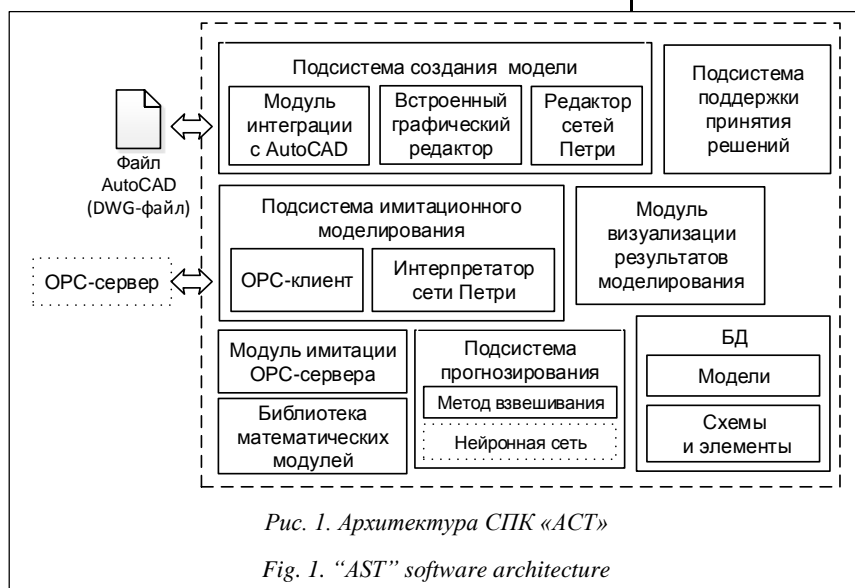


Рис. 1. Архитектура СПК «АСТ»

Fig. 1. "AST" software architecture

ческого средства моделирования (AutoCad или встроенного графического редактора) в форме схемы, обеспечивающей наглядное отображение структуры системы.

- Модель отражает объекты  $P_k$ , события  $T_j$  и параметры  $A_i$  системы, где  $k$  – количество объектов,  $j$  – количество событий,  $i$  – количество параметров:

- объекты ТС имеют параметры, соответствующие значениям атрибутов  $A_i$  метки  $m_n$ , находящейся в позиции  $P_k$ ;  $n$  определяет число меток, принадлежащих (находящихся в) позиции  $P_k$ , значения которых могут изменяться с течением времени в зависимости от данных, получаемых от ОРС-сервера (при отсутствии ОРС-сервера используется модуль генерации данных);

- событие  $T_j$  происходит при выполнении условий на переход меток  $m_n$  между объектами  $P_k$ ;

- значения параметров объектов ТС поступают с первичных приборов или контроллеров на ОРС-сервер и затем могут быть использованы в модели как значения атрибутов  $A_i$  меток  $m_n$ ;

- события происходят или не происходят в соответствии с условиями дуг  $I$ ,  $O$ , которые определены в модели выражениями входных и выходных дуг перехода  $T_j$ ;

- все события в модели происходят в соответствии с порядком функционирования вложенных сетей  $NF$ .

- Модель ТС является иерархической; количество уровней зависит от сложности ТС и поставленной задачи. Разделение ТС на объекты доста-

точно условно и зависит от постановки задачи, например, при анализе технологической линии ее объектами могут считаться отдельные установки и станки, транспортные и загрузочные устройства. В свою очередь, станки и загрузочные устройства также могут считаться ТС и при оценке их надежности должны быть разделены на элементы – узлы, блоки, которые, в свою очередь – на детали и т.д.

2. На основе статической модели создается динамическая модель в форме сети Петри, описывающая процесс функционирования ТС (рис. 2):

- задается начальное состояние системы;

- задаются интервалы номинальных, допустимых и недопустимых значений параметров;

- задаются функции, определяющие значения атрибутов  $A_i$  меток;

- выполняются активные переходы  $T_j$ .

3. Настраивается (задается) ОРС-сервер или имитирующий его модуль.

Успешное завершение данных шагов позволяет получить тренажер, процесс функционирования которого представляет собой взаимодействие динамической модели (представленной в виде сети Петри) и ОРС-сервера, в ходе которого выполняются следующие действия.

1. Обмен данными атрибутов меток и ОРС-сервера во время выполнения переходов. Предыдущие значения атрибутов  $A_i$  сохраняются, количество сохраненных значений определяется настройками модели.

2. Перерасчет значений атрибутов меток.

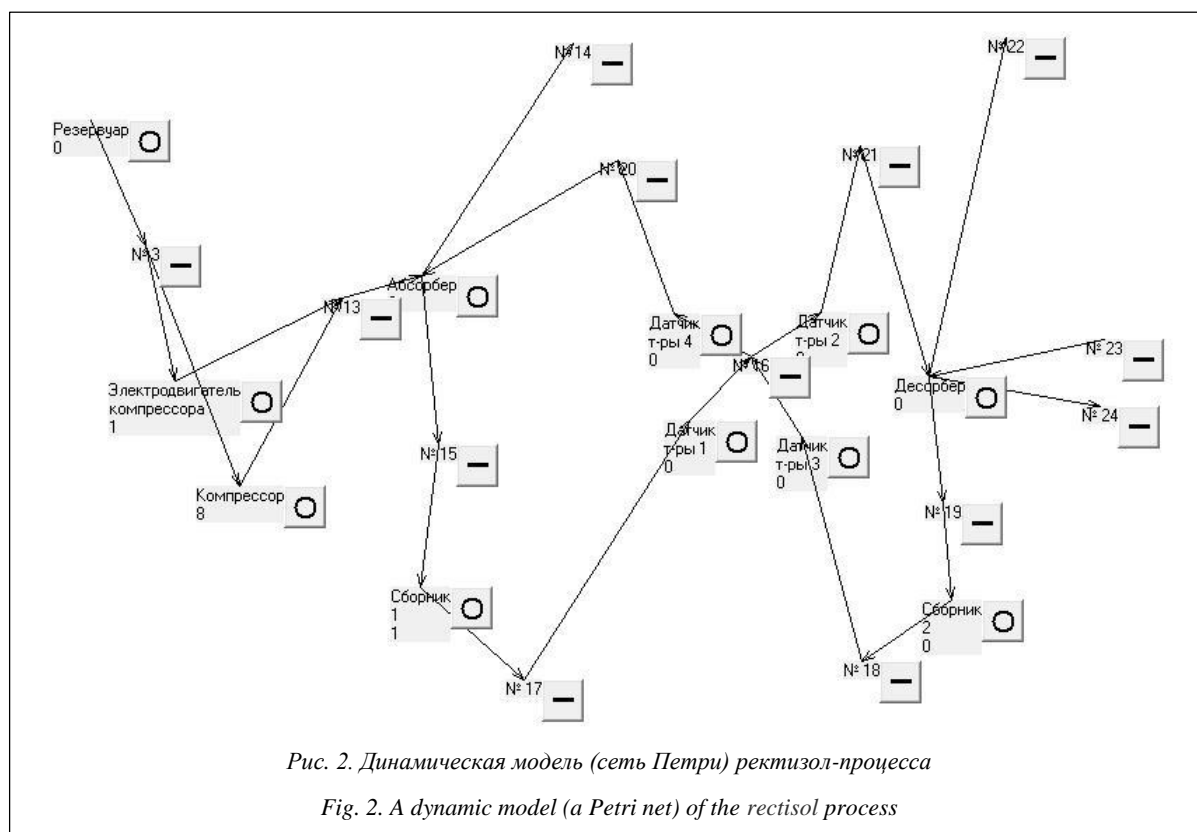


Рис. 2. Динамическая модель (сеть Петри) ректизол-процесса

Fig. 2. A dynamic model (a Petri net) of the rectisol process



3. Прогноз значений атрибута меток (то есть появляется информация о будущих событиях в модели).

4. При выходе значения параметра  $A_i$  объекта  $P_k$  ТС за допустимые или регламентируемые границы  $Z_{Ai}^{Pk}$  определение последовательности произошедших событий:

- определяются наименования и значения параметров ТС, изменение которых влияет на значение атрибута  $A_i$  метки  $m_n$  элемента модели ТС;
- определяются значения параметров объектов модели ТС, достижимые из заданного состояния; выбирается нужное состояние модели;
- определяется последовательность событий, приводящих модель ТС в это состояние.

5. При нарушении регламентируемой последовательности событий определение:

- условия, необходимого, но не выполненного для возникновения события;
- условий возникновения нерегламентируемого события;
- параметров объектов системы, значения которых влияют на выполнение данного условия.

Обучение и контроль знаний оператора на компьютерных тренажерах заключаются в изучении структурных связей между объектами ТС, действующими в технологическом процессе (например, при передаче вращающего момента, жидкости) с последующей выработкой навыков поведения в аварийных ситуациях. Применение компьютерного тренажера, обеспечивающего прогнозирование значений параметров объектов ТС и определение причин инцидентов (и аварийных ситуаций), повышает эффективность обучения за счет оперативного обнаружения ошибок управле-

ния. Использование аварийных сценариев при работе на тренажере позволяет обучаемому персоналу получить навыки устранения подобных событий. Принудительное задание аварийных значений параметров может быть произведено как при составлении исходного задания оператору, так и во время работы тренажера. При составлении задания также могут быть определены условия возникновения аварийной ситуации – аварийное отключение, разгерметизация и т.д. Оценка действий оператора задается при настройке СПК и определяется тем, какой именно параметр и насколько не соответствует регламентируемым значениям.

Программная реализация СПК осуществлена в среде Borland Delphi.

### Примеры применения СПК

С помощью описанной методики были созданы учебные тренажерные комплексы с модулем поддержки принятия решений для контроля ректизол-процесса при производстве сжиженного природного газа, а также послеремонтного испытания компрессоров и лазерного измерительного комплекса.

**Ректизол-процесс** – абсорбция сероводорода, сераорганических соединений и  $\text{CO}_2$  метанолом при низких температурах и повышенных давлениях; является частью технологического процесса при производстве сжиженного природного газа и применяется для очистки природного газа от  $\text{H}_2\text{S}$  и  $\text{CO}_2$ . Технологическая схема включает две ступени: газ при давлении 20–25 атм. и температуре  $-35^\circ\text{C}$  поступает в абсорбер, орошаемый увлажненным холодным метанолом, и десорбер  $\text{H}_2\text{S}$ .

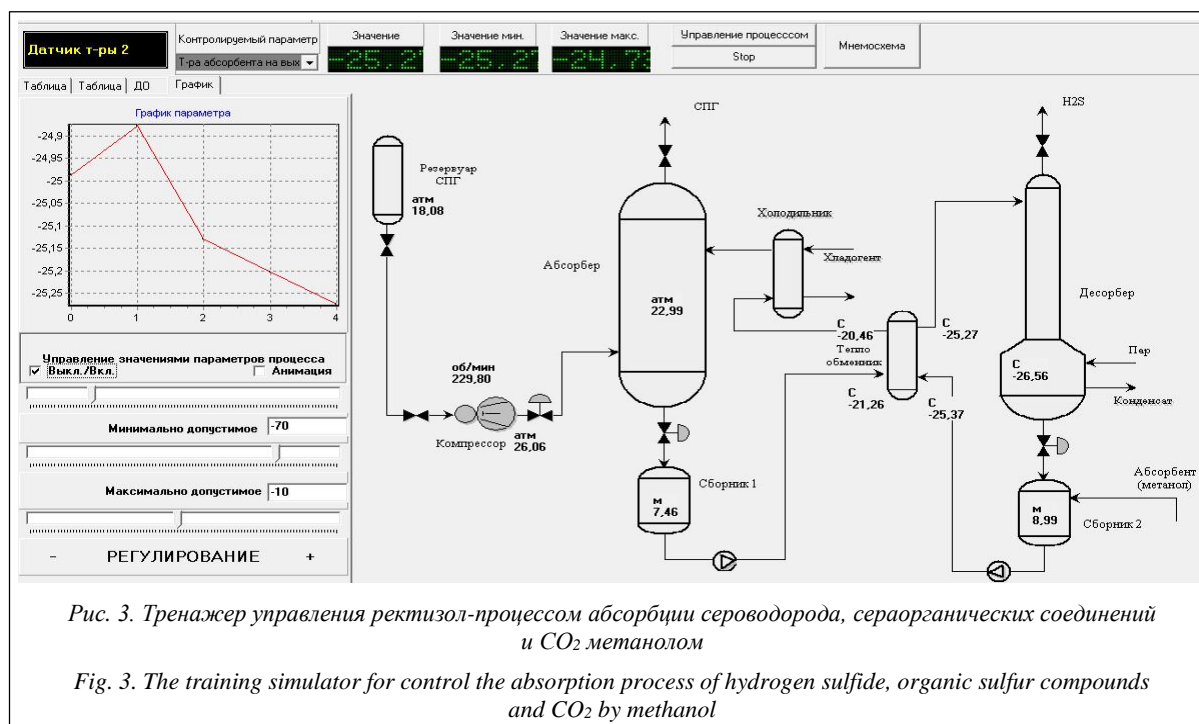


Рис. 3. Тренажер управления ректизол-процессом абсорбции сероводорода, сераорганических соединений и  $\text{CO}_2$  метанолом

Fig. 3. The training simulator for control the absorption process of hydrogen sulfide, organic sulfur compounds and  $\text{CO}_2$  by methanol

На рисунке 3 приведен пример экрана программного тренажера оператора ректизол-процесса.

Основные функции тренажера:

- управление компрессором, расходом сжиженного природного газа, температурой в абсорбере и десорбере, уровнем метанола в сборниках;
- построение графиков значений параметров процесса;
- вычисление значений параметров модели в соответствии с управлением оператора;
- контроль допустимых эксплуатационных параметров оборудования;
- создание отчетов о технологическом процессе и действиях оператора.

На рисунке 2 приведена структура сети Петри для моделирования ректизол-процесса, узлы сети соответствуют основным элементам мнемосхемы.

**Тренажер для автоматизированного испытания компрессоров** разрабатывался для Ростовского вагоноремонтного завода. Основные функции тренажера:

- вычисление значений параметров модели в соответствии с управлением оператора;
- мониторинг испытаний;
- автоматическая оценка действий обучаемого персонала на основе определенных критериев;
- построение графиков значений параметров процесса;
- управление испытаниями;
- контроль допустимых эксплуатационных параметров оборудования;
- создание отчетов о проведенных испытаниях.

На рисунке 4 приведен пример экрана программного тренажера для автоматизированного испытания компрессоров: пример мнемосхемы со значениями приборов при моделировании процесса, график выбранного параметра (давление) значения, а также информация об операторе и дополнительные параметры испытания.

В заключение отметим, что обучение ответственным технологическим операциям до того, как обучаемый оператор столкнется с ними на практике, способствует сохранению оборудования

от возможных последствий ошибок персонала и повышает качество выполняемых работ. Одним из способов обучения является использование программно-аппаратных тренажерных комплексов, обеспечивающих выработку зрительно-моторных навыков при работе на оборудовании.

При общей высокой стоимости готовых тренажерных комплексов перспективны разработка и применение специальных инструментальных средств, позволяющих специалистам предприятий самостоятельно создавать программные тренажеры с учетом специфики производственных процессов.

СПК «АСТ» позволяет использовать формализм цветных сетей Петри для моделирования ТС и обеспечивает следующее: создание статической модели ТС в виде мнемосхем, схем и чертежей; создание динамической модели ТС в виде сети Петри, совмещаемой со схемами; имитацию взаимодействия или реальное использование OPC-серверов; прогнозирование параметров ТС и построение деревьев событий. СПК «АСТ» апробирован при создании тренажеров для контроля ректизол-процесса при производстве сжиженного природного газа и послеремонтного испытания компрессоров и лазерного измерительного комплекса.

Использование СПК «АСТ» и создаваемых с его помощью тренажеров позволяет повысить уровень подготовки персонала, а значит, надежность и безопасность технологических процессов.

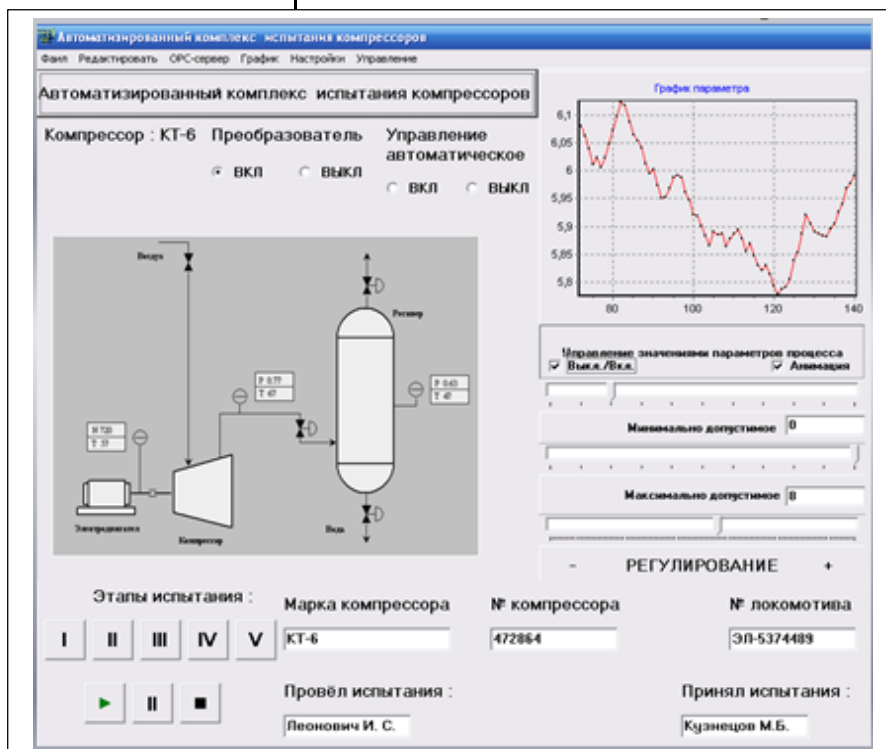


Рис. 4. Тренажер «Автоматизированный комплекс испытания компрессоров»

Fig. 4. The training simulator "Automated complex for compressor testing"



**Литература**

1. Компания SIKE. Системный интегратор, разработчик информационных и обучающих систем. URL: <http://sike.ru/articles/kompyuternyy-trenazher-sike-vyiplovka-stali-v-konvertire/> (дата обращения: 01.11.2015).
2. Берман А.Ф., Николайчук О.А., Вильвер П.Ю. Моделирование функционирования сложных технологических комплексов на основе модифицированной сети Петри // Современные технологии. Системный анализ. Моделирование. 2009. № 4. С. 23–29.
3. Дубинин В.Н., Зинкин С.А. Проектирование вычислительных систем и сетей на основе сетевых формализмов. Кн. 1: Сетевые технологии проектирования и реализации распределенных вычислительных систем на программно-аппаратных платформах локальных и глобальных сетей Ethernet/Internet. Пенза: Изд-во ПГУ, 1998. 322 с.
4. Вильвер П.Ю., Протасов А.В. Имитационное моделирование сложных динамических систем с использованием сетей Петри // Мехатроника, автоматизация, управление. 2011. № 7. С. 35–39.
5. Jensen K. Colored Petri nets – basic concepts, analysis methods and practical use. Springer-Verlag, 1997, vol. 1–3, 673 p.
6. Котов В.Е. Сети Петри. М.: Наука, 1984. 160 с.
7. Подиновский В.В., Гаврилов В.М. Оптимизация по последовательно применяемым критериям. М.: Сов. радио, 1975. 192 с.
8. Орлов А.И. Теория принятия решений: учеб. пособие. М.: Март, 2004. 656 с.
9. Малтугуева Г.С., Юрин А.Ю. Алгоритм коллективного выбора на основе обобщенных ранжировок для поддержки принятия решений // Современные технологии. Системный анализ. Моделирование. 2009. № 3. С. 57–62.

DOI: 10.15827/0236-235X.115.136-141

Received 16.11.15

**“AST” SOFTWARE SUITE FOR CREATING TRAINING SIMULATORS**<sup>1</sup>Wilwer P.Yu., Programmer, wilwer@icc.ru<sup>1</sup>Yurin A.Yu., Ph.D. (Engineering), Associate Professor<sup>1</sup>Institute of System Dynamics and Control Theory SB RAS, Lermontov St. 134, Irkutsk, 664033, Russian Federation

**Abstract.** The paper describes the “AST” software (Automated creation of training simulators) designed for creation of computer training simulators with the decision support functions. The software is based on the colored Petri nets formalism that is used for modeling technical systems (TS).

The software provides the following: creation of a hierarchical static model of technical systems in the form of mnemonic diagrams and drawings using a built-in graphic editor or AutoCad; creation of a dynamic model of technical systems in the form of a Petri net; visual modeling (simulation) of static and dynamic properties of technical systems as a result of interaction (or interaction simulation) with an OPC-Server; forecasting parameters of technical systems and building event trees.

The paper describes the software functions, its architecture and methodology of creating simulator complexes. There are examples of application of the developed software for creating complex software simulators for training specialists of post-repairing testing of compressors and rektizol-process control when producing liquefied natural gas.

**Keywords:** colored Petri nets, software training simulator, computer-aided, heat mechanical, chemical and technological systems, OPC-server, training.

**References**

1. SIKE Company. The system integrator, developer of information and systems. Available at: <http://sike.ru/articles/kompyuternyy-trenazher-sike-vyiplovka-stali-v-konvertire/> (accessed November 3, 2015).
2. Berman A.F., Nikolaychuk O.A., Vilver P.Yu. Modeling of complex technological system (complex) functioning on the basis of modified Petri net. *Sovremennyye tekhnologii. Sistemy analiz. Modelirovaniye* [Modern technologies. System analysis. Modeling]. 2009, no. 4, pp. 23–29 (in Russ.).
3. Dubinin V.N., Zinkin S.A. *Proektirovaniye vychislitelnykh sistem i setey na osnove setevykh formalizmov. Kn. 1: Setevye tekhnologii proektirovaniya i realizatsii raspredelennykh vychislitelnykh sistem na programmno-apparatnykh platformakh lokalnykh i globalnykh setey Ethernet/Internet* [Design of Computer Systems and Networks on the Basis of Net Formalisms. Book 1: Network Technology Design and Implementation of Distributed Computing Systems Hardware and Software Platforms LAN and WAN Ethernet/Internet]. Penza, PSU Publ., 1998, 322 p. (in Russ.).
4. Vilver P.Yu., Protasov A.V. Imitating Modeling of Complex Dynamic Systems Using Petri Nets. *Mekhatronika, avtomatizatsiya, upravleniye* [Mechatronics, Automation, Control]. 2011, no. 7, pp. 35–39 (in Russ.).
5. Jensen K. *Colored Petri Nets – Basic Concepts, Analysis Methods and Practical Use*. Springer-Verlag, 1997, vol. 1–3, 673 p.
6. Kotov V.E. *Seti Petri* [Petri Networks]. Moscow, Nauka Publ., 1984, 160 p. (in Russ.).
7. Podinovskiy V.V., Gavrilov V.M. *Optimizatsiya po posledovatelno primenyaemym kriteriyam* [Optimization on Consistently Applied Criteria]. Moscow, Sovetskoe radio, 1975, 192 p. (in Russ.).
8. Orlov A.I. *Teoriya prinyatiya resheniy* [Decision-Making Theory]. Study guide for univ. Moscow, Mart Publ., 2004, 656 p. (in Russ.).
9. Maltugueva G.S., Yurin A.Yu. Collective choice algorithm on the basis of generalized rankings for decision support. *Sovremennyye tekhnologii. Sistemy analiz. Modelirovaniye* [Modern Technologies. System Analysis. Modeling]. 2009, no. 3, pp. 57–62 (in Russ.).

УДК 004.94

Дата подачи статьи: 30.12.15

DOI: 10.15827/0236-235X.115.142-148

## **ЭТАПЫ РАЗРАБОТКИ ИНТЕГРИРОВАННОЙ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОЙ СИСТЕМЫ ОБРАБОТКИ ЭКОНОМИЧЕСКОЙ ИНФОРМАЦИИ В ОРГАНИЗАЦИЯХ**

*М.А. Беляева, д.т.н., профессор, belyaevamar@mail.ru;*

*О.К. Безотосова, магистрант*

*(Российский экономический университет им. Г.В. Плеханова,  
Стремянный пер., 36, г. Москва, 117997, Россия)*

В статье представлена интегрированная информационно-аналитическая система обработки экономической информации, разработанная с использованием современных механизмов интеграции и программных средств, показан импорт файла данных из офисного приложения MS Excel в имитационную систему AnyLogic. Отражены механизмы интеграции обработки данных, а также этапы создания интегрированной информационно-аналитической системы. В MS Excel проводится расчет экономических показателей, а в имитационной системе те же экономические показатели рассчитываются и одновременно графически отображаются их динамические изменения. На основании результатов можно проводить анализ финансово-хозяйственной деятельности предприятия.

Система включает в себя подсистемы структурно-параметрического моделирования, подсистему расчетов экономических показателей в MS Excel с помощью макросов и кнопочных форм управления в Visual Basic For Application, подсистему динамического моделирования в имитационной системе AnyLogic и подсистему 1С. Предприятие.

Разработанная интегрированная информационно-аналитическая система может оказать руководителям и менеджерам компаний информационную поддержку для принятия оптимальных управленческих решений.

Система универсальная и может найти широкое применение.

**Ключевые слова:** механизмы интеграции, системы поддержки принятия решений, экономические показатели, имитационная система, факторный анализ, информационно-аналитическая система.

Современные информационные технологии и программные продукты, инструментальные средства и механизмы позволяют разрабатывать экономико-математические модели для контроля и регулирования ключевых показателей финансово-хозяйственной деятельности в организациях для принятия на всех уровнях управления оптимальных решений, что способствует интенсификации обработки огромного количества поступающей информации.

Рассматривая организацию с функциональных позиций и уровней управления, можно представить ее в виде информационной пирамиды, отражающей стратегический уровень – его деятельность обеспечивает руководитель предприятия или главный менеджер, функционально-тактический – менеджеры среднего звена, уровень знаний – сотрудники, владеющие знаниями, и операционный уровень – исполнители, и проводить интеграцию, начиная с автоматизации рабочего места (АРМ) сотрудника, локальных информационных систем, внедряемых в различных подразделениях, систем стратегического назначения, тем самым создавая корпоративные информационные системы.

Внедряемые на предприятиях информационные системы как отечественных, так и зарубежных разработчиков можно классифицировать по функциональному признаку: производственно-технические системы, бухгалтерские и финансовые, маркетинговые, системы управления персоналом.

На российском рынке пользуются спросом огромное количество информационных систем и

программных продуктов, которые содержат достаточно эффективные модули (компоненты) по управлению производственными процессами и материальными потоками, такие как SAP, Oracle, IFS, Axapta, Navision, SyteLine, «1С: Предприятие 8.0. Управление производственным предприятием». Рассмотрим некоторые из них.

Oracle Financial Services Applications (OFSA) – преднастроенная интегрированная система поддержки принятия решений, улучшающая процессы бюджетирования, планирования, управления рисками, взаимоотношения с клиентами и измерения эффективности в финансовых организациях различного профиля.

Oracle Financial Data Manager (FDM) – ключевая компонента OFSA, обеспечивающая создание и ведение единого общекорпоративного хранилища финансово-аналитических данных. Как централизованное хранилище данных FDM реализует модель данных, ориентированную на отображение широкого спектра информации (доходы, затраты, бюджеты, многомерная рентабельность, движение денежной наличности) по счетам, клиентам и продуктам. Реализованная в FDM модель данных отражает опыт, накопленный в процессе реализации финансово-аналитических систем в крупнейших банках и страховых компаниях, и решает проблему «разрастания» моделей данных и фрагментации данных по различным информационным системам. Она позволяет извлекать, собирать и поддерживать детальную информацию по счетам из любой транзакционной и финансовой системы, функционирующей в финансовой организации.

Использование системы Oracle Financial Services Applications неактуально в компании, так как система обладает излишней функциональностью, что значительно повышает стоимость внедрения и сопровождения. В связи с этим для развития и совершенствования деятельности организации целесообразнее внедрить систему более узкого направления с возможностью различной детализации в соответствующей области.

Система поддержки принятия решений «Эксперт» предназначена для решения слабоструктурированных и неструктурированных задач на основе компьютерного анализа экспертных суждений. Система представляет собой программный модуль, функционирующий как в среде ГИС, так и независимо в качестве отдельного приложения. Особенности системы: базируется на современных методах поддержки принятия решений, применявшихся в США, Мексике, Канаде и в других странах для задач аналитического планирования; поддерживает работу как с числовыми значениями, так и субъективных вербальных предпочтений пользователя; обеспечивает возможность анализа данных на предмет согласованности и достоверности; позволяет исправлять несогласованности; имеет возможность обработки любых внешних данных и совместных суждений, достижения консенсуса; предусматривает подробные печатные отчеты и низкие системные требования и пр.

Методология, лежащая в основе системы «Эксперт», моделирует ход мышления человека и основывается на представлении всех элементов, факторов и т.д., влияющих на принятие решения, в виде иерархии (производится декомпозиция проблемы, то есть разделение ее на составные части). При этом происходят группировка отдельных элементов в пределах уровней и разделение уровней по сложности. Иерархия строится с вершины (цель решаемой задачи), через промежуточные уровни (критерии принятия решений) к последнему уровню (альтернативы, варианты, сценарии). Элементы каждого уровня являются непосредственными критериями для элементов следующего.

Описанная система представляет собой сложный программный комплекс, нацеленный на анализ слабоструктурированных данных, поэтому такие системы поддержки принятия решений не востребованы. Наоборот, руководителю предприятия удобно графическое представление данных, которое является важным для принятия решений менеджерами, в системе «Эксперт» этому уделено мало внимания.

Система поддержки принятия решений Crystal Info (Seagate Info) 7.5 основана на гибкой технологии доступа к данным и их обработки, обеспечивает стандартизацию документооборота и отчетности в масштабе предприятия. Поддерживается технология OLAP, что позволяет быстро получать ответы на важные для бизнеса вопросы. В Crystal

Info имеются средства для просмотра отчетов и данных, представленных в виде многомерных массивов, через web-браузеры. Это делает Crystal Info практически независимой от используемых серверных платформ. Crystal Info имеет возможность распространения отчетов, в том числе и по Internet.

Отчеты могут выводиться на web-сервер или рассылаться по электронной почте тем, для кого они предназначены. Составление и рассылка отчетов могут производиться по заранее составленному расписанию или инициироваться теми или иными внешними событиями. Crystal Info поддерживает все основные источники данных, многомерные данные, может работать со специфичными для программных продуктов Microsoft данными, такими как Exchange или Outlook.

Crystal Info интегрирована с системой программной архивации и резервного копирования Crystal Backup Exec, что обеспечивает возможность быстрого восстановления данных. Пользователи могут устанавливать опции резервного копирования как глобально для всей системы, так и для отдельных объектов.

Возможности работы в Internet/Intranet: вывод отчетов в формате HTML с возможностью просмотра в браузере; обновление выводимых на веб-сайт отчетов по расписанию; вывод отчетов по статистике обращений к web-сайту.

Crystal Info WebAccess обеспечивает доступ ко всем возможностям анализа данных и просмотра отчетов в безопасной среде web через любой браузер.

Для внедрения данной системы необходимо дополнительное оборудование, что влечет дополнительные затраты. Потребуется осуществлять обмен данными между Sap и Crystal, что приводит к снижению скорости предоставления данных менеджерам.

SAP Business Information Warehouse (SAP BW) предоставляет функции для обработки данных, интеграционную платформу и набор компонентов, позволяющих бизнесу добиваться поставленных целей, таких как построение отчетности и анализ данных. Данные из SAP, а также из внешних источников могут интегрироваться, трансформироваться и консолидироваться в SAP BW, которая, имея гибкие механизмы для построения отчетности и ее анализа, позволяет менеджерам принимать продуманные и хорошо обоснованные решения и определять цели, ориентированные на достижение максимального результата.

SAP BW – это полностью готовое информационное хранилище, которое начинает работать сразу, а не после многомесячных экспериментов. Требование поставки экономического содержания означает привнесение порядка в хранилище данных, ориентированного на нужды пользователей с точки зрения получения необходимой отчетности и поддержки принятия решений. Данные выгру-

жаются в соответствии с определенными, согласованными с бизнес-процессами и требованиями клиентов схемами, собираются и логически объединяются друг с другом. Для этой цели в SAP BW имеется большое количество предопределенных экстракторов, информационных моделей и моделей анализа, показателей и отчетов. Эти готовые отчеты предназначены для определенных групп пользователей.

Администрирование информационного хранилища SAP BW выполняется очень просто – оно соединяется с любым SAP-окружением. Administrator Workbench (инструментарий администратора) гарантирует удобное и эффективное управление информационным хранилищем. Он поддерживает перенос, интеграцию и управление внешними данными, а также разработку различных расширений, обеспечивает администратору хранилища возможность полного контроля над экстрактами данных, агрегатами, отчетами и многое другое. Работа осуществляется с помощью различных несложных функций типа Point&Click.

Документы и метаданные, созданные в SAP BW, могут быть интегрированы при помощи BW Metadata Repository Manager в Knowledge Management SAP Enterprise Portal. Обмен SAP и других источников данных с SAP BW может осуществляться при помощи SAP Exchange Infrastructure (SAP XI). В SAP BI Content предлагаются предварительно настроенные роли, информационные объекты и модели, готовые запросы и отчеты, которые покрывают практически все бизнес-направления и SAP-приложения.

SAP Business Explorer Analyzer предназначен для создания и просмотра аналитических отчетов в информационном хранилище SAP BW.

БЕх является компонентом SAP BW, предоставляющим гибкую систему отчетов и методы анализа, которые можно использовать для стратегического анализа и поддержки процесса принятия решений на предприятии. Сотрудники, имеющие соответствующие полномочия, могут анализировать исторические и актуальные данные на различных уровнях детализации и с различных точек зрения. Для представления данных они могут использовать web или Microsoft Excel.

Среди недостатков в вышеперечисленных программных продуктах отметим высокую стоимость и возможность их внедрения только на крупных предприятиях, а предлагаемая авторами система проста, она может быть внедрена как на малых, так и на больших предприятиях.

Программный продукт AnyLogic компания-разработчик предлагает использовать в течение года бесплатно,

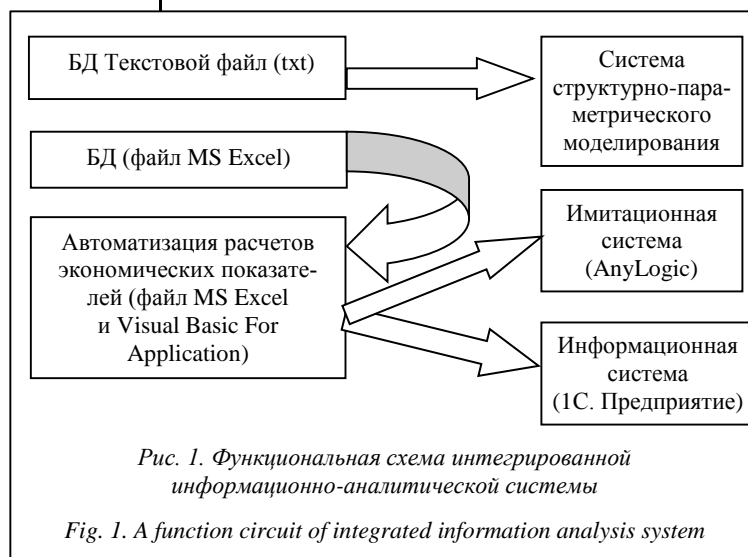
со второго года пользователи оплачивают только консультации по возникающим при эксплуатации моделей вопросам.

Современные программные средства позволяют интегрировать или транспортировать из одной информационной системы или программы в другую. Для этого имеются специальные команды импорта и экспорта, внешние процедуры ввода и вывода, транспортирование данных, обработку которых можно осуществить с использованием различных математических методов: корреляционный и регрессионный анализ, прогнозирование и диагностирование, методы оптимизации, многофакторный анализ. Поэтому на сегодняшний день актуально разрабатывать интегрированные информационные системы, пользователями которых являются руководители и менеджеры среднего звена, на всех уровнях пирамиды предприятия, можно проводить как вертикальную, так и горизонтальную интеграцию [1–5].

Разработанная система включает в себя подсистемы структурно-параметрического моделирования, подсистему расчетов экономических показателей в MS Excel с помощью макросов и кнопочных форм управления в Visual Basic For Application, подсистему динамического моделирования в имитационной системе AnyLogic и подсистему 1C. Предприятие.

Функциональная схема представлена на рисунке 1.

Функционирование системы заключается в том, что экономические данные можно создать в текстовом формате и транспортировать их в любую из подсистем, осуществлять обработку, используя методы структурно-параметрического анализа, и автоматизацию расчетов экономических показателей средствами MS Excel и Visual Basic For Application. Для визуализации вывода результатов вычислений на экран были использованы и элементы MSFlexGrid, что дает возможность проводить фак-



торный анализ, обрабатывать данные в имитационных системах и осуществлять мониторинг и контроль динамического состояния (рис. 1).

Разработанную интеграцию можно расширять, что дает дальнейшие перспективы развития системы, а для руководителя информационную поддержку, интенсифицирующую обработку рутинной информации.

К основным экономическим показателям финансово-хозяйственной деятельности относятся прибыль, среднемесячная выручка, доля денежных средств в выручке и среднемесячная численность работников, показатели платежеспособности и финансовой устойчивости, показатели эффективности использования оборотного капитала, рентабельность, коэффициент задолженности другим организациям, степень платежеспособности по текущим обязательствам и пр.

На рисунках по ссылкам [http://www.swsys.ru/uploaded/image/2016\\_3/2016-3-dop/1.jpg](http://www.swsys.ru/uploaded/image/2016_3/2016-3-dop/1.jpg), [http://www.swsys.ru/uploaded/image/2016\\_3/2016-3-dop/2.jpg](http://www.swsys.ru/uploaded/image/2016_3/2016-3-dop/2.jpg) можно увидеть ввод исходных данных и вывод результатов средствами MS Excel объектно-ориентированного языка программирования VBA.

Деятельность и устойчивое функционирование организации характеризуются получением максимальной прибыли. Разработанная информационная система предназначена для руководителя, чтобы на основе анализа результатов расчета экономических

показателей финансово-хозяйственной деятельности принимать оптимальные управленческие решения [5–7].

Для проведения факторного анализа прибыли организации необходимы исходные данные за два периода в рублях для сравнения показателей прироста в рублях, удельного веса в выручке в % по каждому периоду, прироста удельного веса в %.

Исходные данные для расчета и полученные результаты сведены в таблицу 1.

Для выполнения анализа прибыли организации необходимо задать значение инфляционного роста цен и тарифов (в %) [8].

Визуально все рассчитанные факторы для данной организации можно подразделить на две группы: доходы, увеличивающие прибыль, расходы и налоги, уменьшающие прибыль (с обратным знаком).

Одновременно с этим выводятся влияние инфляции на суммарный объем выручки в абсолютном выражении (в рублях) и влияние выпуска продукции на суммарный объем выручки в рублях (см. [http://www.swsys.ru/uploaded/image/2016\\_3/2016-3-dop/1.jpg](http://www.swsys.ru/uploaded/image/2016_3/2016-3-dop/1.jpg), [http://www.swsys.ru/uploaded/image/2016\\_3/2016-3-dop/2.jpg](http://www.swsys.ru/uploaded/image/2016_3/2016-3-dop/2.jpg), [http://www.swsys.ru/uploaded/image/2016\\_3/2016-3-dop/3.jpg](http://www.swsys.ru/uploaded/image/2016_3/2016-3-dop/3.jpg), [http://www.swsys.ru/uploaded/image/2016\\_3/2016-3-dop/4.jpg](http://www.swsys.ru/uploaded/image/2016_3/2016-3-dop/4.jpg)). Из рассчитанных показателей (табл. 1) видно, что падение прибыли в отчетном году в организации

Таблица 1

## Анализ прибыли за отчетный период

Table 1

## Profit analysis in the period under review

Показатель	Обозначение	Значение показателя, руб.		Прирост показателя, руб.	Удельный вес показателя в выручке, %		Прирост удельного веса, %
		01.01.2013	01.01.2014		01.01.2013	01.01.2014	
Выручка без НДС	R	21847064	21898116	-51052	100	100	0
Затраты	TC	15637658	15562990	74668	71,58	71,07	0,51
В том числе:							
себестоимость	VC	11812503	12632148	-819645	54,07	57,69	-3,62
коммерческие расходы	FC <sub>к</sub>	1856223	1096321	759902	8,5	5,01	3,49
управленческие расходы	FC <sub>у</sub>	1968932	1834521	134411	9,01	8,37	0,64
Прибыль от реализации	P <sub>т</sub>	6209406	6335126	-125720	28,42	28,93	-0,51
Операционные доходы	Do	409318	236687	172631	1,87	1,08	0,79
Операционные расходы	Ro	521625	178133	343492	2,39	0,81	1,57
Внереализационные расходы	Rv	231911	92346	139565	1,06	0,42	0,64
Внереализационные доходы	Dv	322637	239397	83240	1,48	1,09	0,38
Налоги за счет финансовых результатов	T <sub>х</sub>	684354	507104	177250	3,13	2,32	0,82
Балансовая прибыль	P <sub>гб</sub>	5503471	6033627	-530156	25,19	27,55	-2,36

вызвано несколькими факторами, а именно: объемом продаж, полными затратами, коммерческими, управленческими, операционными, внебюджетными расходами и налогами в счет финансовых результатов.

При разработке динамических расчетов экономических показателей была использована имитационная система AnyLogic Professional, поддерживающая все подходы к созданию моделей.

Для расчета экономических показателей в AnyLogic использован метод Хольта, у которого есть преимущество в гибкости, позволяющей выбирать соотношение, отслеживающее и уровень, и наклон. При этом в методе Хольта оба значения сглаживаются непосредственно, но с использованием разных постоянных сглаживания для каждого из них. Помимо двойного сглаживания, существуют также модели тройного экспоненциального сглаживания. Данный способ используется при наличии тренда и сезонности. Для управления сезонностью необходимо ввести третий параметр. Таким образом, вводится третье уравнение в модель, которая известна как модель Хольта–Уинтерса. Различают две главные ее разновидности в зависимости от типа сезонности: модели с аддитивной и с мультипликативной сезонностью. В таблице 2 представлены модели линейного тренда Хольта–Уинтерса.

Для реализации имитационной модели был использован системно-динамический подход, позволяющий понять структуру и динамику сложных систем, инструменты которого дают возможность создавать микромиры-симуляторы, где пространство и время могут быть заданы так, что можно спрогнозировать последствия решений, быстро освоить ме-

тоды, спроектировать тактику и стратегию для большего успеха деятельности организации.

Для создания системы был использован механизм интеграции: так, элемент файл MS Excel предоставляет легкий доступ к файлам с расширением .xls, .xlsx из моделей AnyLogic.

С помощью этого объекта можно читать содержимое файлов MS Excel, считывать значения и формулы из определенных ячеек рабочих книг, считывать данные в табличные функции AnyLogic, считывать данные в одномерные и двумерные массивы AnyLogic, создавать новые ячейки; записывать значения в определенные ячейки и др.

Отметим, что на разработанную интегрированную информационно-аналитическую систему получено Свидетельство о государственной регистрации программ для ЭВМ № 2016614666 «Универсальная интегрированная система расчета экономических показателей предприятия с использованием современных программных средств».

Этапы создания интегрированной информационной системы.

1. Задать файл, из которого будут импортировать данные, для этого из палитры на рабочую область модели выносятся ExcelFile, затем указывается путь.

2. Создать параметры, в которых будут храниться данные из MS Excel (рис. 2; в левой части рисунка – входящие данные, в правой – параметры для прогноза).

Так как входящие параметры содержат восемь кварталов, каждая переменная, обозначающая входные данные, будет записана как массив с размерностью восемь, а прогноз будет – как массив с размерностью четыре, так как прогнозирование

Таблица 2

Модели линейного тренда Хольта–Уинтерса

Table 2

The models of Holt–Winters linear trend

Модель	Рекуррентная форма	Скорректированная на ошибку форма
Несезонная (2-1) (модель Хольта)	$S_t = \alpha X_t + (1 - \alpha)(S_{t-1} + T_{t-1})$ $T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)T_{t-1}$ $\hat{X}_{t(m)} = S_t + mT_t$	$S_t = S_{t-1} + T_{t-1} + \alpha\epsilon_t$ $T_t = T_{t-1} + \gamma\alpha\epsilon_t$ $\hat{X}_{t(m)} = S_t + mT_t$
С аддитивной сезонностью (2-2)	$S_t = \alpha(X_t - I_{t-p}) + (1 - \alpha)(S_{t-1} + T_{t-1})$ $T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)T_{t-1}$ $I_t = \delta(X_t - S_t) + (1 - \delta)I_{t-p}$ $\hat{X}_{t(m)} = S_t + mT_t + I_{t-p+m}$	$S_t = S_{t-1} + T_{t-1} + \alpha\epsilon_t$ $T_t = T_{t-1} + \gamma\alpha\epsilon_t$ $I_t = I_{t-p} + \delta(1 - \alpha)\epsilon_t$ $\hat{X}_{t(m)} = S_t + mT_t + I_{t-p+m}$
С мультипликативной сезонностью (2-3)	$S_t = \alpha(X_t/I_{t-p}) + (1 - \alpha)(S_{t-1} + T_{t-1})$ $T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)T_{t-1}$ $I_t = \delta(X_t/S_t) + (1 - \delta)I_{t-p}$ $\hat{X}_{t(m)} = (S_t + mT_t)I_{t-p+m}$	$S_t = S_{t-1} + T_{t-1} + \alpha\epsilon_t/I_{t-p}$ $T_t = T_{t-1} + \gamma\alpha\epsilon_t/I_{t-p}$ $I_t = I_{t-p} + \delta(1 - \alpha)\epsilon_t/S_t$ $\hat{X}_{t(m)} = (S_t + mT_t)I_{t-p+m}$

Примечание:  $\alpha$  – сглаживающий параметр для уровня ряда;  $\gamma$  – сглаживающий параметр для тренда;  $\delta$  – сглаживающий параметр для сезонных факторов;  $S_t$  – сглаженный уровень ряда, подсчитанный после наблюдаемого  $X_t$ , а также ожидаемая оценка данных в конце периода  $t$  в некоторых моделях;  $T_t$  – сглаженный тренд в конце периода  $t$ ;  $I_t$  – сглаженный сезонный индекс или фактор в конце периода  $t$ ;  $X_t$  – наблюдаемое значение временного ряда в периоде  $t$ ;  $m$  – число периодов в прогнозируемом процессе освоения;  $p$  – число периодов в сезонном цикле;  $\hat{X}_{t(m)}$  – прогноз на  $m$  периодов вперед от первоначального  $t$ ;  $\epsilon_t$  – прогнозная ошибка на шаг вперед.

✓ валовая_выручка	✓ валовая_выручка 1
✓ среднемесячная_численность_чел	✓ среднемесячная_численность_чел 1
✓ долгосрочные_обязательства	✓ долгосрочные_обязательства 1
✓ займы_и_кредиты	✓ займы_и_кредиты 1
✓ поставщики_и_подрядчики	✓ поставщики_и_подрядчики 1
✓ задолженность_перед_персоналом	✓ задолженность_перед_персоналом 1
✓ задолженность_перед_фондами	✓ задолженность_перед_фондами 1
✓ задолженность_по_налогам_и_сборам	✓ задолженность_по_налогам_и_сборам 1
✓ прочие_кредиторы	✓ прочие_кредиторы 1
✓ задолженность_перед_учредителями_по_выплате_доходов	✓ задолженность_перед_учредителями_по_выплате_доходов 1
✓ доходы_будущих_периодов	✓ доходы_будущих_периодов 1
✓ резервы_предстоящих_расходов	✓ резервы_предстоящих_расходов 1
✓ прочие_краткосрочные_обязательства	✓ прочие_краткосрочные_обязательства 1
✓ оборотные_активы	✓ оборотные_активы 1
✓ внеоборотные_активы	✓ внеоборотные_активы 1
✓ капитал_и_резервы	✓ капитал_и_резервы 1
✓ запасы	✓ запасы 1
✓ товары_отгруженные	✓ товары_отгруженные 1
✓ чистая_прибыль	✓ чистая_прибыль 1
✓ прибыль	✓ прибыль 1
✓ выручка_нетто	✓ выручка_нетто 1
✓ незавершенное_строительство	✓ незавершенное_строительство 1
✓ долгосрочные_финансовые_вложения	✓ долгосрочные_финансовые_вложения 1

Рис. 2. Экранная форма списка входных параметров (программная среда AnyLogic)

Fig. 2. A display form of the input parameter list (software environment AnyLogic)

данных осуществляется по четырем кварталам.

3. Задать поля редактирования данных, в которых также будут обновляться соответствующие графические зависимости, с их помощью выводятся конкретный набор значений на экран.

По графическим зависимостям можно выполнить анализ результатов прибыли и выручки по годам, кварталам, провести сравнительный анализ.

На рисунке 2 представлен перечень переменных для их расчета – список входных параметров.

4. Создать функции. Данные считываются с файла с помощью функции `getCellNumericValue` (лист, строка, столбец).

`data.валовая_выручка[i]=excelFile.getCellNumericValue(3, 12, 13+i)` инициализации данных написана функция прогноза (в агенте Data), прогноз строится методом Винтерса–Хольта, для каждой динамической переменной написаны расчетные формулы. (Некоторые из них см. [http://www.swsys.ru/uploaded/image/2016\\_3/2016-3-dop/6.jpg](http://www.swsys.ru/uploaded/image/2016_3/2016-3-dop/6.jpg).)

Таким образом, при наличии руководства пользователя система может быть внедрена в любой организации, так как является универсальной, и рекомендована руководителям и менеджерам в качестве компьютерной поддержки принятия опти-

мальных управленческих решений на основе анализа экономических показателей финансово-хозяйственной деятельности.

### Литература

1. Шеер А.-В. Управляя менеджерами. М.: Альпина бизнес-букс, 2012. 224 с.
2. Беляева М.А. Моделирование систем. М.: Изд-во МГУП, 2012. Ч. 1. 188 с.
3. Беляева М.А. Имитационное моделирование социально-экономических систем для поддержки принятия решений // Пищевая промышленность, 2011. № 4. С. 86–87.
4. Беляева М.А., Бурляева О.К., Сырова В.И. Формирование мультимодельной системы для принятия оптимальных управленческих решений на предприятии // Программные продукты и системы. 2014. № 2 (106). С. 181–187.
5. Савицкая Г.В. Анализ финансово-хозяйственной деятельности. М.: Инфра-М, 2009. 536 с.
6. Шеремет А.Д., Сайфулин Р.С., Негашев Е.В. Методика финансового анализа. М.: ИНФРА-М, 1999. 208 с.
7. Уткин В.Б., Балдин К.В. Информационные системы и технологии в экономике. М.: Изд-во ЮНИТИ, 2003. 395 с.
8. Хаммер М., Хершман Л. Быстрее, лучше, дешевле. Десять методов реинжиниринга бизнес-процессов. М.: Альпина Паблишер, 2012. 356 с.
9. Тельнов Ю.Ф. Интеллектуальные информационные системы в экономике. М.: СИНТЕГ, 1999. 214 с.
10. Бажин Н.Н. Информационные системы менеджмента. М.: Изд-во ВШЭ, 2000. 688 с.

DOI: 10.15827/0236-235X.115.142-148

Received 30.12.15

**DEVELOPMENT STAGES OF INTEGRATED INFORMATION ANALYSIS SYSTEM  
OF PROCESSING ECONOMIC INFORMATION IN ORGANIZATIONS**<sup>1</sup>*Belyaeva M.A., Dr.Sc. (Engineering), Professor, belyaevamar@mail.ru*<sup>1</sup>*Bezotosova O.K., Undergraduate*<sup>1</sup>*Plekhanov Russian University of Economics, Stremyanny Line 36, Moscow, 117997, Russian Federation*

**Abstract.** The article presents an integrated information analysis system of processing economic information, which is developed using modern integration mechanisms and software. It also shows data file import from an office application such as MS Excel into a simulation system AnyLogic.

The article describes the mechanisms of data processing integration and the stages of integrated information-analytical system creation. MS Excel calculates economic indicators. The simulation system calculates the same economic indicators and simultaneously displays dynamic graphic changes of these indicators. On the basis of the results it is possible to analyse financial and economic activities of an enterprise.

The system includes a structure-parametric modeling subsystem, a subsystem of economic indicators calculations in MS Excel with macroses and button control forms in Visual Basic For Application, a dynamic modeling subsystem in AnyLogic simulation system and subsystem 1C.Enterprise.

The developed integrated information analysis system can be an information decision-making support for heads and managers in company activities. The system is universal as it can be implemented in any company.

**Keywords:** integration mechanisms, decision-making support system, economic indicators, simulation system, factorial analysis, information and analytical system.

**References**

1. Scheer A.-W. *Upravlyaya menedzherami* [Managing Managers]. Moscow, Alpina Business Books Publ., 2012, 224 p.
2. Belyaeva M.A. *Modelirovanie sistem* [System Modeling]. Moscow, MGUP Publ., 2012, part 1, 188 p.
3. Belyaeva M.A. Simulation modeling of socio-economic systems for decision support. *Pishchevaya promyshlennost* [Food Industry]. 2011, no. 4, pp. 86–87 (in Russ.).
4. Belyaeva M.A., Burlyayeva O.K., Syrova V.I. Forming a multi-model system to make better management decisions at an enterprise. *Programmnye produkty i sistemy* [Software and Systems]. 2014, no. 2 (106), pp 181–187.
5. Savitskaya G.V. *Analiz finansovo-khozyaystvennoy deyatel'nosti* [Analysis of Financial and Economic Activity]. Moscow, INFRA-M Publ., 2009, 5th ed., 536 p.
6. Sheremet A.D., Sayfulin R.S. *Metodika finansovogo analiza* [Methodology of financial analysis]. Moscow, INFRA-M Publ., 1996.
7. Utkin V.B., Baldin K.V. *Informatsionnye sistemy i tekhnologii v ekonomike* [Information Systems and Technologies in Economics]. Moscow, YUNITI Publ., 2003.
8. Hammer M., Hershman L.W. *Faster Cheaper Better: The 9 Levers for Transforming How Work Gets Done*. Crown Publ., 320 p. (Russ.ed.: Moscow, Alpina Publ., 2012, 356 p.).
9. Telnov Yu.F. *Intellektualnye informatsionnye sistemy v ekonomike* [Intelligent Information Systems in Economics]. Moscow, SINTEG Publ., 1999, 214 p.
10. Bazhin N.N. *Informatsionnye sistemy menedzhmenta* [Information Management Systems]. Moscow, HSE Publ., 2000, 688 p.

---

---

**НИИ «Центрпрограммсистем»  
на Международном военно-техническом форуме «АРМИЯ-2016»**

НИИ «Центрпрограммсистем» принимал участие в Международном военно-техническом форуме «АРМИЯ-2016» 6–11 сентября в г. Кубинка Московской области.

В своей экспозиции институт представил наиболее перспективные новейшие разработки в области информационных и автоматизированных систем, тренажеростроения и защиты информации.

*Контакты для партнеров и заказчиков:*

*Телефон: (4822) 44-32-44*

*Официальный сайт: [www.cps.tver.ru](http://www.cps.tver.ru)*

*Электронная почта: [info@cpr.tver.ru](mailto:info@cpr.tver.ru)*



## **АСПЕКТНО-ОРИЕНТИРОВАННОЕ ПРОГРАММИРОВАНИЕ В КОНТЕКСТЕ РЕШЕНИЯ ВОПРОСОВ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ ИТ-ПРОЕКТОВ**

*А.А. Гончаров, аспирант, andre\_goncharov@hotmail.com;*

*Н.А. Семенов, д.т.н., профессор, is@tstu.tver.ru  
(Тверской государственный технический университет,  
наб. Аф. Никитина, 22, г. Тверь, 170026, Россия)*

Оценка бюджета разрабатываемого проекта, классификация расходов, планирование сроков разработки и анализ затрат сопровождаются неточностями и искажениями в силу того, что исходный код промышленных информационных систем характеризуется необоснованной сложностью, излишним дублированием, шаблонными ошибками, неопределенностью, наличием сильной взаимосвязи между модулями.

Статья посвящена способу решения актуальной задачи повышения эффективности экономических показателей ИТ-проектов. В ней описан современный подход к проектированию крупных информационных систем с использованием аспектно-ориентированного программирования.

Аспектно-ориентированное программирование является современным развитием объектно-ориентированного программирования и предназначено для отделения бизнес-логики информационных систем от сквозных функций. Бизнес-логика описывает реализацию бизнес-правил для конкретного промышленного предприятия.

В статье подчеркнута актуальность использования аспектно-ориентированного программирования в крупных промышленных информационных системах, а также описаны его преимущества и недостатки с точки зрения наибольшей практической ценности. Предложен механизм, позволяющий абстрагировать бизнес-процессы предприятия на всех итерациях стадий проектирования и реализации ПО от несущественных (вспомогательных) аспектов системы.

Кроме этого, в статье приводится практический пример использования аспектно-ориентированного программирования.

Абстрагирование совокупности задач в аспекты позволяет получать корректное отображение связей между бизнес-процессами и требованиями к реализации объектов предметной области проектируемой информационной системы.

**Ключевые слова:** *аспектно-ориентированное программирование, задача, информационная система, проектирование, программирование, эффективность.*

Тенденция к проникновению *информационных технологий* (ИТ) в подавляющее большинство областей деятельности человека, особенно в промышленность, развитие средств массовой коммуникации и масштабность бизнеса находят свое отражение в сложности принимаемых решений и неопределенности представления предметной области при разработке ПО.

Общепризнано, что предметная область является инвариантной составляющей в современном рекуррентном подходе к проектированию промышленного ПО, который характеризуется компоновкой различных взглядов в отношении реализации функциональности бизнес-процессов будущей *информационной системы* (ИС).

Таким образом, необходим механизм, позволяющий абстрагировать каждый бизнес-процесс предприятия на всех итерациях стадий проектирования и реализации ПО от несущественных (вспомогательных) аспектов системы.

### **Методологии программирования**

Среди основных методологий программирования (императивного, функционального, логического программирования и *объектно-ориентированного программирования* (ООП)) наиболее востребованным подходом при проектировании и

разработке ПО является ООП (Object-Oriented Programming) [1].

ООП – методология программирования, основанная на представлении программного продукта в виде совокупности объектов, каждый из которых является экземпляром конкретного класса. ООП использует в качестве базовых элементов взаимодействие объектов – именованных моделей реальной сущности, обладающих конкретными значениями свойств и проявляющих свое поведение [2].

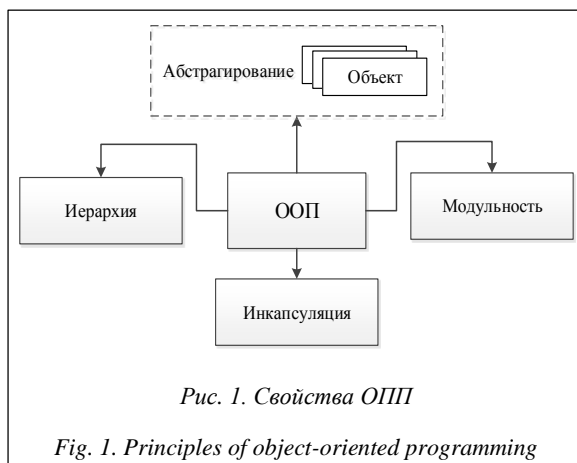
ООП можно описать четырьмя главными характеристическими свойствами: абстрагирование, инкапсуляция, модульность, иерархия (рис. 1).

Абстрагирование позволяет выделить наиболее значимые свойства объекта и тем самым моделировать объект, отличающийся от других объектов. С помощью абстракции разработчики промышленных ИС могут решать разнообразные сложные проблемы, последовательно дифференцируя их на более простые [3].

Инкапсуляция предназначена для сокрытия реализации элементов, описывающих поведение объекта.

Под модульностью подразумевается свойство разделения программы на независимые составные части – модули.

Иерархия позволяет упорядочить абстракции и модули, формируя из них уровни взаимодействия.



Рассматривая ООП в контексте проектирования крупных промышленных ИС, можно отметить, что исходный код данных ИС характеризуется необоснованной сложностью, излишним дублированием, шаблонными ошибками, неопределенностью, наличием сильной взаимосвязи между модулями. Как следствие, оценка бюджета разрабатываемого проекта, классификация расходов, планирование сроков разработки и анализ затрат сопровождаются неточностями и искажениями.

В связи с этим особую актуальность приобретает применение инновационного подхода к проектированию ПО, который с точки зрения разработки ПО позволит

- значительно повысить уровень программной абстракции;
- разрабатывать ПО как независимую совокупность различных надежных компонентов функциональности и предметной области;
- повысить качество исходного кода: уменьшить дублирование и, как следствие, объем, повысить прозрачность (улучшить структуру) и читаемость, упростить тестирование;
- не зависеть от возможного изменения конфигурации ИС (масштабирование);
- систематически добавлять и модифицировать функциональность (адаптивность);
- упростить разработку документации и сопровождение ИС.

С точки зрения экономических показателей позволит

- повысить точность оценки предполагаемого бюджета разрабатываемой ИС;
- улучшить качество решений при планировании сроков разработки ИС;
- рациональнее распределять расходы по аспектам разработки ИС;
- корректно отражать величину трудозатрат при разработке ИС на каждом из этапов жизненного цикла ПО.

Парадигмой программирования, нацеленной на решение указанных проблем, является *аспектно-ориентированное программирование* (АОП).

## Описание АОП

АОП является современным развитием ООП и предназначено для отделения бизнес-логики ИС от сквозных функций.

Сквозная функциональность (Cross-cutting concern) – это функциональность, рассеянная по всему исходному коду ПО, систематически независимая от предметной области. АОП предоставляет возможность вызывать код сквозных функций без изменения исходного кода программы в определенный момент работы программы [4]:

- перед вызовом метода;
- после вызова метода независимо от результата;
- после успешного вызова (выполнение метода завершилось успешно);
- после исключения (вызванный метод вызвал исключение);
- вокруг (до и после вызова метода).

К общесистемным требованиям обычно относятся сквозные функции (рис. 2).

К сокращению количества вариантов сквозной функциональности приводит реализация главной цели АОП, которая состоит в том, что типовой код (не связанный с решением прикладных задач предметной области и повторяющийся от компонента к компоненту) не должен смешиваться с прикладной бизнес-логикой [4].

## Описание предлагаемой модели ПО

Наиболее общую модель архитектуры промышленной ИС без использования АОП можно представить множеством  $ИС_M = \left\{ \underbrace{СПО}_{СФ}, \underbrace{\sum Б}_{СФ} \right\}$ , где  $ИС_M$  –

обозначение модели промышленной ИС; СПО – сведения о предметной области;  $\sum Б$  – набор реализации функций бизнес-логики ИС; СФ – сквозная функциональность.

Сведения о предметной области представляют собой формализованную структуру, отображающую взаимодействие бизнес-процессов предприятия.

Под *бизнес-логикой* понимается набор функций, необходимый для корректного и полного описания алгоритма взаимодействия бизнес-процессов предприятия на каком-либо языке программирования.

Общая модель архитектуры промышленной ИС без использования АОП в соответствии с представленным множеством отражена на рисунке 3.

В данном подходе целевой код будущей ИС зависит от общесистемных требований (сквозной функциональности). Это приводит к увеличению сложности модификации кода и невозможности повторного использования кода, реализующего конкретную, специфичную для данной предметной области поведенческую особенность ИС.

Модель ПО с использованием АОП можно транслитерировать в следующее представление:  $ИС_{МАОП} = \{СПО, \Sigma Б, СФ\}$ , где  $ИС_{МАОП}$  – обозначение модели промышленной ИС с использованием АОП; СПО – сведения о предметной области;  $\Sigma Б$  – набор реализаций функций бизнес-логики ИС; СФ – сквозная функциональность.

Общая модель архитектуры промышленной ИС с использованием АОП представлена на рисунке 4.

Очевидно, что данный подход позволяет изолировать целевой код сквозной функциональности от предметной области и бизнес-логики, следовательно, вынести реализацию функций сквозной функциональности в отдельный, независимый процесс разработки.

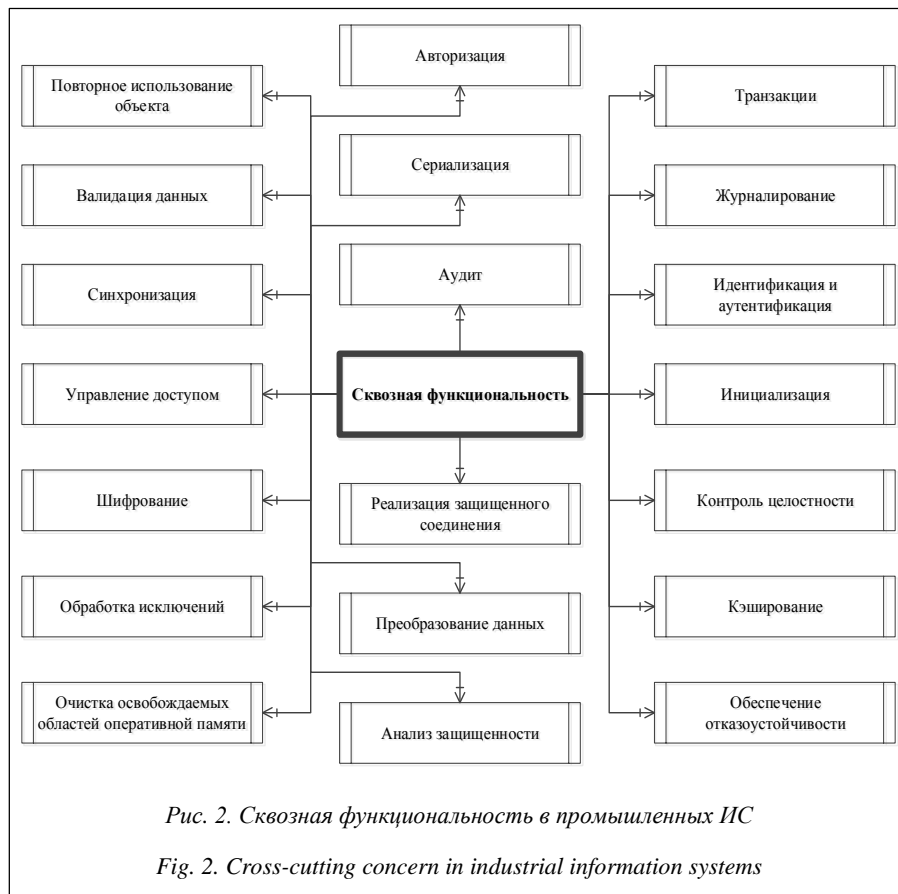
### Иллюстрация использования АОП

В качестве иллюстрации использования АОП рассмотрим следующий фрагмент программы на языке Java:

```
public class UserFileService {
    ...
    public void getCustomInfo(File file) {
        if (file == null) {
            throw new FileNotFoundException(
                file.getPath() + " not found...");
        }
        if (this.id == file.id) {
            System.out.println(file.getInfo());
        }
        else {
            throw new SecurityException(
                "Not allowed");
        }
        ...
    }
    ...
}
```

Данный модуль (метод getCustomInfo объекта класса UserFileService) реализует алгоритмы двух аспектов:

- вывод на экран титульных сведений о содержимом объекта класса File – главный аспект;
- обеспечение безопасности (в данном случае проверка совпадения значений идентификаторов файлов, доступных для чтения текущему пользователю) – сквозная функциональность.



Поскольку к одному модулю программы предъявляются требования сразу двух аспектов, один из которых реализует сквозную функциональность, средствами АОП необходимо изолировать данные аспекты в независимые модули (для примера был выбран язык AspectJ, который является расширением языка Java):

```
public class UserFileService {
    ...
    public void getCustomInfo(File file) {
        System.out.println(file.getInfo());
    }
    ...
}

public aspect Security {
    ...
    void around(UserFileService service,
        File file): call(void
        UserFileService.getCustomInfo(File))
        && args(file) && target(service) {
        if (file == null) {
            throw new FileNotFoundException(
                file.getPath() + " not found...");
        }
        if (file.id == service.id) {
            proceed(service, file);
        }
        else {
            throw new SecurityException(
                "Not allowed");
        }
        ...
    }
    ...
}
```

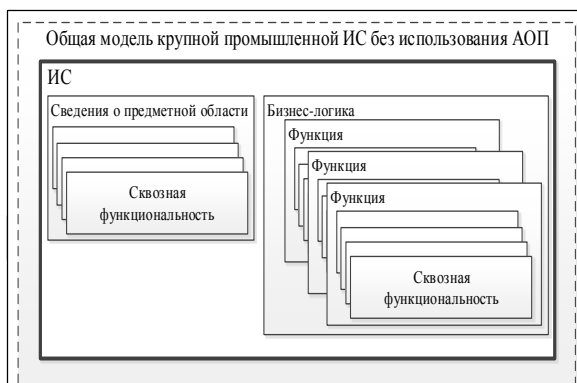


Рис. 3. Общая модель ИС без использования АОП

Fig. 3. A general model of an information system without AOP

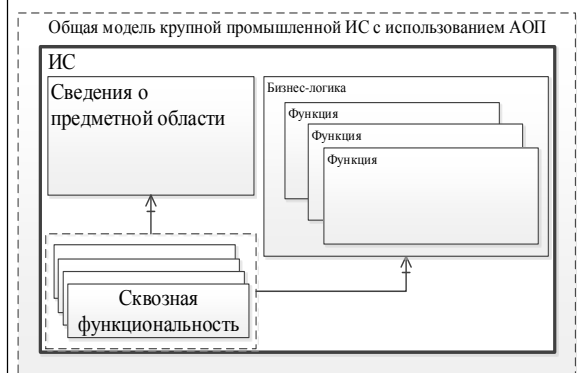


Рис. 4. Общая модель ИС с использованием АОП

Fig. 4. A general model of an information system with AOP

Как видно, реализация чтения титульных сведений о содержимом объекта класса File теперь изолирована от сквозной функциональности (проверки совпадения значений идентификаторов файлов, доступных для чтения текущему пользователю).

Параметр *aspect* предназначен для инициализации нового аспекта – Security.

С помощью конструкции *around* объявляется, что вместо метода `UserFileService.getCustomInfo(File)` будет выполнен метод (алгоритм), указанный в фигурных скобках [5]. Алгоритм реализует сквозную функциональность (простую проверку совпадения идентификаторов файлов).

Служебное слово *proceed* выполняет метод `getCustomInfo` объекта `service` класса `UserFileService` с аргументом `file` класса `File` в случае успеха выполнения алгоритма сквозной функциональности [6].

**Преимущества и недостатки методологии АОП** с точки зрения наибольшей практической ценности приведены в таблице.

В целом можно отметить, что основная польза АОП заключается в повышении прозрачности кода

бизнес-логики и в упрощении прикладных модулей, поскольку они будут содержать только базовую функциональность, а вторичные проблемы будут вынесены в аспекты.

### Преимущества и недостатки АОП

#### AOP advantages and disadvantages

Преимущество	Недостаток
Более прозрачный код логики (уменьшение связанности классов)	Отсутствие структурированной информации о разработке ПО с использованием данной методологии
Легкость модификации кода	Высокий порог вхождения (трудности с пониманием концепции)
Меньшая вероятность возникновения ошибок	Ограниченная помощь среды разработки (IDE)
Соблюдение принципа DRY (Don't repeat you're self – не повторяйся) [7]	Отсутствие подробной документации на русском языке
Уменьшение шаблонных ошибок	Незначительное снижение скорости выполнения кода
Упрощение разработки документации, и, как следствие, улучшение ее читаемости	
Улучшение модульности ИС (модули содержат только базовую функциональность, а вторичные проблемы вынесены в аспекты)	
Упрощение модульного тестирования (unit testing)	

Недостатки в контексте использования методологии АОП при разработке крупных промышленных ИС не являются существенными.

В заключение сделаем следующие выводы. Реализация сквозной функциональности в отдельные автономные модули, называемые аспектами, является существенным отличием АОП от иных парадигм разработки ПО.

Аспектно-ориентированный сценарий жизненного цикла программного продукта, с одной стороны, предполагает упрощение адаптации к каждому конкретному IT-проекту, с другой, поднимает качество оценки эффективности экономических показателей на новый уровень.

Обобщая сказанное, можно констатировать, что реализация большей части будущей функциональности в виде аспектов – потенциальный шаг к повышению качества принимаемых экономических решений, вследствие чего возможны

- оценка предполагаемого бюджета разрабатываемой ИС без фактического проектирования;
- варьирование численностью команды разработчиков при наличии эквивалентной работы;

– гибкое распределение сроков разработки и соблюдение требуемого качества программного продукта независимо от количества функциональных точек [8];

– уменьшение трудности поиска разумного компромисса между функционалом и затратами ресурсов на создание ИС;

– прозрачность понимания в каждый конкретный момент времени реальной продуктивности разработчиков.

Абстрагирование и модуляризация совокупности задач в аспекты позволяет получить корректное отображение связей между бизнес-процессами и требованиями к реализации объектов предметной области проектируемой ИС.

Независимость целевого кода и аспектов способствует реализации более гибкого процесса разработки ПО, тем самым предполагается улучшение качества принимаемых экономических решений на всех итерациях жизненного цикла программного продукта.

## Литература

1. Programming Community Index – TIOBE Index. URL: [http://www.tiobe.com/tiobe\\_index?page=index](http://www.tiobe.com/tiobe_index?page=index) (дата обращения: 28.02.2016).
2. Блинов И.Н., Романчик В.С. Java. Методы программирования. Минск: Четыре четверти, 2013. 896 с.
3. Стелтинг С., Маасен О. Применение шаблонов Java. Библиотека профессионала. М.: Вильямс, 2002. 576 с.
4. Йенер М., Фидом А. Java EE. Паттерны проектирования для профессионалов. СПб: Питер, 2016. С. 122–123.
5. Laddad R., Johnson R. AspectJ in Action: Enterprise AOP with Spring Applications, Manning Publ., 2009. 568 p.
6. Gradecki J.D., Lesiecki N. Mastering AspectJ: Aspect-Oriented Programming in Java, Wiley Publ., 2003. 456 p.
7. The Don't Repeat Yourself Principle and the Wormhole Anti-Pattern. URL: <http://codebetter.com/jeremymiller/2007/03/22/the-dont-repeat-yourself-principle-and-the-wormhole-anti-pattern/> (дата обращения: 28.02.2016).
8. Уоллис К. Spring в действии. М.: ДМК Пресс, 2013. С. 202–203.
9. Буч Г. Объектно-ориентированный анализ и проектирование. М.: Вильямс, 2010. С. 506–509.
10. Репин В.В., Елиферов В.Г. Процессный подход к управлению. Моделирование бизнес-процессов. М.: Манн, Иванов и Фербер, 2013. 544 с.

DOI: 10.15827/0236-235X.115.149-153

Received 05.04.16

## ASPECT-ORIENTED PROGRAMMING AS AN APPROACH TO IMPROVE THE EFFICIENCY OF IT PROJECTS ECONOMIC PERFORMANCE

<sup>1</sup>Goncharov A.A., Postgraduate Student, [andre\\_goncharov@hotmail.com](mailto:andre_goncharov@hotmail.com)

<sup>1</sup>Semenov N.A., Dr.Sc. (Engineering), Professor, [is@tstu.tver.ru](mailto:is@tstu.tver.ru)

<sup>1</sup>Tver State Technical University, Nikitin Quay 22, Tver, 170026, Russian Federation

**Abstract.** Budget evaluation of the developed project, expenditure classification, time development planning and cost analysis are accompanied by inaccuracies and distortions due to the fact that the initial code of industrial information systems is characterized by undue complexity, excessive duplication, formulaic errors, uncertainty, strong relationships between modules.

The article is devoted to the method of solving the current problem of increasing efficiency of IT-projects economic performance. It describes the modern approach to large-scale information system design using aspect-oriented programming.

Aspect-oriented programming is a modern development of object-oriented programming. It is intended to separate business logic from information systems through features. Business logic describes business rules implementation for particular industrial enterprises.

The article highlights the relevance of using aspect-oriented programming in large industrial information systems. It also describes its advantages and disadvantages in terms of the greatest practical value. The authors propose a mechanism that allows abstracting business processes at all stages of design iterations and implementation of non-essential (auxiliary) aspects of the system.

In addition, the article provides a practical example of using aspect-oriented programming.

Abstracting a set of problems into aspects allows obtaining the correct display of the links between business processes and requirements to the implementation of the designed information system domain objects.

**Keywords:** aspect-oriented programming, AOP, task, information system, engineering, programming, efficiency.

## References

1. Programming Community Index – TIOBE Index. Available at: [http://www.tiobe.com/tiobe\\_index?page=index](http://www.tiobe.com/tiobe_index?page=index) (accessed February 28, 2016).
2. Blinov I.N., Romanchik V.S. Java. Metody programmirovaniya [Java. Programming Methods]. Minsk, Chetyre Chetverti Publ., 2013, 896 p.
3. Stelting S., Maassen O. Applied Java Patterns. Prentice Hall Publ., 2002, 608 p. (Russ.ed.: Moscow, Williams Publ., 2002, 576 p.).
4. Yener M., Theedom A. Professional Java EE Design Patterns. Wrox Publ., 2015, 264 p. (Russ.ed.: St. Petersburg, Piter Publ., 2016, pp. 122–123).
5. Laddad R., Johnson R. AspectJ in Action: Enterprise AOP with Spring Applications. Manning Publ., 2009, 568 p.
6. Gradecki J.D., Lesiecki N. Mastering AspectJ: Aspect-Oriented Programming in Java. Wiley Publ., 2003, 456 p.
7. The Don't Repeat Yourself Principle and the Wormhole Anti-Pattern. Available at: <http://codebetter.com/jeremymiller/2007/03/22/the-dont-repeat-yourself-principle-and-the-wormhole-anti-pattern/> (accessed February 28, 2016).
8. Walls C. Spring in Action. Manning Publ., 2011, 424 p. (Russ.ed.: Moscow, DМК Press Publ., 2013, 752 p.).
9. Booch G. Object-oriented analysis and design with applications. Pearson Education Publ., 2000 (Russ.ed.: Moscow, Williams Publ., 2010, 720 p.).
10. Repin V.V., Eliferov V.G. Protsessnyy podkhod k upravleniyu. Modelirovanie biznes-protsessov [Process Approach to Management. Modeling Business Processes]. Moscow, Mann, Ivanov i Ferber Publ., 2013, 544 p.

УДК 004.942

Дата подачи статьи: 29.02.16

DOI: 10.15827/0236-235X.115.154-162

## **СИСТЕМА ВИРТУАЛЬНЫХ ИНСТРУМЕНТОВ И ПРИБОРОВ ДЛЯ АВТОМАТИЗАЦИИ УЧЕБНЫХ И НАУЧНЫХ ЭКСПЕРИМЕНТОВ**

*В.М. Дмитриев, д.т.н., профессор;*

*Т.В. Ганджа, к.т.н., доцент, gandgatv@gmail.com;*

*С.А. Панов, аспирант*

*(Томский государственный университет систем управления и радиоэлектроники,  
просп. Ленина, 40, г. Томск, 634050, Россия)*

В статье рассматривается структура системы виртуальных инструментов и приборов (системы ВИП), представляющей собой комплекс программ автоматизации реально-виртуальных экспериментов. Такие эксперименты могут проводиться как на реальном объекте, так и на его компьютерной модели с использованием разработанных виртуальных приборов, каковыми являются генераторы и измерительные приборы со встроенными в них блоками математической обработки результатов измерения и моделирования. Методологическую основу системы ВИП составляет метод многоуровневого компьютерного моделирования, в формате которого строится многоуровневая компьютерная модель системы ВИП. Предназначенная для автоматизации реально-виртуальных экспериментов, она состоит из трех взаимосвязанных уровней: визуального, на котором расположены средства визуализации и интерактивного изменения значений параметров, логического, где располагается алгоритм проведения реально-виртуального эксперимента, и объектного, представляющего собой реальный технический объект и его компьютерную модель, сформированную в формате метода компонентных цепей. Будучи универсальным методом компьютерного моделирования, метод компонентных цепей позволяет автоматизировать моделирование механических, гидравлических, теплоэнергетических, электрических и электромеханических, физико-химических и химико-технологических систем. Его программной реализацией является универсальное вычислительное ядро, осуществляющее формирование и решение систем алгебродифференциальных уравнений, составленных на основе компонентных цепей, представляющих собой компьютерные модели анализируемых объектов в формате метода компонентных цепей. Для максимального приближения виртуального эксперимента к натурному используются виртуальные приборы, представленные своими многоуровневыми моделями. Для автоматизации натуральных экспериментов над реальными объектами в рамках системы ВИП реализован блок компонентов взаимодействия с измерительно-управляющим контроллером аппаратно-программного комплекса «Лабораторное автоматизированное рабочее место». Подключаемый к компьютеру посредством шины USB, он включает в себя источник постоянного напряжения, генератор напряжения произвольной формы, вольтметр, амперметр, а также двухканальный осциллограф. Они осуществляют генерацию или измерение характеристик реальных электрических сигналов с последующей их передачей в компьютер, математической обработкой и визуализацией с помощью разработанных виртуальных инструментов и приборов.

**Ключевые слова:** *реально-виртуальный эксперимент, виртуальные инструменты и приборы, многоуровневое компьютерное моделирование, технические объекты.*

Для автоматизации научно-технических исследований и практических форм обучения студентов технических вузов широко применяются средства компьютерного моделирования и созданные на их основе автоматизированные учебные практикумы и лаборатории [1]. При этом активно развиваются два направления. Одно из них, называемое *виртуальным*, основано на анализе *компьютерных моделей* (КМ) *технических объектов* (ТО) в рамках комплексов программ, позволяющих формировать виртуальные лаборатории [2]. Второе – *реально-виртуальное* [3] – предполагает создание и использование для исследований реальных ТО аппаратно-программных комплексов с универсальными возможностями сбора, компьютерной обработки и отображения информации на компьютере.

Предлагаемые в настоящее время на российском рынке программные и программно-аппаратные продукты либо узконаправленные, либо очень широкого назначения. К отечественным программным продуктам, предназначенным для автоматизации учебных и научно-исследовательских экспериментов, можно отнести виртуальные приборы

фирмы «АКТАКОМ» [4] и ПО ZetLab [5]. Созданные в этих программных комплексах виртуальные приборы позволяют интегрироваться с реальными исследуемыми объектами с помощью плат сбора данных и универсальных контроллеров. Но они не содержат средства компьютерного моделирования мультифизических ТО, что ограничивает сферу их применения при автоматизации экспериментов учебного и научно-исследовательского характера.

Одной из известных в России зарубежных программных систем для формирования *виртуальных инструментов и приборов* (ВИП) с целью их использования при исследовании реальных ТО является система LabView [6] фирмы National Instrument. Созданные в ней виртуальные приборы широко применяются для автоматизации реальных экспериментов над ТО с использованием аппаратно-программного комплекса Evlis. Являясь системой имитационного моделирования, LabView позволяет разрабатывать в формате графического G-языка модели достаточно простых ТО, которые могут быть явно разрешены относительно наблю-

даемых переменных. Их создание трудоемко, а полученная в результате модель зачастую не отражает топологическую структуру объекта, сложна для понимания обучающимися и не позволяет анализировать модель объекта в различных режимах.

Таким образом, актуально построение комплекса программ, позволяющего формировать ВИП [7], предназначенные для максимального приближения виртуальных экспериментов, выполняемых над моделями исследуемых мультифизических объектов с возможностью интерактивного изменения значений параметров исследуемых моделей, к реальным экспериментам. Причем эти же ВИП должны использоваться для автоматизации реальных экспериментов, выполняемых над ТО. Таким образом, появляется возможность замены дорогостоящих измерительных приборов, источников и генераторов виртуальными приборами, выполняющими визуализацию значений и характеристик наблюдаемых сигналов, а также максимально приближающими виртуальные эксперименты к реальным.

Для решения этой задачи реализована система ВИП [8] как комплекс программ автоматизации реально-виртуальных экспериментов. На ее базе разрабатываются автоматизированные лабораторные комплексы [9] по различным техническим (в основном электротехническим) дисциплинам. Методологической основой системы ВИП является метод многоуровневого компьютерного моделирования [10], позволяющий формировать многоуровневые КМ виртуальных приборов со встроенными в них блоками математической обработки результатов экспериментов, компонентами визуализации численных значений и графических зависимостей, а также средствами интерактивного управления параметрами прибора и управляющими воздействиями на исследуемый объект или модель.

Для проведения экспериментов над реальными ТО наряду с исследованием их КМ в рамках системы ВИП предусмотрена интеграция с аппаратно-программным комплексом *Лабораторное автоматизированное рабочее*

*место* (ЛАРМ). Он подсоединяется к компьютеру посредством шины USB и включает в себя источник постоянного тока и генератор сигналов различной формы, параметры которых управляются подаваемыми с компьютера командами, а также набор измерительных приборов: вольтметр и амперметр постоянного тока, осциллограф, измеряющий два канала напряжения относительно общей для всех приборов земли. Результаты измерения в пакетном режиме передаются в компьютер, где подхватываются определенными функциональными блоками системы ВИП. Они осуществляют предварительную математическую обработку входящих в пакеты данных, их подготовку и визуализацию в числовом или графическом виде.

### Архитектура многоуровневой КМ для формирования виртуальных приборов

Система ВИП позволяет с применением одних и тех же программно-алгоритмических принципов исследовать реальные ТО и их КМ, автоматизируя тем самым *реально-виртуальный эксперимент*, то есть эксперимент, проводимый над реальным ТО и/или его КМ, представляющий собой совокупность реального и виртуального экспериментов с возможностями их одновременного проведения. Сформированная в формате метода многоуровневого компьютерного моделирования многоуровневая КМ состоит из объектного, логического и визуального взаимосвязанных уровней (рис. 1), каждый

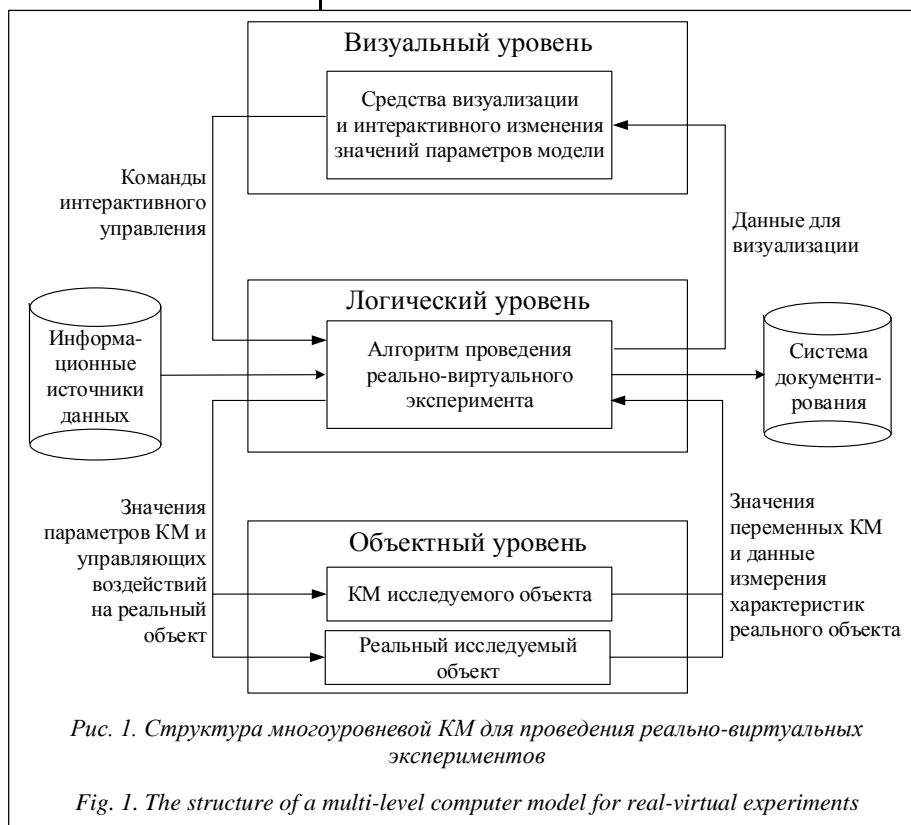


Fig. 1. The structure of a multi-level computer model for real-virtual experiments

из которых поддерживается соответствующим графическим языком представления модели.

- На *объектном уровне* располагаются средства интеграции с реальным ТО и его КМ, представленная в формате языка моделирования сложных ТО, допускающих декомпозицию на компоненты с неоднородными векторными связями, позволяющими одновременное описание энергетических и информационных потоков между компонентами. Использование классов данного типа открывает возможности построения и исследования моделей многофазных электрических цепей, многосвязных механических и многокомпонентных химико-технологических систем. Данному уровню также принадлежат компоненты, осуществляющие взаимодействие многоуровневой модели с аппаратно-программным комплексом ЛАРМ. Они выполняют такие функции, как подключение к комплексу ЛАРМ, прием результатов измерений наблюдаемых характеристик реального объекта и передача ему команд управления параметрами источников, генераторов и измерительных приборов. Данные измерения реальных сигналов наравне с результатами анализа КМ ТО передаются с объектного уровня многоуровневой КМ на ее логический уровень.

- На *логическом уровне* с помощью выражений языка математико-алгоритмических конструкций, являющегося развитием языка представления математических выражений в формат метода компонентных цепей, формируются алгоритмы выполнения реально-виртуальных экспериментов с использованием значений, поступающих из информационных источников данных, и возможностью формирования документов на основе интерактивных отчетных форм. Их внутримашинной формой представления являются алгоритмические компонентные цепи, для анализа которых используется алгоритм передачи сообщений.

- На *визуальном уровне* с помощью визуальных компонентов формируются лицевые панели виртуальных приборов, предназначенные для визуализации результатов и управления параметрами исследуемых объектов и моделей. Элементами языка ВИП являются визуальные компоненты: компоненты-визуализаторы, отображающие поступающую к ним информацию, и компоненты-регуляторы, позволяющие пользователю изменять значения соответствующих параметров исследуемой модели и переменных, входящих в алгоритмы управления.

Интеграция представленных уровней многоуровневой КМ (рис. 1) осуществляется с помощью

- измерительных компонентов, реализующих функцию передачи данных измерения характеристик объекта и результатов анализа КМ с объектного уровня на логический;

- компонентов-атрибутов, осуществляющих передачу сформированных команд управления ТО

и вычисленных значений параметров компонентов КМ объекта с логического уровня на объектный;

- компонентов-визуализаторов, выполняющих отображение переданной с логического уровня информации на визуальном уровне;

- компонентов-регуляторов, с помощью которых с визуального на логический уровень передаются пользовательские команды, направленные на изменение параметров КМ и значений управляющих воздействий на реальный ТО.

Представленные многоуровневые КМ, состоящие из моделей объектов и алгоритмов управления, а также включающие ВИП [11], предназначенные для численной обработки и визуализации результатов измерения и моделирования с возможностями интерактивного управления параметрами исследуемых объектов, моделей и алгоритмов, позволяют автоматизировать

- проведение экспериментов учебного и научно-исследовательского характера, выполняемых над реальными управляемыми ТО и (или) их КМ;

- решение задач исследования и функционального проектирования сложных ТО и систем управления ими, в основе которых лежит многократный анализ их КМ при различных значениях параметров входящих в них компонентов;

- формирование алгоритмов функционирования устройств управления, на основе которых осуществляется автоматическая генерация кода программы функционирования управляющего контроллера;

- этапы разработки и отладки ТО, измерение характеристик которых производится измерительно-управляющими контроллерами с их последующей математической обработкой и визуализацией.

### Структура системы ВИП

Структура данной системы (рис. 2), входящей в состав комплекса программ автоматизации реально-виртуальных экспериментов, включает в себя следующие программно-инструментальные блоки:

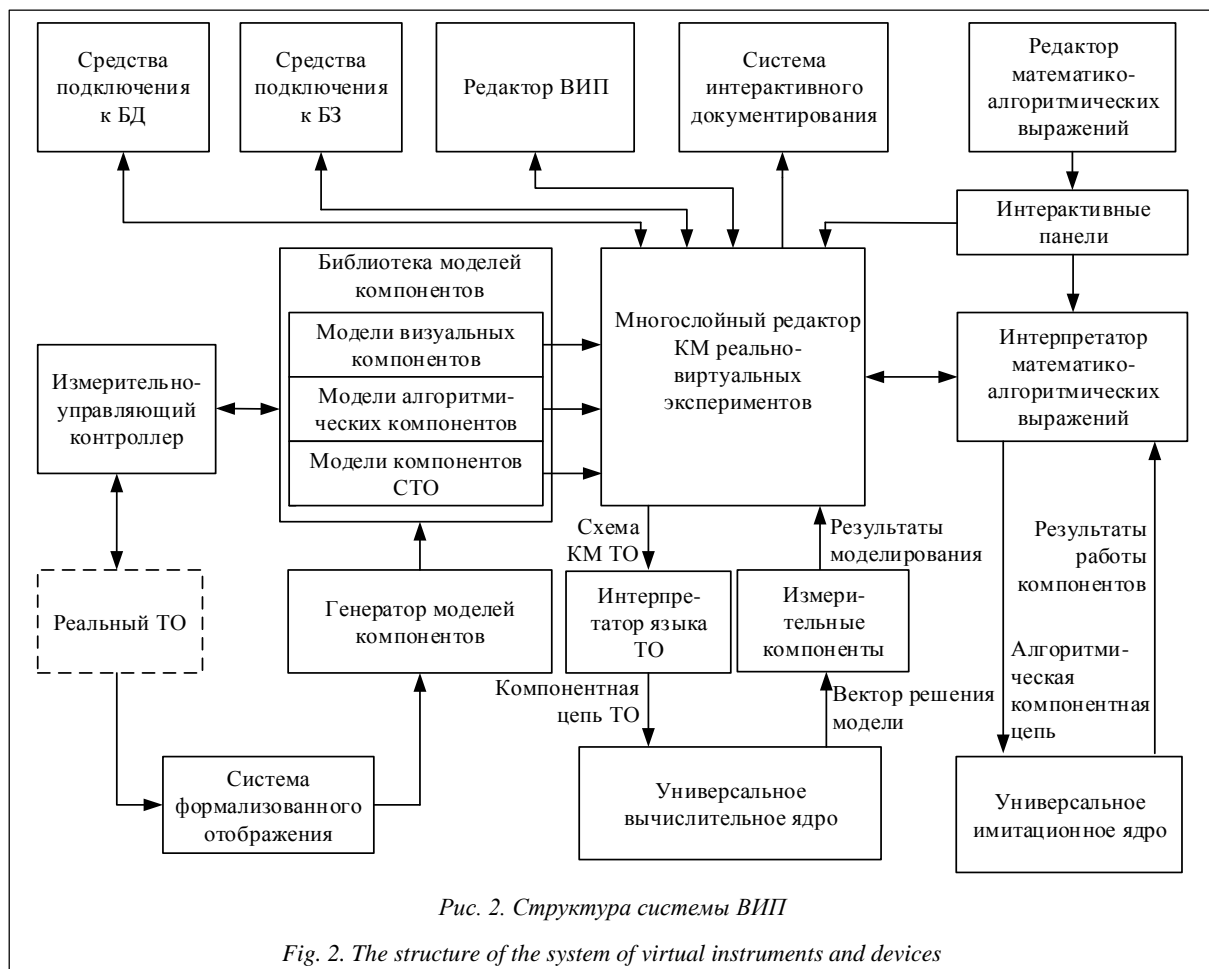
- многослойный редактор КМ, позволяющий в графическом виде формировать многоуровневые КМ реально-виртуальных экспериментов, каждый слой которого соответствует определенному уровню многоуровневой КМ экспериментов;

- интерпретатор языка ТО, предназначенный для построения компонентной цепи ТО на основе схемы его КМ, представленной в виде компонентов на объектном слое редактора;

- универсальное вычислительное ядро, осуществляющее анализ модели ТО, представленной в формате метода компонентных цепей;

- измерительные компоненты, выполняющие измерения значений соответствующих перемен-





ных и их передачу в алгоритмические компонентные цепи логического уровня;

- интерпретатор математико-алгоритмических выражений, осуществляющий преобразование сформированных схем алгоритмов в алгоритмические компонентные цепи, подлежащие обработке универсальным имитационным ядром;

- универсальное имитационное ядро, осуществляющее анализ представленных на логическом слое редактора алгоритмических компонентных цепей на основе алгоритма передачи сообщений;

- библиотека моделей компонентов, включающая в себя программно-реализованные модели визуальных и алгоритмических компонентов, а также модели компонентов ТО, обладающих неоднородными векторными связями;

- интерактивные панели, представляющие собой компоненты логического и объектного уровней и обладающие графическим интерфейсом для формирования и редактирования математических выражений и уравнений, являющихся моделями соответствующих компонентов;

- генератор моделей компонентов, представляющий собой инструментальное средство для автоматизированного создания (генерации) программно-алгоритмических текстов моделей визуальных

и алгоритмических компонентов, а также моделей компонентов ТО;

- измерительно-управляющий контроллер, представляющий собой аппаратно-программный комплекс ЛАРМ для подключения исследуемого реального объекта к компьютеру с установленной системой ВИП;

- средства подключения к хранилищам данных и знаний, представляющие собой отдельную группу компонентов логического уровня многоуровневой КМ, предназначенных для подключения к БД и БЗ;

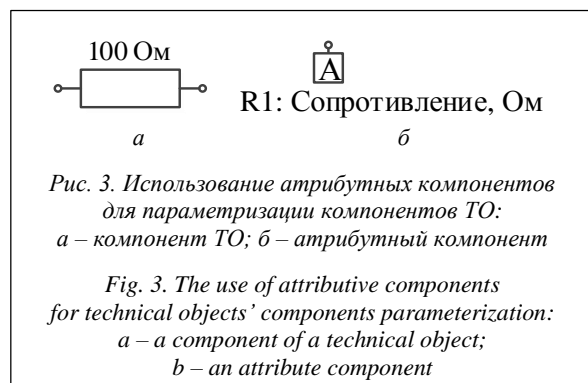
- система интерактивного документирования, предназначенная для формирования отчетов о реально-виртуальных экспериментах, проводимых с помощью многоуровневой КМ.

### Учебный эксперимент с помощью учебно-иллюстративных модулей

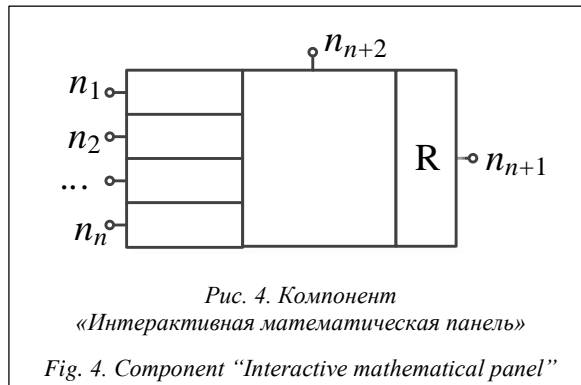
Основными задачами учебного эксперимента являются демонстрация некоторого физического эффекта, явления, а также иллюстрация принципов, законов и теорем изучаемых дисциплин. Для его выполнения в системе ВИП реализована возможность разработки учебно-иллюстративных

модулей (УИМ) [12]. Они представляют собой многоуровневые КМ, предназначенные для пояснения и визуализации основных положений и законов технических дисциплин. Такие модели разрабатываются в формате представленной на рисунке 1 структуры многоуровневой КМ. В этом случае на ее объектном уровне располагается КМ ТО, демонстрирующая изучаемый физический закон или эффект. Логический уровень содержит модель алгоритма функционирования УИМ. На визуальном уровне располагаются средства визуализации результатов выполняемых в УИМ вычислительных экспериментов и данные их математической обработки. Также на этом уровне представлены компоненты интерактивного управления значениями параметров модели исследуемого объекта.

Для варьирования параметров компонентов объектного уровня каждому изменяемому компоненту может быть поставлен в соответствие компонент «Атрибут» располагающийся на логическом уровне. Представленный на рисунке 3б компонент «Атрибут» предназначен для изменения параметра «Сопротивление» компонента «Резистор», представленного на рисунке 3а и отображаемого на объектном уровне многоуровневой КМ УИМ.



На логическом уровне многоуровневой КМ из алгоритмических компонентов и интерактивных математических панелей могут быть набраны выражения предварительного расчета параметров компонентов исследуемой модели УТО, представленной на объектном уровне, и блоки обработки результатов моделирования. Если математическое выражение, предназначенное для вычисления параметров или обработки результатов моделирования, достаточно простое, оно может быть набрано из взаимосвязанных компонентов, каждый из которых на логическом уровне редактора отображает определенную математическую операцию (оператор или функцию). Для ввода и вычисления достаточно сложного математического выражения реализован компонент «Интерактивная математическая панель» (рис. 4). Он представляет собой компонент с переменным числом связей  $n_1, n_2, \dots, n_n$ , находящихся слева от него. Для каждой из них в его свойствах задается имя переменной, которым

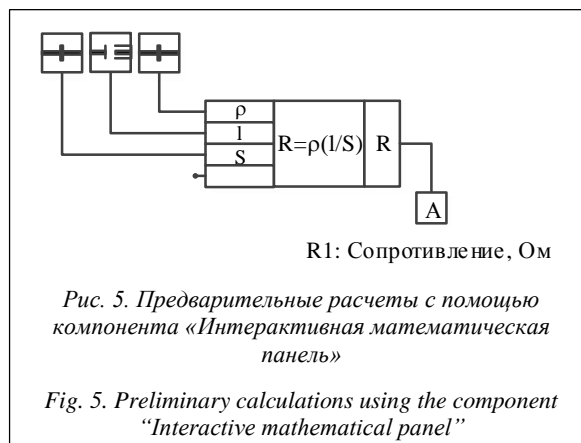


она представлена в математическом выражении, введенном пользователем в редакторе математико-алгоритмических выражений. Он открывается по двойному щелчку мыши по изображению данного компонента в многослойном редакторе КМ реально-виртуальных экспериментов.

Также компонент «Интерактивная математическая панель» имеет одну связь  $n_{n+1}$ , находящуюся справа от него. С его помощью осуществляется передача результата вычисления математического выражения, введенного в редактор математико-алгоритмических выражений. Узел  $n_{n+2}$  предназначен для передачи данного выражения в компонент «Визуализатор математических выражений», который отображает его на визуальном уровне многоуровневой КМ УИМ.

К узлам  $n_1, n_2, \dots, n_n$  могут быть подключены компоненты-регуляторы, предоставляющие возможность пользователю задавать значения соответствующих параметров анализируемой в УИМе модели ТО, и компоненты-измерители результатов первичных измерений значений переменных расположенной на объектном уровне КМ исследуемого объекта.

Результаты выполняемых с помощью компонента «Интерактивная математическая панель» вычислений могут являться новыми значениями параметров модели объекта либо итоговыми результатами и выводиться на средства визуализации. На рисунке 5 приведен пример расчета значения со-



противления резистора по удельному сопротивлению материала  $\rho$ , его длине  $l$  и диаметру  $d$ .

В случае применения компонента «Интерактивная математическая панель» для обработки результатов вычислительных экспериментов и вычисления их косвенных результатов на основе прямых он может быть подключен так, как показано на рисунке 6, где  $V$  и  $A$  – измерительные компоненты вольтметр и амперметр, осуществляющие передачу результатов моделирования с объектного уровня на логический.

Таким образом, предназначенный для интерактивной демонстрации некоторого теоретического положения, закона, теоремы или расчетной формулы некоторой учебной дисциплины УИМ представляет собой многоуровневую КМ, созданную в системе ВИП.

На объектном уровне системы располагается КМ подлежащего изучению ТО. На рисунке 7 представлена модель объекта, предназначенная для изучения закона Ома.

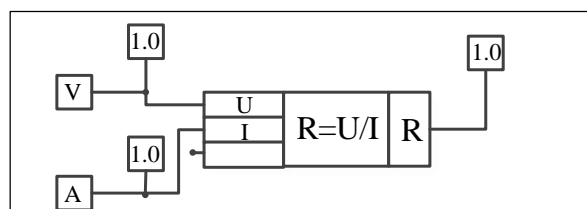


Рис. 6. Пример построения блока обработки результатов вычислительных экспериментов

Fig. 6. Example of constructing a processing unit of the result of computational experiments

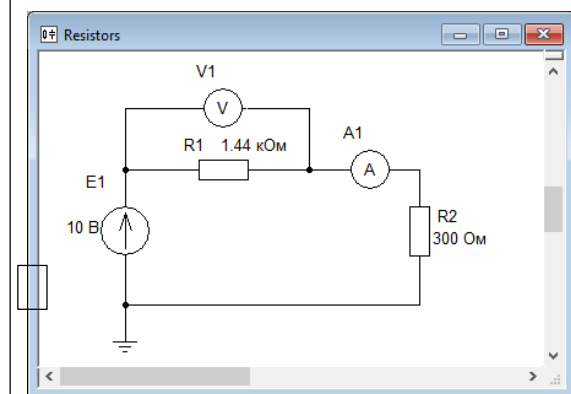


Рис. 7. Модель объекта для изучения закона Ома

Fig. 7. The object model for studying Ohm's Law

На логическом уровне многоуровневой КМ УИМ с помощью компонента «Интерактивная математическая модель» и других математико-алгоритмических компонентов сформированы расчетные схемы, представленные на рисунках 5 и 6. Его шагами являются предварительный расчет параметров компонентов модели исследуемого объекта, а также компоненты математико-алгоритмической обработки результатов ее анализа.

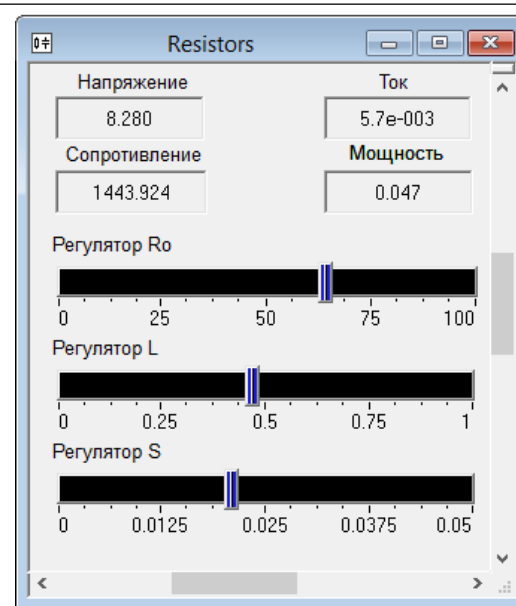


Рис. 8. Блок визуализации и интерактивного управления параметрами УИМ

Fig. 8. The unit for imaging and interactive control of educational and illustrative module parameters

Визуальный уровень включает в себя средства визуализации результатов расчетов и экспериментов, а также средства интерактивного изменения варьируемых параметров. В рассматриваемом случае, представленном на рисунке 8, – удельного сопротивления  $R_0$ , длины сопротивления  $l$  и площади поперечного сечения  $S$ .

Таким образом, построенный УИМ, сформированный на трех взаимосвязанных уровнях многоуровневой КМ виртуального прибора, включает в себя модель исследуемого объекта, блоки предварительного расчета параметров ее компонентов и блоки обработки результатов выполняемых вычислительных экспериментов.

### Реально-виртуальные эксперименты

Для проведения реально-виртуальных экспериментов с помощью аппаратно-программного комплекса ЛАРМ в системе ВИП разработан виртуальный прибор «Измерительно-генераторный стенд». Его основными задачами являются отображение измеренных временных характеристик реальных сигналов с управлением масштабами визуализации по оси абсцисс и оси ординат, а также варьирование значений характеристик задающих приборов: источника постоянного напряжения и генератора переменного напряжения. Лицевая панель ВИП «Измерительно-генераторный стенд» представлена на рисунке 9.

Взаимодействие виртуального прибора «Измерительно-генераторный стенд» с аппаратно-программным комплексом ЛАРМ, представленным на рисунке 10, осуществляется с помощью USB.

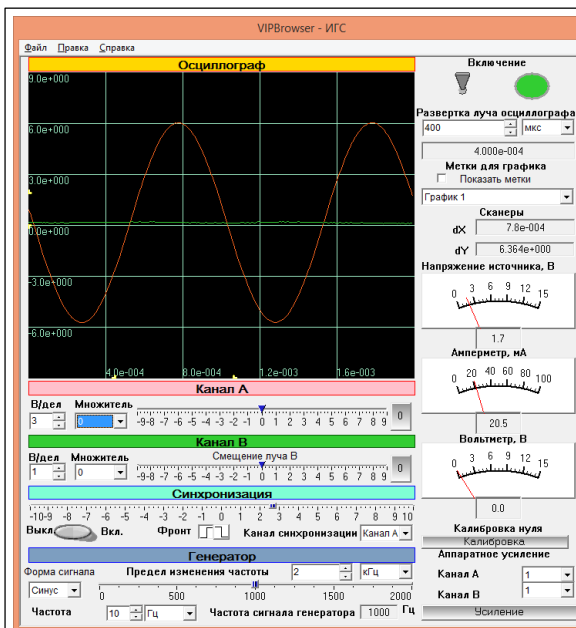


Рис. 9. Лицевая панель виртуального прибора «Измерительно-генераторный стенд»

Fig. 9. The front panel of a virtual device "Measuring and generating stand"

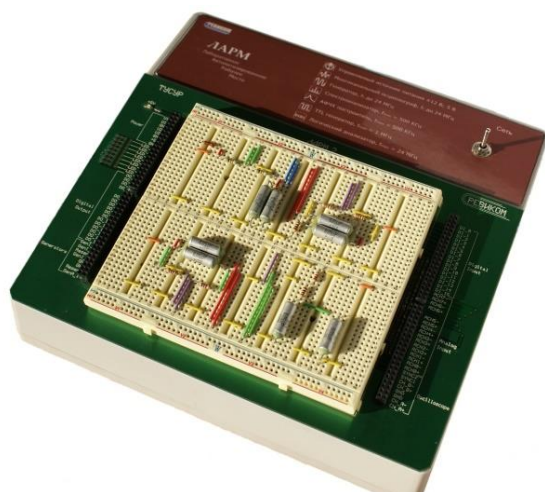


Рис. 10. Аппаратно-программный комплекс ЛАРМ

Fig. 10. Hardware-software complex "LaboratoryWorkstation"

Аппаратно-программный комплекс ЛАРМ является универсальным измерительным комплексом широкого назначения. Области его возможного использования – учебные лаборатории, автоматизация научных исследований в физике, электротехнике, теории и системах автоматического и автоматизированного управления, диагностика в медицине, настройка электронных схем. Он может быть использован для создания автоматизированных измерительных систем, являющихся основой рабочих мест исследователя, настройщика, метролога, а также учащегося и студента. В его состав

входят источник постоянного напряжения, генератор переменного периодического напряжения различной формы, осциллограф двухканальный цифровой, вольтметр и амперметр. На макетной плате ЛАРМ может быть собрана электрическая цепь из реальных элементов, к которой с помощью коммутационных проводов подключаются источники, генератор и измерительные приборы. Аппаратно-программный комплекс ЛАРМ реализован на основе микроконтроллера X-Mega, который функционирует на основе сценария, сформированного на языке X-Robot [4].

Управление характеристиками генератора заключается в изменении формы, амплитуды и частоты его сигнала. Варьирование амплитуды генератора производится аппаратным способом с помощью органа управления «Генератор», выведенного на панель ЛАРМ. Изменение формы сигнала и его частоты осуществляется с помощью компонентов на лицевой панели виртуального прибора «Измерительно-генераторный стенд», образующих блок управления «Генератор». Форма сигнала генератора изменяется с помощью соответствующего комбинированного списка и может принимать значения: синус, треугольник и прямоугольник. Плавное изменение частоты сигнала осуществляется с помощью регулятора с бегунком, а его максимальное значение изменяется с помощью цифрового табло со спином и комбинированного списка, образующих группу визуальных компонентов «Предел изменения частоты». Для точного задания частоты, которая может изменяться от 10 Гц до 1000 кГц, на лицевой панели виртуального прибора «Измерительно-генераторный стенд» предусмотрена группа визуальных компонентов «Частота».

Характеристикой осциллографа является развертка лучей, которая одинакова как для луча А, так и для луча В, подаваемых по двум одноименным каналам осциллографа. Для каждого из каналов по отдельности задаются усиление и смещение сигнала.

Под разверткой луча осциллографа в данном случае понимается цена одного деления по оси абсцисс. Она может задаваться в секундах, миллисекундах или микросекундах. Управление разверткой лучей осциллографа осуществляется с помощью группы визуальных компонентов «Развертка луча осциллографа».

Параметрами лучей осциллографа, позволяющими изменять их визуализацию по оси ординат, являются усиление и смещение, которые устанавливаются для каждого канала в отдельности. Усилением называется величина, соответствующая цене деления оси. Она задается визуальными компонентами. С помощью цифрового табло «В/дел» задается значение усиления, изменяющееся от 1 до 10. Комбинированным списком «Множитель» задается степень числа 10, имеющая возможность изменяться от –6 до 6. Число, полученное в резуль-

тате возведения цены деления на множитель, является истинным усилением сигнала.

Смещением является величина, на которую относительно центра окна визуализации смещается ось абсцисс, то есть ноль синусоидального сигнала без смещения. Эта величина задается с помощью регулятора «Смещение луча А» или «Смещение луча В». Около каждого из них установлена кнопка «0», с помощью которой производится автоматическая установка смещения сигнала в 0.

Для получения неподвижного изображения временных диаграмм на экране каждые последующие траектории луча должны пробегать по экрану в циклах развертки по одной и той же кривой. Эта возможность обеспечивается схемой синхронизации переводом выключателя в положение «Вкл.». Также с помощью группы визуальных компонентов «Синхронизация» выбираются сигнал, по которому осуществляется синхронизация, и его фронт – передний или задний, а также задается уровень синхронизации, который может варьироваться с помощью соответствующего регулятора в диапазоне от –9 до 9.

Для исследования размахов, амплитуд и временных (фазовых сдвигов) сигналов используются сканеры, которые представлены уголками на левой и нижней сторонах поля визуализации осциллограмм. На каждой стороне используются по два сканера, разница значений между которыми выводится в цифровые табло, входящие в группу визуальных компонентов «Сканеры».

С помощью соответствующих цифровых табло и представленных цифровых табло на лицевой панели прибора «Измерительно-генераторный стенд» осуществляется визуализация значений источника сигнала постоянного тока, амперметра и вольтметра, с помощью которых измеряются ток и напряжение цепей постоянного тока.

Для калибровки нулей измерительных приборов аппаратно-программного комплекса ЛАРМ, которые требуется устанавливать всякий раз в силу особенностей измерительных каналов, на лицевую панель прибора выведена кнопка «Калибровка». С помощью группы визуальных компонентов «Аппаратное усиление» осуществляется установка значения усиления измерительных каналов, что позволяет с помощью ЛАРМ исследовать сигналы достаточно малой амплитуды.

Таким образом, виртуальный прибор «Измерительно-генераторный стенд» содержит в себе все необходимые функциональные возможности для исследования электрических сигналов, проходящих в реальных цепях и измеряемых с помощью аппаратно-программного комплекса ЛАРМ. На базе компонентов, осуществляющих связь виртуального прибора с ЛАРМ, могут быть реализованы и другие виртуальные приборы, расширяющие возможности реально-виртуального эксперимента.

В заключение сделаем следующие выводы. Система ВИП позволяет выполнять реально-виртуальные эксперименты учебного и научно-исследовательского характера. На ее базе могут реализовываться виртуальные и реально-виртуальные лаборатории, компьютерные задачки и интерактивные тренажеры, применяемые в средних и высших учебных заведениях технического профиля, а также автоматизироваться процессы подготовки, переподготовки и аттестации управленческих кадров на предприятиях.

Система ВИП, являющаяся комплексом программ автоматизации реально-виртуального эксперимента, реализована на основе многоуровневой КМ, разработанной в формате метода многоуровневого компьютерного моделирования. Он расширяет перспективы компьютерного моделирования, открывая дополнительные возможности использования одних и тех же виртуальных инструментов и приборов для исследования реальных объектов и их КМ.

### Литература

1. Черкашин С.В., Любенко А.Ю., Баран Е.Д., Кухто А.В., Марченко И.О. Автоматизированные учебные практикумы и лаборатории // Образовательные, научные и инженерные приложения в среде LabView и технологии National Instruments: сб. тр. V Междунар. науч.-практич. конф. М.: Изд-во РУДН, 2006. С. 465.
2. Тумковский С.Р. Виртуальная лаборатория как элемент электронного обучения // Инновации на основе информационных и коммуникационных технологий. 2014. № 1. С. 115–117.
3. Зеленский В.И. Изучение закона Ома и определение удельного электрического сопротивления вещества проводника (Реально-виртуальная лаборатория по физике). Ханты-Мансийск: Изд-во Югорского гос. ун-та, 2006. 5 с.
4. АКТАКОМ – программное обеспечение. URL: <http://www.aktakom.ru/support/software/> (дата обращения: 04.02.2016).
5. Программное обеспечение ZetLAB. URL: <http://www.zetlab.ru/catalog/programmnnoe-obespechenie-zetlab/> (дата обращения: 04.02.2016).
6. Федосов В.П., Нестеренко А.К. Цифровая обработка сигналов в LabVIEW: учеб. пособие. М.: ДМК Пресс, 2007. 456 с.
7. Пец А.В., Раздорский И.И. Дидактические функции виртуальных приборов в системе подготовки будущих инженеров // Изв. Балтийской гос. акад. рыбопромышленного флота: психолого-педагогич. науки. 2015. № 2 (32). С. 103–108.
8. Дмитриев В.М., Ганджа Т.В., Ганджа В.В., Мальцев Ю.И. СВИП – система виртуальных инструментов и приборов. Томск: В-Спектр, 2014. 216 с.
9. Дмитриев В.М., Кураколов А.Н., Мальцев Ю.И., Коротина Т.Ю. Автоматизированные лабораторные комплексы в учебном процессе: монография. Томск: В-Спектр, 2007. 182 с.
10. Дмитриев В.М., Ганджа Т.В. Принцип формирования многоуровневых компьютерных моделей SCADA-систем для управления сложными технологическими объектами // Информатика и системы управления. 2013. № 2 (36). С. 24–36.
11. Пец А.В. Виртуальные приборы – инструмент исследования двух реальностей // Вестн. БФУ им. И. Канта. 2013. № 10. С. 144–146.
12. Шутенков А.В. Интерпретация граничных условий учебно-иллюстративным модулем в анализе переходных процессов // Электронные средства и системы управления: докл. XI Междунар. науч.-практич. конф. Томск: Изд-во ТУСУР. 2015. № 1–2. С. 268–271.

DOI: 10.15827/0236-235X.115.154-162

Received 29.02.16

**VIP. THE SYSTEM OF VIRTUAL INSTRUMENTS AND DEVICES FOR EDUCATION  
AND SCIENTIFIC EXPERIMENT AUTOMATION**<sup>1</sup>*Dmitriev V.M., Dr.Sc. (Engineering), Professor*<sup>1</sup>*Gandzha T.V., Ph.D. (Engineering), Associate Professor, gandgatv@gmail.com*<sup>1</sup>*Panov S.A., Postgraduate Student*<sup>1</sup>*Tomsk State University of Control Systems and Radio electronics,  
Lenina Ave. 40, Tomsk, 634050, Russian Federation*

**Abstract.** The paper describes the structure of the system of virtual instruments and devices (System VID), which is a software complex for automating real-virtual experiments. These experiments can be carried out both on the real object and its computer model using developed virtual instruments, e.g. generators and measuring devices with the built-in units for mathematical processing of measurement and modeling results. The multi-level VID system is based on a computer simulation approach, which is used for constructing a VID computer model. This computer model is designed to automate real-virtual experiments. It consists of three interrelated levels: visual (with visualization and interactive change of parameters' values), logic (with the algorithm of real-virtual experiment) and object (a real technical object and its computer model in the format of the method of component circuits (MCC)). As a universal computer modeling method, MCC allows automating modeling of mechanical, hydraulic, thermal energy, electrical and electro-mechanical, physical and chemical processes. It is implemented as a general purpose computing kernel that forms and solves of algebraic and differential equation systems from component circuits, which are computer models of analyzed objects in the MCC format. Virtual instruments represented by their multi-level models are used to maximize the virtual approach to a full-scale experiment. To automate field experiments on real objects within the VID system there is a block of components interacting with measuring and control hardware-software complex "Laboratory workstation". It connects to a computer via USB and includes a constant-voltage source, a free form constant-voltage generator, a voltmeter, an ammeter and a dual-channel oscilloscope. They generate and measure characteristics of real electrical signals and then transmit them to a computer, use mathematical processing and rendering using virtual instruments and devices.

**Keywords:** real-virtual experiment, virtual instrument and devices, multilevel computer modeling, technical objects.

**References**

1. Cherkashin S.V., Lyubenko A.Yu., Baran E.D., Kukhto A.V., Marchenko I.O. Automated education practicums and laboratories *Obrazovatelnye, nauchnye i inzhenernye prilozheniya v srede LabView i tekhnologii National Instruments: sb. tr. V Mezhdunar. nauch.-praktich. konf.* [Educational, Scientific and Engineering Applications in LabView and National Instruments Technology. Proc. 5th Int. Science and Practice Conf.]. Moscow, Peoples' Friendship Univ. of Russia Publ., 2006, p. 465 (in Russ.).
2. Tumkovsky S.R. Virtual Laboratory as a part of the e-learning. *Innovatsii na osnove informatsionnykh i kommunikatsionnykh tekhnology* [Innovations Based on Information and Communication Technologies]. 2014, no. 1, pp. 115–117.
3. Zelensky V.I. *Izuchenie zakona Oma i opredelenie udelnogo elektricheskogo soprotivleniya veshchestva provodnika (Realno-virtualnaya laboratoriya po fizike)* [Studying Ohm's Law and the Definition of the Specific Electrical Resistance of a Conductor Material (Real-virtual Physics Laboratory)]. Khanty-Mansiysk, Yugra State Univ. Publ., 2006, 5 p.
4. AKTAKOM – Software. Available at: <http://www.aktakom.ru/support/software/> (accessed February 4, 2016).
5. ZetLAB Software. Available at: <http://www.zetlab.ru/catalog/programmnoe-obespechenie-zetlab/> (accessed February 4, 2016).
6. Fedosov V.P., Nesterenko A.K. *Tsifrovaya obrabotka signalov v LabVIEW* [Signal Digital Processing in LabVIEW]. Tutorial. Moscow, DMK Press, 2007, 456 p.
7. Pets A.V., Razdorsky I.I. Didactic functions of virtual instruments in training future engineers. *Izvestiya Baltiyskoy gos. akademii rybopromyslovogo flota: psikhologo-pedagogicheskie nauki* [Proc. of the Baltic Fishing Fleet State Academy: Psychological and Pedagogical Sciences]. 2015, no. 2 (32), pp. 103–108 (in Russ.).
8. Dmitriev V.M., Gandzha T.V., Gandzha V.V., Maltsev Yu.I. *SVIP – sistema virtualnykh instrumentov i priborov* [SWID. The System of Virtual Instruments and Devices]. Tomsk, V-Spectr Publ., 2014, 216 p.
9. Dmitriev V.M., Kurakolov A.N., Maltsev Yu.I., Korotina T.Yu. *Avtomatizirovannye laboratornye komplekсы v uchebnom protsesse* [Automated Laboratory Facilities in the Educational Process]. Monograph. Tomsk, V-Spectr Publ., 2007, 182 p.
10. Dmitriev V.M., Gandzha T.V. The principle of forming multilevel computer models of SCADA-systems to control complex technological objects. *Informatika i sistemy upravleniya* [Informatics and Control Systems]. 2013, no. 2 (36), pp. 24–36 (in Russ.).
11. Pets A.V. Virtual Instruments. A research tool of two realities. *Vestnik Baltiyskogo fed. univ. im. I. Kanta* [Herald of the Baltic Federal Univ. n.a. I. Kant]. 2013, no. 10, pp. 144–146 (in Russ.).
12. Shutenkov A.V. Interpretation of the boundary conditions using an education and exemplary module in transient analysis. *Elektronnye sredstva i sistemy upravleniya. Dokl. XI Mezhdunar. nauch.-praktich. konf.* [Electronic Instruments and Control Systems. Proc. 11 Int. Science and Practice Conf.]. Tomsk, TUSUR Publ., 2015, no. 1–2, pp. 268–271 (in Russ.).

УДК 004.946

Дата подачи статьи: 02.03.16

DOI: 10.15827/0236-235X.115.163-167

### **3D-МОДЕЛИРОВАНИЕ МЕТОДОВ СЪЕМКИ МОБИЛЬНЫМИ ВИДЕОСИСТЕМАМИ**

*В.В. Извозчикова, к.т.н., доцент, viza-8.11@mail.ru*

*(Оренбургский государственный университет, просп. Победы, 13, г. Оренбург, 460018, Россия);*

*А.В. Меженин, к.т.н., доцент, mejenin@mail.ru*

*(Санкт-Петербургский национальный исследовательский университет  
информационных технологий, механики и оптики (Университет ИТМО),  
Кронверкский просп., 49, г. Санкт-Петербург, 197101, Россия)*

Проводимые исследования направлены на разработку виртуальных симуляторов, имитирующих реальные условия эксплуатации, для тестирования и моделирования методов съемки мобильными видеосистемами. При этом предполагается, что средствами 3D-графики эмулируются не только окружающая среда и условия съемки, но и сами видеосистемы: средства технического зрения, бортовые камеры роботов и беспилотных летательных аппаратов. Использование таких симуляторов позволит создавать, совершенствовать и тестировать различные алгоритмы прикладных интеллектуальных систем без создания сложных дорогостоящих стендов.

Рассматриваются вопросы оценки качества работы систем распознавания и реконструкции, использующих для представления данных полигональную сетку. Для оценки качества реконструкции и точности получения полигональной сети в качестве математического аппарата предлагается использовать размерность Хаусдорфа.

Приведены результаты апробации рассматриваемых методов: использование виртуальных сред и оценка подобия полигональных сеток. На этапе сегментации и нахождения вершин и отрезков исследуемых объектов продемонстрировано использование как реальных изображений, полученных реальной камерой, так и синтезированных. На этапе реконструкции для тестирования одной из систем реконструкции также был использован набор синтезированных изображений. Для оценки подобия полигональных объектов на основе размерности Хаусдорфа разработаны соответствующие функции и приведен результат их работы.

**Ключевые слова:** *мобильные роботы, видеосистемы беспилотных летательных аппаратов, виртуальные симуляторы, методы съемки, 3D-реконструкции, фотограмметрия.*

Фото- и видеосъемка с помощью различных мобильных систем (роботов, беспилотных летательных аппаратов (БПЛА)) находит все большее применение, особенно в таких областях, как разведка, аэрофотосъемка, картографирование. Для получения качественных изображений, несущих в себе достаточное количество данных, на основе которых может быть произведена достоверная реконструкция объектов, необходимо решить задачи оптимального выбора места фото- или видеосъемки, направления и расстояния между исследуемыми объектами и мобильной видеосистемой, а также количества фото- и видеок кадров. Данные, полученные в результате съемки, являются исходным материалом для последующей обработки. От их качества во многом зависит результат. Определяющими являются не только параметры используемых камер, но и условия, методы и геометрические параметры съемки. Средства компьютерного моделирования и виртуальные симуляторы позволяют до проведения реальных съемок моделировать этот процесс [1, 2]. При этом предполагается, что средствами 3D-графики осуществляется эмуляция не только окружающей среды, но и результатов работы различных сенсоров, в частности, изображений, получаемых средствами технического зрения, изображений с бортовых телекамер роботов, БПЛА и т.д. Планирование съемки позволит получить лучшие результаты, оптимизировать процесс и избежать ошибок при последующей обработке данных. Кроме этого, такой подход можно использо-

вать при проектировании: разрабатывать, совершенствовать и тестировать различные алгоритмы прикладных мобильных интеллектуальных систем без создания сложных дорогостоящих стендов.

В первую очередь, виртуальные среды успешно применяются в области робототехники. Рассмотрим ряд разработок в этой области. Компанией «Ланит-Терком» (г. Санкт-Петербург) совместно с кафедрой системного программирования и математико-механическим факультетом СПбГУ были реализованы проекты QReal:Robots и Embox. Среди базовых возможностей можно назвать среду графического программирования и моделирование поведения робота на экране. В МГТУ МИРЭА разработаны виртуальные лабораторные стенды, позволяющие на основе кинематики типа ПУМА моделировать траекторию движения робота и оценивать влияние изменений параметров элементов приводов на динамику робота. Известны разработки исследовательской группы Института промышленных управляющих систем Политехнического университета Валенсии: ABB RobotStudio программирование с использованием симуляционных 3D-программ (Virtual Robot Technology и VirtualRobot Simulator). На основе MATLAB и Simulink разработана среда Embedded Coder Robot for LEGO Mindstorms NXT, которая позволяет создавать систему управления роботом NXT, моделировать его поведение, обрабатывать и отображать данные, используя пакет трехмерной визуализации данных Virtual Reality Toolbox.



Для реализации виртуальных сред могут быть использованы как универсальные системы 3D-моделирования Autodesk 3ds Max, Maya, Cinema 4D, так и специализированные Microsoft Robotics Developer Studio, Virtual Reality Toolbox в среде MATLAB и Simulink. Реальный мир не статичен, поэтому проектирование и использование виртуальных сред с динамически изменяемой конфигурацией актуальны и перспективны.

Одним из путей решения задачи обнаружения препятствий средствами машинного зрения является подход на базе стереозрения. Важное преимущество стереоскопических систем заключается в возможности из-за различия в положении камер различать трехмерные (3D) объекты, отстоящие от поверхности, и объекты, принадлежащие этой поверхности (блики, тени, специальный рисунок), что потенциально снижает вероятность ложных обнаружений. Использование 3D-моделирования при проектировании и испытаниях таких систем позволяет значительно повысить эффективность этого процесса.

Технологии фотограмметрии, построение 3D-моделей по набору изображений находят все большее применение не только у профессионалов [3, 4], но и у широкого круга различных компьютерных пользователей, интересующихся 3D-графикой. Примером могут служить облачный сервис 123D Catch компании Autodesk [5], Agisoft Photoscan, Strata Foto 3D-компании 3DSom [6]. Выбор оптимального метода и инструментария для восстановления реального объекта по фотографиям довольно труден [7, 8]. Кроме этого, способы съемки во многом влияют на результат. Использование виртуальных сред и цифровое моделирование изображений позволяют сравнивать результаты работы того или иного алгоритма и системы, планировать фотограмметрическую съемку для получения лучшего результата.

Для оценки условий съемки и работоспособности алгоритмов сегментации, нахождения вершин и построения линий были использованы изображения исследуемой сцены, полученные реальной камерой, и цифровые, полученные виртуальной камерой в системе трехмерного моделирования 3ds Max (рис. 1).

Дальнейшая обработка изображений выполнялась в среде MATLAB Image Processing Toolbox (IPT). Результаты работы алгоритмов сегментации на основе изображений, полученных на исследовательском стенде и с помощью виртуальной студии, показаны на рисунке 2.

Как известно, сегментация декомпозирует изображение на составляющие его области или объекты. В идеале методы сегментации в данной задаче должны достоверно выделять в изображении только пиксели, лежащие на контурах. На практике это множество пикселей редко отображает контур достаточно точно по причине шумов, разрывов

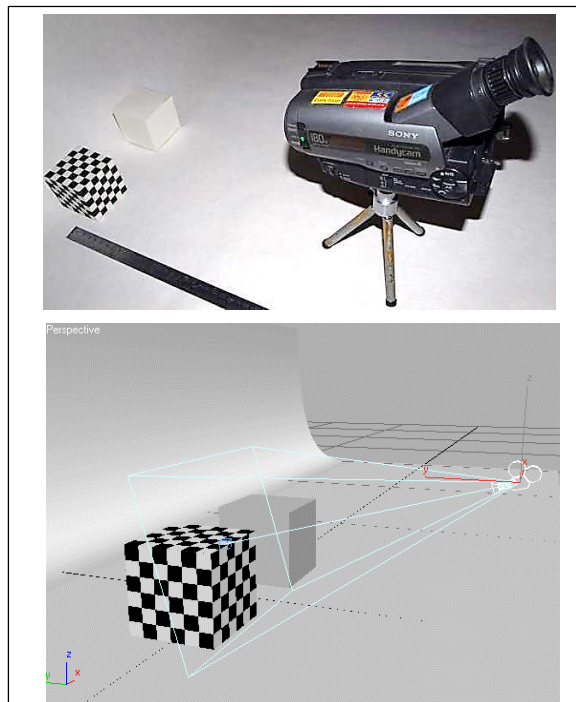


Рис. 1. Исследовательский стенд и виртуальная студия

Fig. 1. A research stand and a virtual studio

контуров из-за неоднородности освещения, а также прочих эффектов, нарушающих непрерывность яркостной картины. На рисунке 2 показаны результаты сегментации синтезированных изображений в условиях помех (смаз изображения) и неравномерного освещения (туман).

Для исследований возможностей систем фотограмметрии и оценок условий съемки была выбрана система Strata Foto 3D, формирующая 3D-контент из фотоизображений. С помощью фотокамеры осуществляется многоплановая съемка объектов на монохромном фоне. Полученные данные используются для построения 3D-объектов. Понятно, что результат моделирования во многом зависит от исходных изображений. Определяющими являются количество изображений, ракурсы съемки и параметры освещения. Для тестирования системы была создана 3D-сцена – box, размещенный на калибровочном столе, с нанесенными метками. Полученная с помощью виртуальной камеры серия изображений была импортирована в систему Strata Foto 3D. В процессе экспериментов определялись оптимальное количество изображений, количество источников света и положение камеры. Результат реконструкции показан на рисунке 3. Качество полученной модели вполне удовлетворительное, поверхность граней достаточно ровная, текстура однородная, хорошо заполняет поверхность куба.

Использование виртуальных стендов позволяет производить объективную оценку результатов фо-



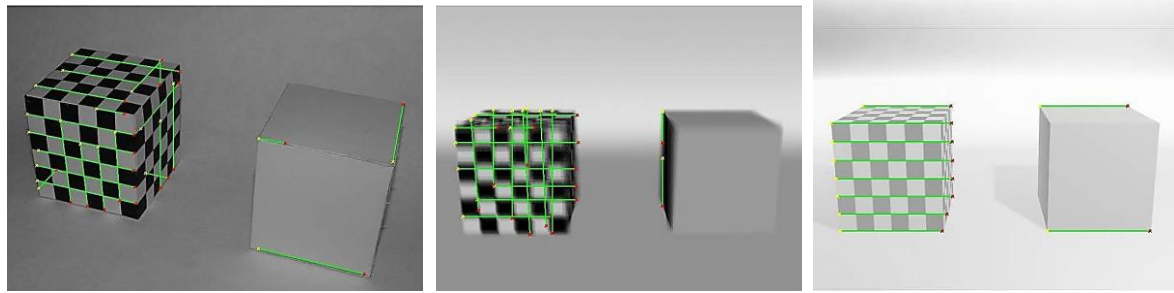


Рис. 2. Результаты обработки изображений

Fig. 2. Image processing results

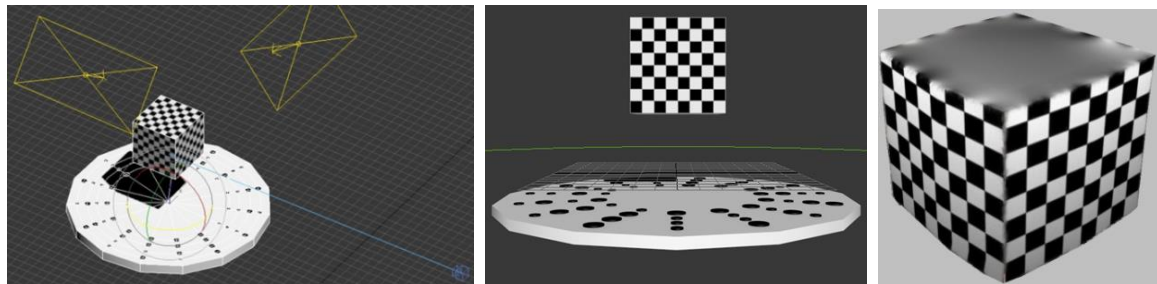


Рис. 3. Виртуальная студия и результаты реконструкции

Fig. 3. A virtual studio and reconstruction results

тограмметрии [10]. Исследуемая и реконструированная модели приводятся к полигональному виду и сравниваются. Для таких сравнений используют различные подходы, базирующиеся на вычислении евклидова расстояния и RMSE [11, 12]. Такая оценка, по мнению авторов, имеет приближенный характер, поскольку не учитывает топологические особенности сравниваемых полигональных моделей [13].

Для повышения точности оценки предлагается использовать размерность Хаусдорфа как меру отклонения сравниваемых моделей [14, 15]. В частности, эта метрика используется для оценки качества упрощения полигональных объектов при реализации технологии LOD (Levels Of Detail).

Для определения размерности Хаусдорфа рассмотрим две полигональные модели: оригинальную модель  $M$  и полученную в результате реконструкции модель  $M'$ . Для обозначения величины подобия введем показатель  $E$ , значение которого  $E(M, M')$  показывает величину отклонения одной формы от другой [16]. Топологическое пространство  $X$  называется *хаусдорфовым*, если любые две различные точки  $x, y$  из  $X$  обладают непересекающимися окрестностями  $U(x), V(x)$ . Пусть даны два набора точек:  $A=\{a_1, a_2, \dots, a_m\}$  и  $B=\{b_1, b_2, \dots, b_n\}$ . Тогда хаусдорфово расстояние определяется как  $H(A, B) = \max(h(A, B), h(B, A))$ , где  $h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$  и  $\|\bullet\|$  – евклидова норма. Значение функции  $h(A, B)$  несимметрично и называется направленным хаусдорфовым расстоянием между

$A$  и  $B$ . Эта величина может служить основой для сравнения двух полигональных поверхностей  $S_1$  и  $S_2$ . Предлагаемый авторами метод расчета метрики базируется на построении векторов нормалей к исследуемым поверхностям. Для повышения точности расчета можно воспользоваться методом усреднения – расчет средневзвешенного значения векторов нормалей, образованных парами соседних треугольников. В этом случае для нахождения координат точек пересечения векторов нормалей, построенных от одной поверхности к другой, воспользуемся следующими предположениями. Пусть  $ABC$  – произвольный треугольник;  $a, b, c$  – длины сторон, лежащие против вершин  $A, B$  и  $C$  соответственно;  $M$  – точка пересечения биссектрис. Тогда для любой точки  $O$  верно равенство

$$\overline{OM} = \frac{a \cdot \overline{OA} + b \cdot \overline{OB} + c \cdot \overline{OC}}{a + b + c}.$$

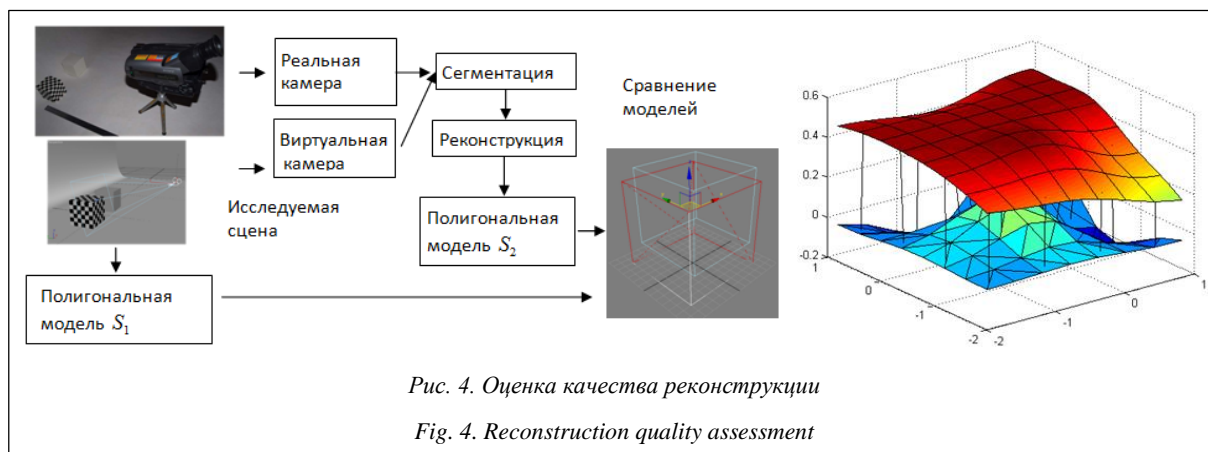
Из этой теоремы следует: если  $O$  – начало координат, то

$$x_M = \frac{a \cdot x_A + b \cdot x_B + c \cdot x_C}{a + b + c},$$

$$y_M = \frac{a \cdot y_A + b \cdot y_B + c \cdot y_C}{a + b + c},$$

$$z_M = \frac{a \cdot z_A + b \cdot z_B + c \cdot z_C}{a + b + c}.$$

Таким образом, если векторы  $\mathbf{a}$  и  $\mathbf{b}$  определены своими прямоугольными декартовыми координатами, то есть представлены в ортонормированном базисе  $\mathbf{a}=(a_x, a_y, a_z)$  и  $\mathbf{b}=(b_x, b_y, b_z)$ , а система коор-



динат правая, то их векторное произведение можно описать следующим образом:  $[\mathbf{a}, \mathbf{b}] = (a_x b_z - a_z b_y, a_z b_x - a_x b_z, a_x b_y - a_y b_x)$ .

Предложенные методы сравнения полигональных объектов на основе вычисления Hausdorff dimension были реализованы в среде MATLAB Image Processing Toolbox (IPT) в виде  $m$ -функций, в результате выполнения которых визуализируются рассчитанные векторы нормалей и величина отклонений одной поверхности от другой. На рисунке 4 представлен фрагмент программного кода, демонстрирующий использование разработанных  $m$ -функций для сравнения полигональных поверхностей, являющихся результатом двух математических трехмерных функций. Приведенный график демонстрирует использование разработанных  $m$ -функций для сравнения полигональных поверхностей, являющихся результатом двух математических трехмерных функций, где показаны векторы нормалей, построенные от полигонов одной поверхности к другой, и величина отклонения представлена в виде графика палитры цветов.

Проведенные исследования показывают целесообразность и перспективность использования виртуальных сред различной конфигурации, созданных средствами 3D-компьютерного моделирования, на этапах проектирования и планирования применения видеосистем в таких областях, как робототехника, фотограмметрия и системы компьютерного зрения. Их использование позволит совершенствовать разработку таких систем, повысит эффективность их применения. Для более точной оценки эффективности и качества рассмотренных подходов планируется дополнительное тестирование рассматриваемых методов моделирования различных условий съемки: условия неравномерного освещения, помех с учетом отражающих свойств материалов исследуемых объектов.

#### Литература

1. Piatti E., Lerma J., A virtual simulator for photogrammetry. ISPRS, Proc. V Sympos. Image Eng. and Vision Metrology, Dresden, Germany, 2006, vol. 36, no. 5, 4 p.

den, Germany, 2006, vol. 36, no. 5, 4 p.

2. Меженин А.В. Виртуальные 3D-среды как средство верификации и тестирования при проектировании // Приоритетные научные направления: от теории к практике. 2016. № 21. С. 105–110.

3. Красьников Н.Н. Метод получения 3D-изображений, основанный на диффузном отражении света сканируемыми объектами // Информационно-управляющие системы. 2009. Вып. 6. С. 7–12.

4. Eisert P., Steinbach E., and Girod B. Multi-Hypothesis, Volumetric Reconstruction of 3-D Objects From Multiple Calibrated Camera Views. Proc. of the Int. Conf. on Acoustics, Speech, and Signal Processing, 1999, pp. 3509–3512.

5. Autodesk 123D Catch | Generate 3d model from photos. URL: <http://www.123dapp.com/catch> (дата обращения: 28.02.2016).

6. Foto 3D CX for Mac PC - Create High-Quality Textured 3D Models From Your Digital Camera – Strata URL: <https://www.strata.com/foto-3d-cx-create-textured-3d-models-from-your-digital-camera/> (дата обращения: 28.02.2016).

7. Klasing K., Wollherr D., and Buss M. A clustering method for online segmentation of 3d laser data. Proc. ICRA, 2008, pp. 4043–4048.

8. Sebastian T.B., Klein P.N., and Kimia B.B. Recognition of shapes by editing shock graphs. Proc. ICCV, 2001, pp. 755–762.

9. Hielsberg M., Tsai R., Guo P., Chen C. Visibility-based urban exploration and learning using point clouds. Proc. IEEE/RSJ Int. Conf. Intell. Robots Sys., 2013.

10. Свешникова Н.В., Юрин Д.В. Оценка погрешностей восстановления трехмерных сцен алгоритмами факторизации. 2006. URL: [http://www.graphicon.ru/html/2006/proceedings/papers/fr09\\_32\\_Sveshnikova\\_Yurin.pdf](http://www.graphicon.ru/html/2006/proceedings/papers/fr09_32_Sveshnikova_Yurin.pdf) (дата обращения: 28.02.2016).

11. Cignoni P., Rocchini C., Scopigno R. Metro: measuring error on simplified surfaces. Proc. Comp. Graphics Forum 1998, vol. 17, no. 2, pp. 167–174.

12. GIS-LAB: Среднеквадратичная ошибка (RMSE). URL: <http://gis-lab.info/qa/rmse.html> (дата обращения: 28.02.2016).

13. Меженин А.В., Извозчикова В.В. Методы построения векторов нормалей в задачах идентификации объектов // Кибернетика и программирование. 2013. № 4. С. 51–58.

14. Меженин А.В., Извозчикова В.В. Размерность Хаусдорфа в задачах анализа подобия полигональных объектов // Интеллект. Инновации. Инвестиции. 2016. № 2. С. 109–112.

15. Yedjour H., Meftah B., Yedjour D. and Benyettou A. Combining spiking neural network with Hausdorff distance matching for object. Tracking Asian Journal of Applied Sc., 2014, vol. 1, pp. 63–71.

16. Teng Y.C. Remote-sensing image processing and recognition using wavelet transform and Hausdorff distance. Ph.D. Thesis, Inst. of Comp. Sc. and Inform. Eng. National Central Univ. Chung-Li, Taiwan, 2002.

**3D MODELING FOR MOBILE VIDEO SYSTEMS**<sup>1</sup>**Izvozhikova V.V.**, Ph.D/ (Engineering), Associate Professor, [viza-8.11@mail.ru](mailto:viza-8.11@mail.ru)<sup>2</sup>**Mezhenin A.V.**, Ph.D/ (Engineering), Associate Professor, [mejenin@mail.ru](mailto:mejenin@mail.ru)<sup>1</sup>Orenburg State University, Pobedy Ave. 13, Orenburg, 460018, Russian Federation<sup>2</sup>The National Research University of Information Technologies, Mechanics and Optics, Kronverksky Ave. 49, St. Petersburg, 197101, Russian Federation

**Abstract.** The research focuses on the development of virtual simulators that simulate actual service conditions for testing and simulation of shooting methods by mobile video systems. It is assumed that 3D-graphics means emulate not only the environment, shooting conditions, but also video systems themselves: visual facilities, onboard robot cameras and UAVs. The use of such simulators allows creating, refining and testing various algorithms of applied intelligent systems without creating complex and expensive stands.

The article considers the issues of performance evaluation of recognition and reconstruction systems that use a polygonal network for data presentation. To assess the reconstruction quality and accuracy of the polygonal network the authors propose the Hausdorff dimension as a mathematical tool.

The results of testing these methods are: virtual environments use and assessment of polygonal networks similarity. At the stage of segmentation and location of vertices and segments of the objects the authors show the use of both real images obtained by real camera, as well as synthesized images. At the reconstruction phase to test one of reconstruction systems they also used a set of synthesized images. In order to assess the similarity of polygonal objects based on the Hausdorff dimension the authors designed corresponding functions and showed the results of their work.

**Keywords:** mobile robots, mobile video systems, virtual simulators, shooting methods, 3D reconstruction, photogrammetry.

**References**

1. Piatti E., Lerma J. A Virtual Simulator for Photogrammetry. *ISPRS, Proc. V Sympos. Image Eng. and Vision Metrology*. Dresden, Germany, 2006, vol. 36, no. 5, 4 p.
2. Mezhenin A.V. Virtual 3D environments as a tool for verification and testing when designing. *Prioritetnye nauchnye napravleniya: ot teorii k praktike* [Foreground Research Areas: from Theory to Practice]. 2016, no. 21, pp. 105–110 (in Russ.).
3. Krasilnikov N.N. Method of 3D imaging based on diffuse light reflection by scanned objects. *Informatsionno-upravlyayushchie sistemy* [Information and Control Systems]. 2009, no. 6, pp. 7–12 (in Russ.).
4. Eisert P., Steinbach E., Girod B. Multi-Hypothesis, Volumetric Reconstruction of 3-D Objects From Multiple Calibrated Camera Views. *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing*. 1999, pp. 3509–3512.
5. Autodesk 123D Catch | Generate 3d model from photos. Available at: <http://www.123dapp.com/catch> (accessed February 28, 2016).
6. Foto 3D CX for Mac PC - Create High-Quality Textured 3D Models From Your Digital Camera – Strata. Available at: <https://www.strata.com/foto-3d-cx-create-textured-3d-models-from-your-digital-camera/> (accessed February 28, 2016).
7. Klasing K., Wollherr D., Buss M. A clustering method for online segmentation of 3d laser data. *Proc. ICRA*. 2008, pp. 4043–4048.
8. Sebastian T.B., Klein P.N., Kimia B.B. Recognition of shapes by editing shock graphs. *Proc. ICCV*. 2001, pp. 755–762.
9. Hielsberg M., Tsai R., Guo P., Chen C. Visibility-based urban exploration and learning using point clouds. *Proc. IEEE/RSJ Int. Conf. Intell. Robots Sys*. 2013.
10. Sveshnikova N.V., Yurin D.V. The factorization algorithms: results reliability and application for the epipolar geometry recovery. *Proc. Int. Conf. Graphicon 2006*. Novosibirsk, Russia. Available at: [http://www.graphicon.ru/html/2006/proceedings/papers/fr09\\_32\\_Sveshnikova\\_Yurin.pdf](http://www.graphicon.ru/html/2006/proceedings/papers/fr09_32_Sveshnikova_Yurin.pdf) (accessed February 28, 2016).
11. Cignoni P., Rocchini C., Scopigno R. Metro: measuring error on simplified surfaces. *Computer Graphics Forum*. 1998, no. 17 (2), pp. 167–174.
12. GIS-LAB: Srednekvadratichnaya oshibka (RMSE). Available at: <http://gis-lab.info> (accessed February 28, 2016).
13. Mezhenin A.V., Izvozhikova V.V. Methods of normal vector constructing in problems of polygonal objects' similarity analysis. *Kibernetika i programmirovaniye* [Cybernetics and Programming]. 2013, no. 4, pp. 51–58 (in Russ.).
14. Mezhenin A.V., Izvozhikova V.V. The Hausdorff dimension in problems of polygonal objects' similarity analysis. *Intellekt. Innovatsii i Investitsii* [Intelligence. Innovations. Investments]. 2016, no. 2, pp. 109–112 (in Russ.).
15. Yedjour H., Meftah B., Yedjour D., Benyettou A. Combining spiking neural network with Hausdorff distance matching for object. *Tracking Asian Journ. of Applied Sc.* 2014, vol. 1, pp. 63–71.
16. Teng Y.C. *Remote-sensing image processing and recognition using wavelet transform and Hausdorff distance*. Ph.D. Thesis, Inst. of Comp. Sc. and Inform. Eng. National Central Univ., Chung-Li, Taiwan, 2002.

УДК 621.396.96

Дата подачи статьи: 24.06.16

DOI: 10.15827/0236-235X.115.168-174

## **МЕТОД ОБНАРУЖЕНИЯ И СОПРОВОЖДЕНИЯ ВОЗДУШНЫХ ОБЪЕКТОВ ПО ОТРАЖЕННЫМ РАДИОСИГНАЛАМ СТОРОННИХ ИСТОЧНИКОВ В ПАССИВНО-АКТИВНЫХ СИСТЕМАХ РАДИОЛОКАЦИИ**

*(Работа выполнена при поддержке гранта Президента РФ № НШ-6831.2016.8)*

*С.А. Батчев, к.т.н., зам. начальника факультета  
(Череповецкое высшее военное инженерное училище радиоэлектроники,  
Советский просп., 126, г. Череповец, 162622, Россия);*

*А.Г. Зайцев, к.т.н., ведущий научный сотрудник;*

*А.Б. Талалаев, д.т.н., профессор, генеральный директор;*

*Д.А. Тимаков, к.т.н., зам. начальника отдела, timakov\_da@rambler.ru  
(Радиотехнические и информационные системы воздушно-космической обороны,  
ул. Зинаиды Коноплянниковой, 89, корп. 1, г. Тверь, 170041, Россия)*

В рамках проведенных исследований дается оценка уровня развития современных пассивно-активных систем радиолокации. Показано, что применение таких систем повышает эффективность контроля воздушного пространства и снижает его стоимость. В работе дано описание разработанного метода обнаружения и сопровождения воздушных объектов по отраженным радиосигналам сторонних источников в пассивно-активных системах радиолокации, основанного на измерении временной задержки между опорным и отраженным сигналами и относительной радиальной скорости контролируемого объекта. В отличие от известных подходов данный метод учитывает факторы, влияющие на качество обнаружения и определения траекторных характеристик воздушных целей, и позволяет расширить зону контроля воздушного пространства.

В ходе проведенных исследований доказано, что основными факторами, снижающими эффективность работы системы, являются уровень подавления опорного сигнала в целевом канале и отраженного сигнала в основном канале, изменение формы отраженного сигнала, уровень мощности отраженных сигналов от местных неподвижных объектов. Предложены пути уменьшения влияния данных факторов, основанные на применении алгоритмов амплитудно-фазовой коррекции сигналов в приемных каналах и децимации цифровых потоков. Применение данных подходов сокращает время выполнения процедур и повышает соотношение сигнал/шум радиолокационного отклика от воздушного объекта. Для сравнения разработанного метода с известными методами проведены эксперименты по приему отраженных от реальных воздушных объектов сигналов.

Представлены результаты данных экспериментально-теоретических исследований, подтверждающие эффективность разработанного метода. Предлагаемый к реализации метод может быть использован при разработке пассивно-активных комплексов радиолокации по сигналам сторонних источников для обнаружения и сопровождения целей с выдачей координатно-трассовой информации.

**Ключевые слова:** *пассивно-активная радиолокация, параметры траектории движения, доплеровская частота, обнаружение воздушных объектов.*

В последние годы большое развитие получили пассивно-активные средства радиолокации воздушных надводных и наземных целей. Неоспоримые преимущества, которыми обладают данные системы по сравнению с активными средствами, предопределили появление в ряде иностранных государств мобильных и стационарных приемных модулей и постов приема и обработки, отраженных от подвижных объектов сигналов сторонних источников подсвета, таких как вышки аналогового и цифрового телевидения и радиовещания, а также базовые станции сотовых систем связи и передачи данных различных стандартов [1, 2]. Работы по данному направлению в настоящее время ведутся и в отечественных организациях промышленности. Основной проблемой при решении задачи приема отраженного от объекта сигнала является влияние мощного основного сигнала от передающих устройств, работающих в непрерывном режиме, а также от неподвижных источников помех. Кроме этого, высокие требования к вычисли-

тельным ресурсам при реализации существующих методов не позволяют полностью использовать потенциал пассивных систем [3].

В связи с этим возникает актуальная научно-техническая задача разработки метода обнаружения и сопровождения воздушных объектов по отраженным радиосигналам сторонних источников. Метод востребован в условиях влияния мощного основного сигнала подсвета и при ограничениях на вычислительные ресурсы системы.

Общий принцип пассивной радиолокации основан на двухканальной когерентной обработке опорного сигнала, излучаемого передающей станцией, и целевого сигнала, отраженного от объекта [4–6].

В основе обработки, как и для большинства аналогичных зарубежных и отечественных пассивных систем, лежит классический подход свертки сигнала с выхода двух каналов (опорного и целевого) с возможностью измерения доплеровской частоты. Используется свойство комплексно-сопряженного умножения сигналов, представленных в цифровом

квадратурном виде. Основная формула обработки для дискретных цифровых сигналов записывается в виде

$$Z(m, n) = \frac{1}{N} \sum_{k=0}^{N-1} S_o^*(k) \cdot S_{\text{ц}}(k-n) \ell^{-j2\pi \frac{m}{N} k}, \quad (1)$$

где  $S_o$  – отсчеты сигналов опорного канала;  $S_{\text{ц}}$  – отсчеты сигналов целевого канала;  $N$  – количество накопленных дискретных отсчетов;  $m$  – номер канала доплеровской частоты;  $n$  – сдвиг данных целевого канала относительно опорного [7].

Решение о наличии отраженного сигнала в выборке  $N$  принимается на основе пороговой обработки двумерной функции  $Z(m, n)$ , вычисляемой для

$$m \in [0; m_{\max}]; n \in (0; n_{\max}], \quad (2)$$

где  $m_{\max}$  – максимальное значение частотного канала, определяемое максимально возможной частотой Доплера  $f_D$ ;  $n_{\max}$  – максимальное значение сдвига, определяемое максимально возможной временной задержкой сигнала  $\tau_3$ .

Выражение (1) можно разделить на две последовательно выполняемые операции: комплексно-сопряженное умножение и расчет дискретного преобразования Фурье для заданных частотных каналов.

Результатом операции комплексно-сопряженного умножения для идентичных сигналов, в том числе и широкополосных, но сдвинутых по частоте на  $f_D$ , является комплексный синусоидальный сигнал, частота которого равна  $f_D$ , а амплитуда – произведению амплитуд сигналов  $S_o$  и  $S_{\text{ц}}$ . Для доказательства воспользуемся следующим представлением комплексного сигнала во времени [8]:

$$z(n) = x(n) + jy(n). \quad (3)$$

При этом сигнал  $x(n)$  определяет действительную часть комплексного сигнала  $z(n)$ , то есть  $\text{Re}[z(n)] = x(n)$ , сигнал  $y(n)$  – его мнимую часть:  $\text{Im}[z(n)] = y(n)$ , а  $j$  – мнимая единица, равная  $(-1)^{1/2}$ .

В соответствии с [9] сигнал  $z(n)$  может быть сдвинут в спектральной области на величину  $f_D$  и преобразован в сигнал  $z_c(n)$ :

$$z_c(n) = x_c(n) + jy_c(n), \quad (4)$$

$$x_c = x(n) \cos(2\pi n W) - y(n) \sin(2\pi n W), \quad (5)$$

$$y_c(n) = y(n) \cos(2\pi n W) + x(n) \sin(2\pi n W), \quad (6)$$

где  $W = f_D / F_{\text{дискр}}$ ,  $F_{\text{дискр}}$  – частота дискретизации сигнала.

Тогда в результате операции комплексно-сопряженного умножения сигналов (3) и (4) получим новый сигнал  $z_{\text{рез}}(n)$ :

$$\begin{aligned} z_{\text{рез}}(n) &= z(n) \cdot z_c^*(n) = (x(n) + jy(n)) \times \\ &\times (x_c(n) - jy_c(n)) = \\ &= (x(n)x_c(n) + y(n)y_c(n) + \\ &+ (x_c(n)y(n) - x(n)y_c(n))j. \end{aligned} \quad (7)$$

При подстановке (5) и (6) в (7) получаем:

$$\begin{aligned} z_{\text{рез}}(n) &= (x^2(n) + y^2(n)) \cos(2\pi n W) - \\ &- j(x^2(n) + y^2(n)) \sin(2\pi n W). \end{aligned} \quad (8)$$

Выражение (8) представляет собой комплексно-сопряженное представление гармонического колебания амплитудой  $x^2(n) + y^2(n)$  и частотой  $W$ . Основным преимуществом операции комплексно-сопряженного умножения является сохранение гармонического закона, определяемого  $f_D$ , результирующего сигнала при фазовом и амплитудном рассогласовании идентичных сигналов, которое влияет только на амплитуду и начальную фазу сигнала, что инвариантно для процедуры дискретного преобразования Фурье.

Одним из основных преимуществ пассивно-активных систем с учетом того, что сигналы подсвета чаще всего работают в непрерывном режиме, является возможность длительного накопления сигнала [10]. При этом время накопления зависит только от размеров элемента разрешения в пространстве, определяемых диаграммами направленности антенн и шириной спектра сигнала. Так как диаграммы направленности передающих антенных систем имеют круговую форму по азимуту, а приемные системы, как правило, секторные, длительность накопления определяется только разрешением по дальности или шириной спектра принимаемого сигнала. При этом может обрабатываться как часть спектра, так и весь спектр сигнала подсвета.

Проведенные исследования показали эффективность многоканальной узкополосной (1–3 МГц) обработки для обнаружения целей и широкополосной одноканальной обработки для селекции и распознавания целей. Выбор оптимальных режимов обработки сигналов в пассивно-активных системах осуществляется на этапе проектирования системы в целом.

С учетом этого размер буфера данных накопления в зависимости от времени накопления может лежать в пределах от одного до нескольких десятков миллионов отсчетов. Обработка такого объема данных в квазиреальном масштабе времени не представляется возможной. Применение алгоритма быстрого преобразования Фурье вместо дискретного преобразования Фурье избыточно, поскольку возможные значения частоты Доплера лежат в начале спектра и занимают не более 5 % от всего диапазона частот.

Другим немаловажным фактором, определяющим качество пассивной локации, является присутствие опорного сигнала в целевом канале и отраженного сигнала в опорном канале, а также сигналов, переотраженных от различных объектов, что приводит к искажениям принимаемых сигналов.

Существующие методы, основанные на (1), не в полной мере решают указанные проблемы, поэтому не обеспечивают требуемую эффективность пассивной радиолокации.

С учетом этого необходимо разработать метод обнаружения и сопровождения воздушных объектов по отраженным радиосигналам сторонних ис-

точников, минимизирующий влияние сигнала подсвета и переотраженных сигналов от неподвижных источников и обеспечивающий требуемое соотношение сигнал/шум на входе решающего устройства при ограничении на вычислительные ресурсы системы.

### Структура метода решения задачи

Для решения поставленной задачи необходимо реализовать алгоритм компенсации опорного сигнала в целевом канале и снизить количество вычислительных операций.

В ряде работ [11] для решения задачи минимизации влияния сигнала подсвета предлагается восстанавливать опорный сигнал и, используя его как опорный, обрабатывать данные целевого канала методом адаптивной фильтрации. Указанный подход при всей его эффективности имеет ограниченное использование, поскольку реализуем только для цифровых передач, а коэффициенты адаптивных фильтров не являются универсальными для всех сигналов подсвета. Кроме этого, такой подход требует большого объема вычислений.

С учетом этого в работе для минимизации влияния сигнала подсвета предложено использовать алгоритм амплитудно-фазовой компенсации, а для сокращения требований к вычислительным ресурсам – алгоритм децимации сигналов.

Структура метода обнаружения и сопровождения воздушных объектов по отраженным радиосиг-

налам сторонних источников в пассивно-активных системах радиолокации представлена на рисунке 1.

В соответствии с данной структурой оцифрованные сигналы в квадратурном виде сохраняются в виде блоков данных размерами  $N$  для опорного канала и  $N+m_{\max}$  для целевого канала. Значение  $m_{\max}$  определяется с учетом возможной максимальной задержки отраженного от цели сигнала для выбранной частоты дискретизации. Далее последовательно со сдвигом в один отсчет на основе представленной выше свертки формируется двумерный массив данных, в котором по одной оси представлено распределение временной задержки принятого сигнала, а по второй – частоты Доплера.

На основе пороговой обработки двумерного массива с учетом заданной вероятности ложной тревоги принимается решение о наличии отраженного сигнала и вычисляются время задержки  $\tau_z$  и частота Доплера  $f_D$ .

Для сокращения количества элементарных вычислительных операций предложено после комплексно-сопряженного умножения с применением операции децимации прореживать отсчеты сигнала с учетом возможного значения частоты Доплера.

Рассмотрим процесс понижения частоты дискретизации сигнала  $z(n)$  в целое число раз, равное  $M$ . В этом случае новая частота дискретизации

$$f'_D = \frac{f_D}{M}. \quad (9)$$

Предположим, что сигнал  $z(n)$  характеризуется полной полосой частот; иными словами, его спектр

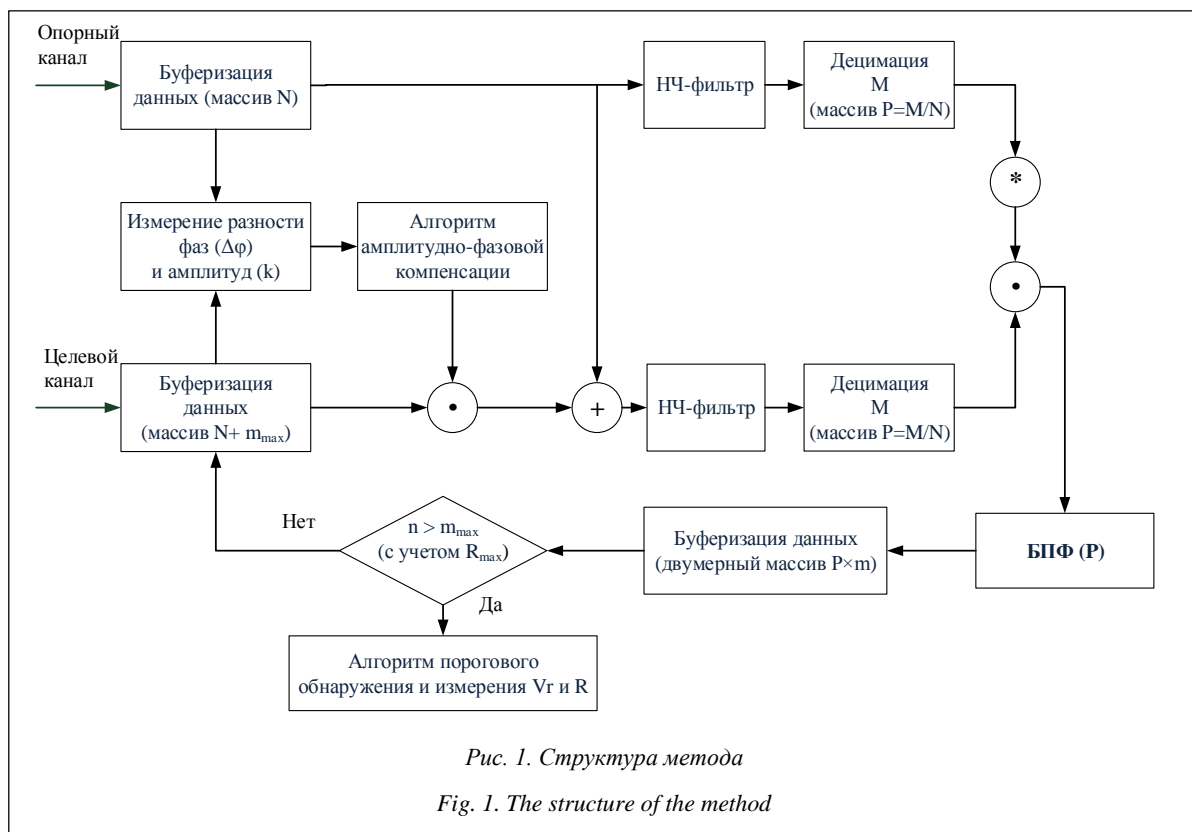


Рис. 1. Структура метода

Fig. 1. The structure of the method



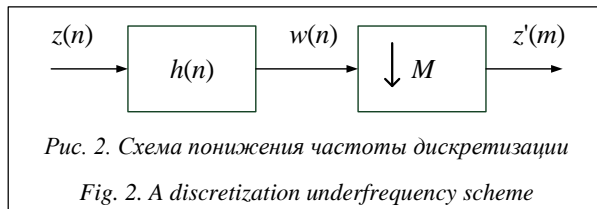
отличен от нуля на всех частотах в диапазоне  $-f_D/2 \leq f \leq f_D/2$ , то есть

$$|X(e^{j\omega})| \neq 0 \text{ при } |\omega| = \left| \frac{2\pi f}{f_D} \right| < \pi. \quad (10)$$

Согласно теореме отсчетов [9], для понижения частоты дискретизации и устранения эффекта наложения на более низкой частоте (НЧ) дискретизации необходимо пропустить сигнал  $z(n)$  через цифровой фильтр нижних частот, который в идеале должен иметь следующую частотную характеристику:

$$\tilde{H}(e^{j\omega}) = \begin{cases} 1 & \text{при } |\omega| = \left| \frac{2\pi f}{f'_D} \right| < \frac{\pi}{M}, \\ 0 & \text{при других значениях.} \end{cases} \quad (11)$$

Понижение частоты дискретизации достигается в данном случае посредством формирования последовательности  $z'(n)$ , которая получается в результате удаления каждого  $M$ -го отсчета из отфильтрованного выходного сигнала. Этот процесс графически представлен на рисунке 2.



Если отклик фильтра нижних частот на единичный отсчет обозначить через  $h(n)$ , то

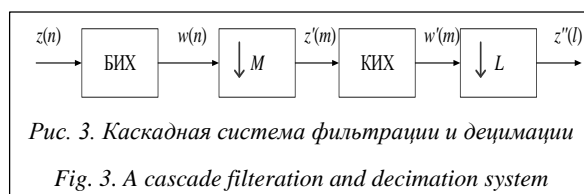
$$w(n) = \sum_{k=-\infty}^{\infty} h(k)z(n-k), \quad (12)$$

где  $w(n)$  – выход фильтра.

Выходной сигнал  $z'(m)$  будет иметь вид  $z'(m) = w(Mn)$ . (13)

Выходной сигнал  $z'(m)$  получается после выполнения операции во втором блоке (рис. 2). Этот блок называется компрессором частоты дискретизации и соответствует операции повторной выборки, описываемой выражением (13).

Реализация НЧ-фильтра зависит от требуемого коэффициента децимации. Так, для малых коэффициентов децимации (до 32) эффективно применение фильтров с конечной импульсной характеристикой (КИХ) из-за их стабильности и универсальности. Для более высоких коэффициентов децимации возможно применение каскадной системы фильтрации и децимации с последовательным включением фильтров с бесконечной импульсной характеристикой (БИХ) и КИХ (рис. 3).



Как будет показано далее, выполнение процедуры децимации приведет к многократному сокращению времени выполнения процедуры, а применение НЧ-фильтра улучшит соотношение сигнал/шум.

В разработанном методе для решения задачи минимизации влияния опорного сигнала и искажений от переотражений предложено использовать алгоритм амплитудно-фазовой коррекции как наиболее универсальный и менее требовательный к вычислительным ресурсам. Коррекции подвергается сигнал целевого канала на основе измерения разности амплитуд и фаз между целевым и опорным каналами с последующей аддитивной обработкой. Особенно эффективно его применение для приемных антенных систем с идентичными характеристиками.

На первом этапе по выборкам сигналов опорного и целевого каналов производится расчет коэффициента амплитудного рассогласования:

$$K_A = \frac{\sum_{k=0}^{N-1} \sqrt{x_o^2(k) + y_o^2(k)}}{\sum_{k=0}^{N-1} \sqrt{x_{\pi}^2(k) + y_{\pi}^2(k)}}. \quad (14)$$

Далее оценивается разность фаз сигналов опорного и целевого каналов:

$$\Delta\varphi = \arctg \frac{\sum_{k=0}^{N-1} x_o(k)x_{\pi}(k) + y_o(k)y_{\pi}(k)}{\sum_{k=0}^{N-1} y_o(k)x_{\pi}(k) - x_o(k)y_{\pi}(k)}. \quad (15)$$

Алгоритм амплитудно-фазовой коррекции представляет собой процедуру преобразования целевого канала, в соответствии с которой осуществляются домножение значений отсчетов на коэффициент амплитудного рассогласования  $K_A$  и сдвиг фазы сигнала целевого канала на величину  $\Delta\varphi + \pi$  с последующей операцией поразрядного сложения с сигналом опорного канала.

Выходным сигналом алгоритма амплитудно-фазовой коррекции является сигнал  $S'_{\pi}$ , который поступает на процедуру комплексно-сопряженного умножения с принятым или восстановленным опорным сигналом.

Ниже представлены выражения для расчета реальной  $x'(n)$  и мнимой  $y'(n)$  частей комплексного сигнала  $S'_{\pi}$ :

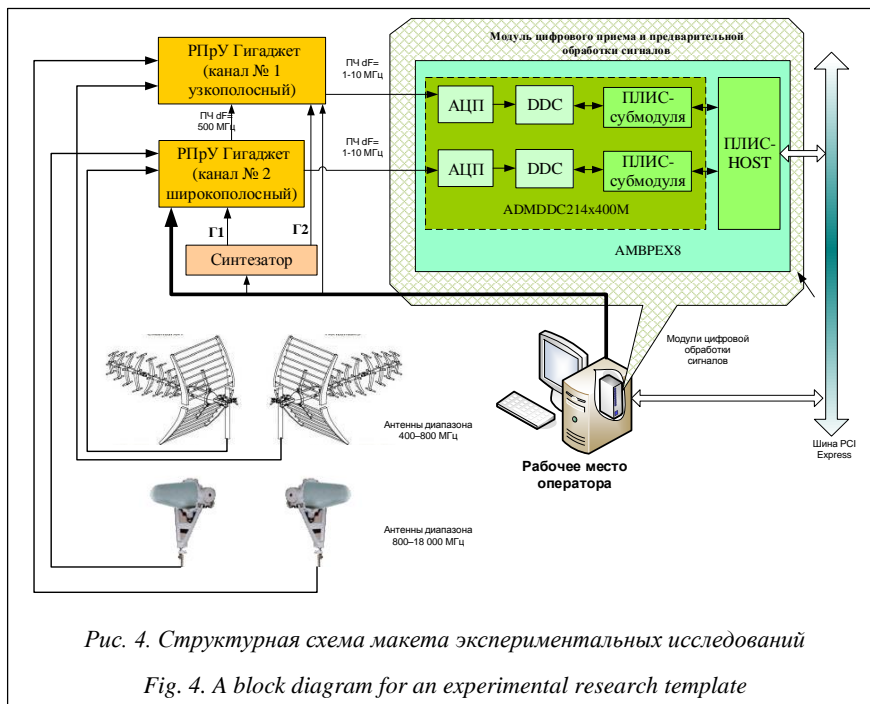
$$x'(n) = x(n)\cos(\Delta\varphi + \pi)K_A + y(n)\sin(\Delta\varphi + \pi)K_A, \quad (16)$$

$$y'(n) = y(n)\cos(\Delta\varphi + \pi)K_A - x(n)\sin(\Delta\varphi + \pi)K_A. \quad (17)$$

Данный алгоритм снижает влияние опорного сигнала, проникающего в целевой канал, повышает соотношение сигнал/шум и, как следствие, вероятность обнаружения цели.

### Результаты экспериментальных исследований

Для оценки эффективности применения разработанного метода были проведены эксперимен-



тальные исследования по обнаружению и определению параметров воздушных целей по отраженным сигналам цифрового телевидения стандарта DVB-T2.

Двухканальная приемная система была реализована на базе радиоприемного устройства из состава подвижного автоматизированного комплекса исследования и анализа радиоизлучений и модуля цифровой обработки производства ЗАО «Инструментальные системы» (г. Москва). В качестве антенн использованы специализированные гражданские антенны дециметрового диапазона частот. Макет позволяет принимать и обрабатывать сигналы в двухканальном когерентном режиме в диапазоне частот 500–2500 МГц в полосе до 4 МГц в реальном масштабе времени и до 32 МГц в квази-реальном.

Структурная схема макета представлена на рисунке 4. Макет включает в себя две антенны ДЦМ-диапазона, двухканальный приемник Гигаджет, двухканальный субмодуль цифрового приема ADMDDC214x400M, модуль сопряжения AMBPEX8, ПЭВМ с программой приема и обработки сигналов.

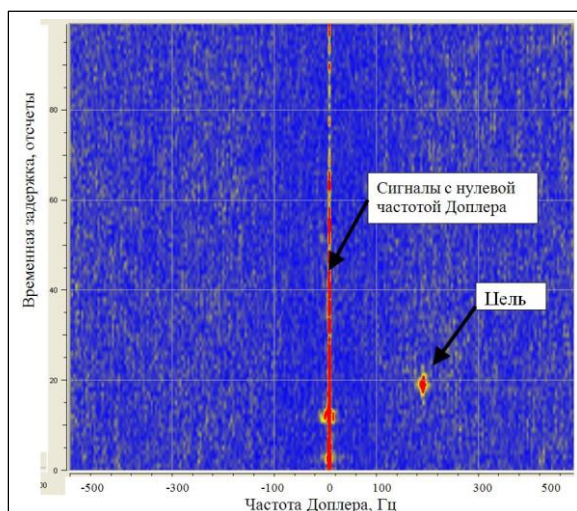
Эксперимент проводился в условиях городской застройки на 10-м этаже здания, в непосредственной близости от воздушного коридора захода на посадку военно-транспортной авиации. Расстояние до телевизионной вышки подсвета составляло около 4 км.

В качестве сигнала подсвета использовался канал цифрового телевидения стандарта DVB-T2, представляющий собой набор перекрывающихся по частоте узкополосных каналов с квадратурно-фазовой модуляцией несущей с общей шириной спектра сигнала 7,62 МГц.

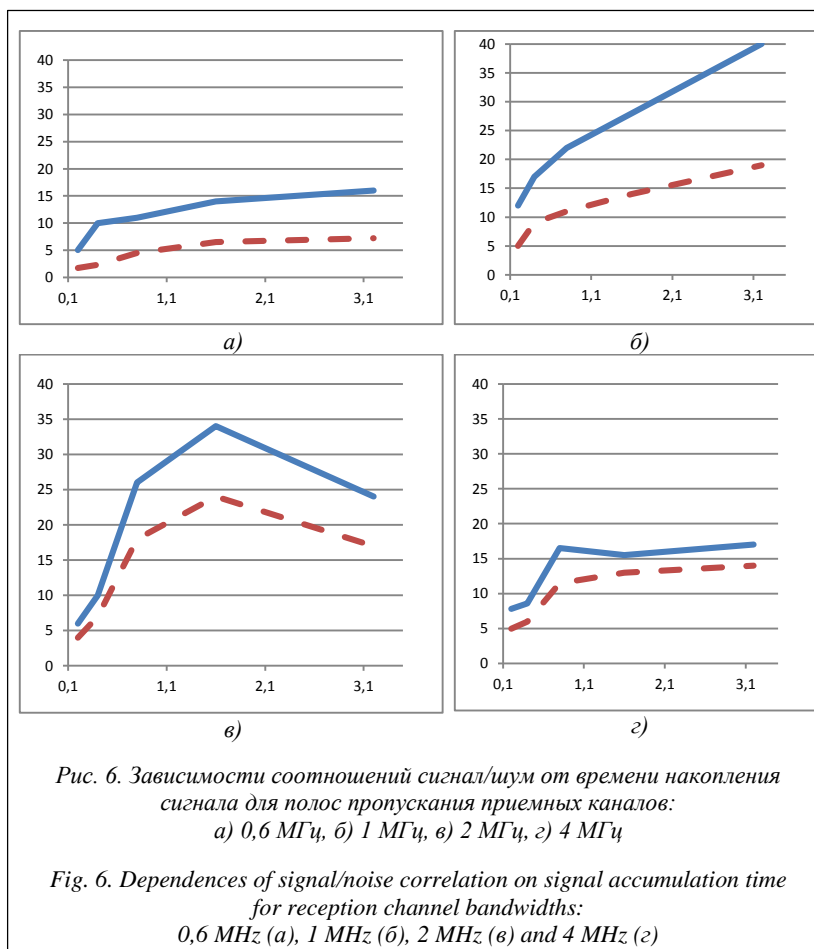
На рисунке 5 представлен результат двухканальной обработки в виде двумерного массива данных, в котором по вертикали отображается распределение по дальности до источника, а по горизонтали – по частоте Доплера. На данной реализации наблюдается отклик от цели, по которому автоматически измеряются относительная радиальная скорость цели  $V'_R$  и временная задержка прихода сигнала по сравнению с опорным  $\tau_3$ .

Линия в центральной части экранной области представляет собой свертку опорного сигнала, а также результат переотражения от стационарных объектов и искажения сигнала.

Для сравнительной оценки эффективности применения разработанного метода были проведены эксперименты по измерению соотношения сигнал/шум и времени выполнения операций для различных полос пропускания приемных каналов (0,6; 1; 2 и 4 МГц), а также с использованием алгоритмов децимации и амплитудно-фазовой коррекции и без. Измерения производились по однотипным воздушным объектам и для близких значений относительной радиальной скорости целей  $V'_R$  и времени задержки прихода сигнала по сравнению с опор-







ным  $\tau_3$ . Графики зависимости соотношений сигнал/шум от времени накопления сигнала для различных полос пропускания приемных каналов с использованием алгоритма амплитудно-фазовой коррекции и без представлены на рисунке 6. Из графиков видно, что соотношение сигнал/шум возрастает со временем накопления обрабатываемого сигнала, а применение алгоритма амплитудно-фазовой коррекции на 6–10 дБ повышает данное соотношение.

Для полос пропускания 1 и 2 МГц наблюдаются ярко выраженные максимумы при значениях времени накопления 3,2 и 1,6 секунды соответственно, что объясняется максимальной длительностью нахождения воздушного объекта в элементе разрешения системы с учетом условия проведения эксперимента.

Наиболее эффективно применение алгоритма компенсации для полос пропускания 2 и менее МГц, что объясняется сильными искажениями сигнала в сложных городских условиях приема.

Оценка времени выполнения единичной операции обработки сигналов на стандартном ПЭВМ с использованием алгоритмов децимации и амплитудно-фазовой коррекции и без них показывает, что использование алгоритма децимации в 2–3 раза сокращает время выполнения операции. При этом при проведении эксперимента использовался

фильтр с конечной импульсной характеристикой. Его замена на фильтр с бесконечной импульсной характеристикой или каскадная фильтрация позволит еще в несколько раз сократить время выполнения процедуры. Кроме этого, операции вычисления тригонометрических функций в (16) и (17) могут быть оптимизированы за счет использования их периодических свойств, что также уменьшит время выполнения процедур.

Таким образом, на основе проведенного экспериментально-теоретического исследования для всех полос пропускания (0,6; 1; 2 и 4 МГц) подтверждена возможность получения первичных радиолокационных отметок от воздушных целей по сигналам сторонних источников.

Реализованный в программе технического анализа сигналов метод обнаружения и сопровождения воздушных объектов по отраженным радиосигналам сторонних источников обеспечивает обнаружение воздушных объектов, находящихся в сек-

торе обзора антенны, и оценку (в каждый момент времени) относительной радиальной скорости цели и времени задержки прихода отраженного сигнала.

Кроме того, в ходе проведенного исследования определены факторы, влияющие на эффективность пассивно-активных систем, в частности, уровень подавления опорного сигнала в целевом канале и отраженного сигнала в основном канале, а также искажения сигнала в опорном канале за счет перелотражений.

Метод может быть использован при разработке пассивно-активных комплексов радиолокации по сигналам сторонних источников для обнаружения и сопровождения целей с выдачей координатно-трассовой информации.

#### Литература

1. Аношкин И.М. Зарубежные многопозиционные радиолокационные системы скрытного контроля воздушного пространства // Наука и военная безопасность. 2007. № 1. С. 28–33.
2. Алифанов Р.Н., Стародубцев П.А., Дементьев С.Г., Дементьев С.С. Системы общедоступного радиовещательного передатчика для обнаружения и сопровождения цели // Наукосведение. 2014. № 6. URL: <http://naukovedenie.ru/PDF/74TVN614.pdf> (дата обращения: 16.05.2016).
3. Willis N.J. Bistatic radars and their third resurgence: passive coherent location. IEEE Radar Conf., Long Beach, USA, 24 April 2002.
4. Кузьмин С.З. Цифровая радиолокация. Введение в теорию. Киев: Изд-во КВЦ, 2000. 428 с.

5. Черняк В.С., Заславский Л.П., Осипов Л.В. Многопозиционные радиолокационные станции и системы // Зарубежная радиоэлектроника. 1987. № 1. С. 9–69.
6. Черняк В.С. Многопозиционная радиолокация. М.: Радио и связь, 1993. 416 с.
7. Christiansen J.M. DVB-T based passive bistatic radar. Master's thesis, 2009, 76 p.
8. Голд Б.Е., Рейдер Ч.А. Цифровая обработка сигналов. М.: Сов. радио и связь, 1978. 368 с.
9. Гольденберг Л.М., Матюшкин Б.Д., Поляк М.Н. Циф-

ровая обработка сигналов: справочник. М.: Радио и связь, 1985. 312 с.

10. Kyovtorov V., Vespe M., Fortuny Guasch J., Tarchi D., Giuliani R., Broussolle J. Passive automatic identification system for maritime surveillance – JRC, 2012. URL: <http://publications.jrc.ec.europa.eu/repository/bitstream/JRC68405/lbna25169enn.pdf> (дата обращения: 16.05.2016).

11. Фам Х.Т. Полуактивная радиолокационная система с использованием сигналов DVB-T2: дис. ... канд. технич. наук. СПбГУ «ЛЭТИ», 2015. 150 с.

DOI: 10.15827/0236-235X.115.168-174

Received 24.06.16

## METHOD OF DETECTING AND TRACKING AERIAL OBJECTS BY REFLECTED RADIO SIGNALS FROM THIRD-PARTY SOURCES IN A PASSIVE-ACTIVE RADAR SYSTEM

(Acknowledgements. This work was financially supported by grant of the President of the Russian Federation no. HIII-6831.2016.8)

<sup>1</sup>Batchev S.A., Ph.D. (Engineering), Deputy Head of Chair; <sup>2</sup>Zaycev A.G., Ph.D. (Engineering), Leading Researcher;

<sup>2</sup>Talalaev A.B., Dr.Sc. (Engineering), Professor, Director General;

<sup>2</sup>Timakov D.A., Ph.D. (Engineering), Deputy Head of Department, timakov\_da@rambler.ru

<sup>1</sup>Cherepovets Higher Military Engineering School of Radioelectronics,

Sovetsky Ave. 126, Cherepovets, 162622, Russian Federation);

<sup>2</sup>Radiotechnical and Information Systems Aerospace Defence,

Z. Konoplyannikova St. 89/1, Tver, 170041, Russian Federation

**Abstract.** The article assesses the level of development of modern passive-active radar systems within the performed research. It shows that the use of such systems increases the airspace control efficiency and reduces its costs. The paper describes the developed method of detecting and tracking aerial objects by reflected radio signals from third-party sources in passive-active radar systems. The method is based on measuring the time delay between reference and reflected signals and the relative radial velocity of the controlled object. In contrast to known approaches, this method considers the factors that affect the quality of detection and determination of air targets' trajectory characteristics and allows expanding the zone of air space control. The research proved that the main factors that reduce the efficiency of the system are: the level of suppression of the reference signal in the target channel and the reflected signal in the main channel; changing the shape of the reflected signal; the power level of reflected signals from local stationary objects. The authors suggest the ways to reduce the impact of these factors based on algorithms for amplitude-phase correction signals in the receiving channels and decimation of digital streams.

The application of these approaches reduces the procedure execution time and improves the signal/noise ratio of radar response from the aircraft object. The proposed method can be used in development of passive-active complexes of radiolocation signals of third-party sources for detection and tracking targets with coordinate and en-route information. To compare the developed method with other known methods the authors made the experiments with reception of signals reflected from real aerial objects. The paper presents the results of these experimental and theoretical studies that prove the effectiveness of the developed method.

**Keywords:** : passive-active radar, parameters of a movement trajectory, Doppler frequency, aerial objects detection.

### References

1. Anoshkin I.M. Foreign multiposition radar systems for mid-air covert breach control. *Nauka i voennaya bezopasnost* [Science and Military Security]. 2007, no. 1, pp. 28–33 (in Russ.).
2. Alifanov R.N., Starodubtsev P.A., Dementev S.G., Dementev S.S. Systems of a public domain broadcast transmitter to target detecting and tracking. *Naukovedenie*. 2014, no. 6. Available at: <http://naukovedenie.ru/PDF/74TVN614.pdf> (accessed June 23, 2016).
3. Willis N.J. Bistatic radars and their third resurgence: passive coherent location. *IEEE Radar Conf.* Long Beach, USA, 2002.
4. Kuzmin S.Z. *Tsifrovaya radiolokatsiya. Vvedenie v teoriyu* [Digital Radio Detecting and Ranging. Introduction to Theory]. Kiev, KVITs Publ., 2000, 428 p.
5. Chernyak V.S., Zaslavsky L.P., Osipov L.V. Multiposition radar stations and systems. *Zarubezhnaya radioelektronika* [Telecommunications and Radio Engineering]. 1987, no. 1, pp. 9–69 (in Russ.).
6. Chernyak V.S. *Mnogopozitsionnaya radiolokatsiya* [Multiposition Radio Detecting and Ranging]. Moscow, Radio i svyaz Publ., 1993.
7. Christiansen J.M. *DVB-T based Passive Bistatic Radar*. Master's Thesis. Norwegian Univ. of Science and Technology Publ., 2009, 76 p.
8. Gold B., Rader C.M. *Digital Processing of Signals*. NY, McGraw-Hill Publ., 1969 (Russ.ed.: Moscow, Radio i svyaz Publ., 1988).
9. Goldenberg L.M., Matyushkin B.D., Polyak M.N. *Tsifrovaya obrabotka signalov* [Digital Processing of Signals]. Reference book. Moscow, Radio i svyaz Publ., 1985.
10. Kyovtorov V., Vespe M., Fortuny Guasch J., Tarchi D., Giuliani R., Broussolle J. *Passive Automatic Identification System for Maritime Surveillance – JRC*. 2012. Available at: <http://publications.jrc.ec.europa.eu/repository/bitstream/JRC68405/lbna25169enn.pdf> (accessed June 23, 2016).
11. Fam Kh.T. *Poluaktivnaya radiolokatsionnaya sistema s ispolzovaniem signalov DVB-T2* [semi-active radar system using DVB-T2 signals]. Ph.D. Thesis, St. Petersburg, LETI Publ., 2015.

УДК 004.932.2

Дата подачи статьи: 28.06.16

DOI: 10.15827/0236-235X.115.175-180

## **ПРОГРАММНЫЙ КОМПЛЕКС ПРОСТРАНСТВЕННОЙ НАВИГАЦИИ И МОНИТОРИНГА НА ОСНОВЕ АЛГОРИТМА ВИЗУАЛЬНОЙ ОДОМЕТРИИ**

(Работа выполнена при поддержке РФФИ, гранты №№ 16-01-00342, 16-29-04268,  
и гранта Президента РФ № НШ-6831.2016.8)

**А.А. Кочкаров**, к.ф.-м.н., зам. директора Научно-тематического центра, AKochkarov@oaorti.ru  
(Компания «РТИ», ул. 8 Марта, 10, стр. 1, г. Москва, 127083, Россия;  
Финансовый университет при Правительстве Российской Федерации,  
Ленинградский просп., 49, г. Москва, 125993, Россия);

**И.А. Калинов**, аспирант, магистрант, специалист, tej@frtk.ru  
(Московский физико-технический институт (государственный университет),  
Институтский пер., 9, г. Долгопрудный, 141700, Россия;  
Сколковский институт науки и технологий,  
дер. Сколково, ул. Новая, 100, 143025, Россия;  
Компания «РТИ», ул. 8 Марта, 10, стр. 1, г. Москва, 127083, Россия)

В настоящей работе представлен программный комплекс беспилотного летательного аппарата мультироторного типа, ориентированный на мониторинг пространства и поиск человеческих лиц. Программный комплекс разработан на основе алгоритма полупрямой визуальной одометрии. Полупрямой подход избавляет от трудностей, связанных с ресурсоемким извлечением контрольных точек, и в то же время представляет собой надежный метод с точки зрения оценки движения. Алгоритм оперирует непосредственно интенсивностью пикселей, что в результате дает субпиксельную точность при обработке больших объемов кадров. Вероятностный метод отображения, в явном виде распознающий аномальные измерения для оценки трехмерных координат точки, приводит к уменьшению количества аномальных измерений и повышению качества измерений, так и их надежности. Такая оценка передвижения кадров повышает надежность в случаях монотонной, повторяющейся или высокочастотной текстуры. Это алгоритм неявно прямой визуальной одометрии. В работе приведено сравнение данного алгоритма с другим популярным алгоритмом визуальной одометрии. Кроме того, рассмотрены различные подходы к патрулированию пространства и приведены результаты компьютерной реализации этих подходов.

**Ключевые слова:** визуальная одометрия, программный комплекс, мониторинг, БПЛА, патрулирование, квадрокоптер, SLAM, PTAM, фильтры глубины, алгоритм.

Малые беспилотные летательные аппараты (БПЛА) играют важную роль в задачах ликвидации последствий различного рода катастроф, стихийных бедствий, производственного контроля и охраны окружающей среды [1–3]. Однако для таких задач недостаточно навигации на основе только информации, полученной при помощи спутниковых систем: полностью автономная система навигации малого БПЛА должна опираться на локализованную систему. Для получения такой системы с минимальным весом и экономией энергии нужно комбинировать одну камеру, смотрящую вниз, и инерциальное измерительное устройство [4]. Такой набор инструментов позволяет малому БПЛА полностью автономно преодолеть заданный маршрут, при этом попутно сняв необходимые показания. На данный момент большая часть систем монокулярной визуальной одометрии для малых БПЛА использует подходы RGB-D [5] и стереосистемы SLAM [6], однако прямые методы, основанные на минимизации фотометрической ошибки, обретают все большую популярность.

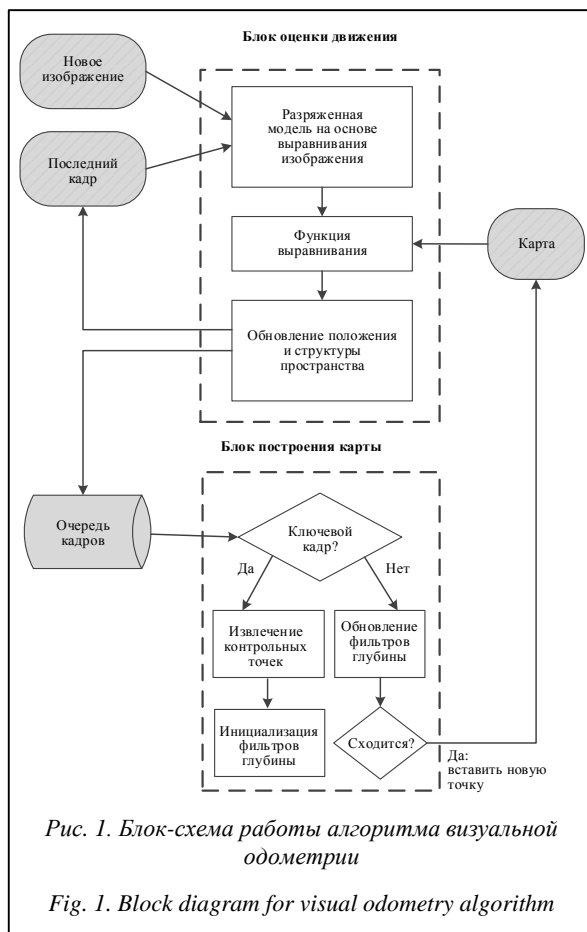
В настоящей работе описан алгоритм полупрямой визуальной одометрии. Он сочетает в себе факторы успеха метода на основе обнаружения контрольных точек, их параллельного отслеживания, составления карты и выбора ключевого кадра.

При всех описанных преимуществах точность и скорость работы алгоритма не хуже, чем у прямых методов. Высокая частота кадров при визуальной одометрии для малого БПЛА повышает надежность и увеличивает скорость маневров в полете.

### **Описание работы системы визуальной одометрии**

На рисунке 1 представлена блок-схема алгоритма *неявно прямой визуальной одометрии* (НПВО). Он использует два параллельных потока: один для оценки передвижения камеры, другой для отображения карты и среды мониторинга в режиме реального времени. Разделение на два потока позволяет расширять карту пространства независимо от отслеживания передвижения, занимающего определенное время.

Поток оценки движения реализует предложенный полупрямой метод оценки относительного сдвига камеры. На первом шаге происходит инициализация позиции камеры с помощью разреженной модели изображения, далее идет выравнивание позиции камеры относительно предыдущего кадра, это происходит через сведение фотометрической ошибки к минимуму, что соответствует прогнозируемому трехмерному местоположению одних то-



чек на разных кадрах. Двухмерные координаты уточняются на следующем этапе через выравнивание характеристик наборов точек. Таким образом, оценка движения изначально находится путем уточнения позиции и структуры пространства через сведение к минимуму ошибки перепроецирования на предыдущем шаге выравнивания по контрольным точкам.

Во втором потоке (отображения карты) вероятностный фильтр глубины инициализируется для каждой контрольной точки на плоскости с целью получения ее третьей координаты. Фильтр глубины инициализируется заново, когда выбран новый ключевой кадр и найдено несколько соответствий трехмерных координат предыдущих контрольных точек с двухмерными на новом ключевом кадре. На каждом следующем кадре обновляются фильтры глубины. Со временем неточность фильтра глубины становится совсем малой, тогда новая контрольная точка с тремя координатами сразу проставляется на карту и используется для оценки движения.

### Сравнение работы алгоритмов визуальной одометрии

Большинство алгоритмов монокулярной визуальной одометрии для малых БПЛА опираются на

технологии параллельного отслеживания и построения карты (PTAM – Parallel Tracking and Mapping) [7]. PTAM, в свою очередь, базируется на методе одновременной локализации и построения карты SLAM, который обеспечивает надежность за счет отслеживания и отображения сотен контрольных точек. Он работает в режиме реального времени и одновременно выполняет задачи отображения пунктов назначения и оценки передвижения, опираясь на эффективную корректировку, основанную на обработке изображений с разных углов обзора. Тем не менее, PTAM был разработан для приложений дополненной реальности в небольших пространствах, но в нескольких модификациях, например, при ограничении числа кадров, что обеспечивает его полноценную работу в условиях пространств с застройкой городского типа [2].

Предлагаемый алгоритм полупрямой визуальной одометрии использует функцию нахождения контрольных точек. При этом нахождение контрольных точек является неявным результатом прямой оценки движения, нежели явного нахождения контрольных точек и сопоставления их положения в каждом кадре. Таким образом, извлечение контрольных точек требуется только тогда, когда ключевой кадр выбран для инициализации новых фильтров глубины (трехмерных точек) (см. рис. 1). Скорость работы алгоритма существенно увеличивается, когда контрольные точки извлекаются не во всех кадрах, а только в ключевых. Увеличивается точность и при использовании субпиксельной функции извлечения контрольных точек. В отличие от прямых методов используется много небольших участков, а не несколько больших плоских участков, что положительно сказывается на скорости работы алгоритма и его надежности. Предлагаемая разреженная модель изображения для оценки и выравнивания движения связана с плотной моделью изображения, однако далее будет показано, что разреженной информации о глубине достаточно, чтобы получить грубую оценку движения и функции поиска контрольных точек. Как только контрольные точки и первоначальное положение камеры будут найдены, алгоритм станет использовать лишь контрольные точки, тем самым объясняется его название «полупрямой». Этот прием позволяет быстро устанавливать рамки для обработки в новом изображении.

Байесовский фильтр, который в явном виде убирает ошибочные измерения, тоже оценивает глубину и расположение контрольных точек. Трехмерная точка наносится на карту только тогда, когда сопутствующий фильтр глубины дал сходящийся, что требует многократных измерений.

Результатом работы является карта с контрольными трехмерными точками, чья надежность была проверена. Научная новизна данной работы в изображении полупрямого алгоритма визуальной одометрии, который работает быстрее и надежнее

(точнее позиционирует БПЛА и не наносит на карту ошибочные точки) установленных на современных малых БПЛА, а также в интеграции вероятностного отображения точек на карте. Алгоритм устойчив к ошибочным измерениям.

При загрузке алгоритм оценивает начальную позицию по первым двум ключевым кадрам и начальной карте – локальному способу позиционирования БПЛА в пространстве. Изначально предполагается локально плоская схема пространства и оценивается начальное проективное преобразование. Исходная карта при этом определяется с первых двух углов обзора методом триангуляции.

Чтобы справиться с большими передвижениями, применяется выравнивание разреженной модели изображения по схеме от большого к малому по принципу пирамиды. Изображение разделяется на два образца для создания пятиуровневой пирамиды изображения. Затем невязки интенсивности оптимизируются на первом уровне пирамиды изображения до сходимости. Впоследствии оптимизация происходит и на следующих уровнях. Для экономии времени обработка остановки оптимизации происходит после сходимости на третьем уровне: на этом этапе оценка достаточна для инициализации функции выравнивания.

Исходя из соображений эффективности алгоритм сохраняет фиксированное число ключевых кадров на карте, которые затем используются в качестве эталонных для функции выравнивания по контрольным точкам и измельчения структуры. Ключевой кадр выбирается, если евклидово расстояние нового кадра по отношению ко всем ключевым кадрам превышает 12 % от средней глубины сцены пространства; когда новый ключевой кадр помещается на карту, удаляется наиболее дальний от текущей позиции ключевой кадр.

В блоке построения карты изображение делится на клетки, например  $30 \times 30$ . Новый фильтр глубины инициализируется в углу, который имеет максимальное значение рейтинга Ши-Томаси [8], если там еще нет соответствия двумерной и трехмерной карт. Это приводит к равномерному распределению контрольных точек по изображению. Эта же сетка используется для перепроецирования карты перед выравниванием. Стоит обратить внимание, что углы с максимальным значением рейтинга Ши-Томаси извлекаются на каждом уровне пирамиды изображения для нахождения наилучших углов независимо от масштаба.

Для сравнения результатов работы использовался модифицированный алгоритм PTAM [9]. Это связано с тем, что оригинальный алгоритм недостаточно надежен в высокочастотных повторяющихся текстурах, модифицированная версия справляется с этой задачей лучше. Алгоритм НПВО показал высокую точность и эффективность, частично это объясняется применяемым фильтром глубины, который отсеивает большую часть ошибочных изме-



Рис. 2. Примеры работы разработанного алгоритма в различных условиях

Fig. 2. Examples of work of the developed algorithm in different conditions

рений и положений трехмерных точек. Также данный фильтр обеспечивает высокую надежность и точность при монотонной, но высокочастотной текстуре (трава, асфальт) в силу высокой частоты обновления кадров. На рисунке 2 показана работа алгоритма как в обычных условиях, так и в условиях высокочастотной монотонной структуры.

На рисунке 3 приведено сравнение работы алгоритмов НПВО и PTAM. PTAM дает очень много ошибочных измерений, что может привести к невыполнению миссии. Стоит заметить, что именно этот алгоритм сейчас применяется на большей части малых БПЛА, таким образом, перед алгоритмом НПВО открываются широкие перспективы применения.

Эксперименты проводились на наборах данных, получаемых с камеры, прикрепленной к малому БПЛА и смотрящей вниз. Видео с камеры обрабатывалось на одноплатном компьютере Odroid-U2 [10], который был установлен непосредственно на БПЛА, а также транслировалось на внешний компьютер для обработки. Стоит заметить, что в результате эксперимента алгоритм работал с двумя разными режимами настроек. На одноплатном компьютере были выставлены настройки, позволяющие уменьшить скорость работы, на внешнем компьютере настройки были ориентированы на вы-

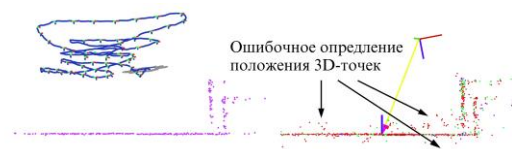


Рис. 3. Сравнение работы НПВО и PTAM

Fig. 3. Comparison of the work of developed algorithm and PTAM



сокую точность. Параметры настройки даны в таблице 1. Для дальнейшего изложения один режим назовем быстрым, другой – точным.

Таблица 1

## Характеристики двух режимов работы

Table 1

## Characteristics of the two operating modes

Параметры настройки	Быстрый	Точный
Максимальное количество контрольных точек в изображении	120	200
Максимальное количество ключевых кадров в изображении	10	50
Локальная оценка с разных углов обзора	Нет	Да

Следующим этапом сравнения были тесты на точность. Точность оценивалась на наборе данных, которые берут начало на земле, имеют длину траектории порядка 100 метров и среднюю высоту

1,5 метра. На рисунке 4 приведено сравнение отношения суммарной ошибки позиционирования к пройденному пути, а на рисунке 5 – сравнение ошибок определения углов связанной системы координат.

В целом обе версии НПВО показали большую точность по сравнению с модифицированным алгоритмом РТАМ. Такой результат обусловлен в основном тем, что РТАМ не извлекает контрольные точки на уровне самого высокого разрешения в пирамиде изображения, а это неизбежно вызывает потерю точности. В работе используется камера с невысоким разрешением, что, в свою очередь, приводит к слишком большому числу ошибочных измерений при работе алгоритма РТАМ в условиях высокочастотной повторяющейся текстуры, предложенный алгоритм эффективно справляется с этой проблемой, используя фильтры глубины. Та-

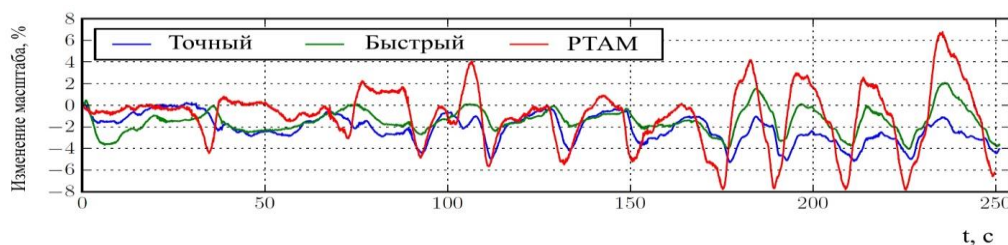


Рис. 4. Отношения суммарной ошибки позиционирования к пройденному пути

Fig. 4. Relationship of the total position error to the distance

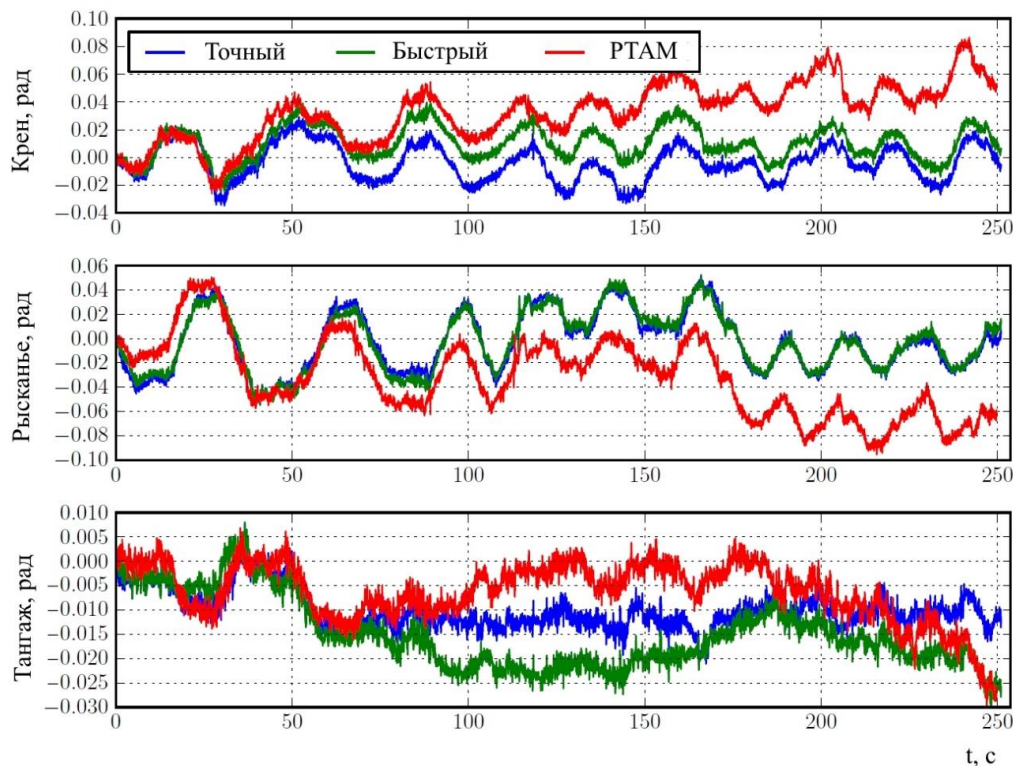


Рис. 5. Сравнение ошибок определения углов связанной системы координат

Fig. 5. Comparison of errors in the determination of angles in related coordinate system

кой подход сокращает число ошибочных измерений десятикратно.

Разница в точности быстрого и точного режимов работы алгоритма не является существенной, быстрый режим достаточно точно определяет позицию БПЛА, что позволяет говорить о его применимости на одноплатном компьютере, установленном на борту БПЛА. Это, в свою очередь, позволяет создать полностью автономный БПЛА.

Сравнение скорости работы алгоритмов показало, что ноутбук способен обрабатывать более 300 кадров в секунду, в то время как на одноплатном компьютере частота работы алгоритма составляет 55 кадров в секунду. Результаты работы алгоритма РТАМ на тех же устройствах составляют 91 и 27 кадров в секунду соответственно. Стоит отметить, что для работы алгоритма НПВО был выбран быстрый режим.

### Методы патрулирования

Для решения задач патрулирования изначально нужно выбрать метод патрулирования и оптимальную высоту. Для нахождения оптимальной высоты была проведена серия опытов по распознаванию лица, напечатанного в реальном масштабе с разных высот. Экспериментально было установлено, что оптимальная высота для детектирования лежит в промежутке от 10 до 20 метров; именно в этом интервале высот детектирование давало 100 %-ный результат. Высота же может меняться в этом промежутке в зависимости от целей миссии и, как следствие, скорости полета.

Когда оптимальная высота выбрана (в нашем случае это 13 метров), можно перейти к расчету площади замета за один пролет. В данном эксперименте использовалась камера с углом обзора по горизонтали в 60°; это подразумевает, что треугольник является равносторонним. Тогда его сторона  $a$  может быть вычислена по формуле  $a = \frac{2h}{\sqrt{3}}$ , при вы-

соте 13 метров  $a=15$  метров. То есть за пролет длиной 100 метров будет замечено  $1\,500^2$  метров. Существуют два принципиально разных подхода по патрулированию территории на БПЛА квадратного типа. Первый условно назовем «змейкой», второй – «спиралью» (рис. 6).

Эти подходы могут применяться в зависимости от различных условий, например, в зависимости от площади и фигуры территории для мониторинга. Для сравнения времени мониторинга при помощи этих двух методов возьмем тестовую территорию 100×100 метров. Вполне понятно, как ее замести «змейкой», если же пользоваться техникой «спирали», то существуют два варианта: квадрат 100×100 вписан в последнюю окружность или последняя окружность вписана в квадрат 100×100.

Результаты сравнения приведены в таблице 2. Как видно из таблицы, наиболее быстрым является

метод «спираль 2», но стоит отметить, что при нем не замечаются края квадрата, так как последняя концентрическая окружность вписана в квадрат. Метод «спираль 1», наоборот, покрывает большую площадь территории, чем необходимо. Оптимальным для площади такой фигуры является замет «змейкой», он дает 100 %-ный замет территории с небольшим увеличением времени по сравнению со «спиралью 2». Стоит отметить, что при мониторинге территории сложной фигуры наиболее оптимально использовать комбинацию этих методов и несколько БПЛА.

Таблица 2

### Сравнение результатов патрулирования территории различными методами

Table 2

#### Comparison of the results of patrolling areas by different methods

Параметр	«Змейка»	«Спираль 1»	«Спираль 2»
Скорость БПЛА	3 м/с~11 км/ч	3 м/с~11 км/ч	3 м/с~11 км/ч
Общая длина траектории	805 м	1 358 м	756 м
Площадь покрытия	10 500 м <sup>2</sup>	15 700 м <sup>2</sup>	7 850 м <sup>2</sup>
Время выполнения миссии	4,47 мин.	7,54 мин.	4,2 мин.

Таким образом, в данной работе предложен программный комплекс для мониторинга пространств на основе алгоритма полупрямой визуальной одометрии, который работает быстрее и точнее текущих алгоритмов. Выигрыш в скорости работы алгоритма получается из-за отсутствия функции извлечения контрольных точек на этапе обработки передвижения в отличие от прямых методов, основанных на интенсивности изображения. Алгоритм особенно полезен для одноплатных компьютеров, которые можно установить на борт БПЛА, так как на таких устройствах он может работать со скоростью более 50 кадров в секунду, что позволяет сделать БПЛА полностью автономным. Высокая частота обновления кадров при оценке движения в сочетании с устойчивостью фильтра ошибочных измерений обеспечивают повышенную надеж-

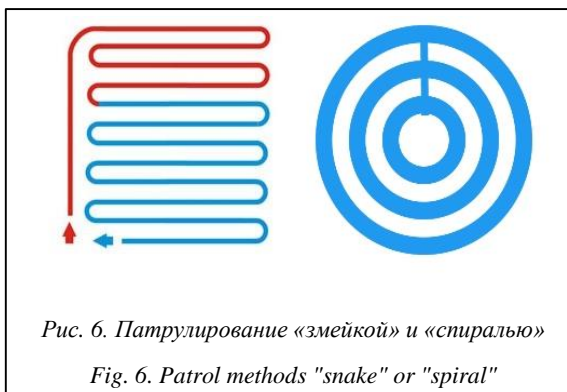


Рис. 6. Патрулирование «змейкой» и «спиралью»

Fig. 6. Patrol methods "snake" or "spiral"



ность в пространствах с высокочастотной монотонной структурой. Также были рассмотрены методы патрулирования пространства и проведено сравнение скорости их работы.

### Литература

1. Каримов А.Х. Цели и задачи, решаемые беспилотными авиационными комплексами нового поколения // Труды МАИ. Вып. № 47. 2011. С. 1–9; URL: <http://www.mai.ru/science/trudy/published.php?ID=26767> (дата обращения: 25.05.2016).
2. Калинов И.А. Создание малого автономного беспилотного летательного аппарата на базе мультиротора с использованием алгоритма полупрямой визуальной одометрии // Интеллектуальные системы, управление и мехатроника-2016: матер. Всерос. научн.-технич. конф. Севастополь: Изд-во СевГУ, 2016. С. 150–154.
3. Малинецкий Г.Г., Кочкаров А.А. Будущее российского оружия и междисциплинарные подходы // Интеллект и технологии. 2014. № 1 (7). С. 48–51.

4. Forster C., Lynen S., Kneip L., and Scaramuzza D. Collaborative monocular SLAM with multiple micro aerial vehicles. *Proc. IEEE/RSJ Int. Conf. IROS*, 2013, pp. 3962–3970.
5. Kerl C., Sturm J., and Cremers D. Robust odometry estimation for RGB-D cameras. *Proc. IEEE ICRA*, 2013, pp. 3748–3754.
6. Meilland M. and Comport A.I. On unifying key-frame and voxelbased dense visual SLAM at large scales. *Proc. IEEE/RSJ Int. Conf. on IROS*, 2013, pp. 3677–3683.
7. Klein G. and Murray D. Parallel tracking and mapping for small AR workspaces. *IEEE and ACM Int. Sympos. on Mixed and Augmented Reality*, Nov. 2007, pp. 1–10.
8. Wu M., Ramakrishnan N., Lam S.-K., and Srikanthan T. Low-complexity pruning for accelerating corner detection. *Proc. Conf. ISCAS-2012, IEEE Intern. Sympos.*, 2012, pp. 1684–1687.
9. Weiss S., Achtelek M.W., Lynen S., Achtelek M.C., Kneip L., Chli M., and Siegwart R. Monocular vision for long-term micro aerial vehicle state estimation: a compendium. *Journ. of FieldRobotics*, 2013, vol. 30, no. 5.
10. Odroid U2 user manual. URL: <https://astataruf.files.wordpress.com/2015/11/odroid-x2-user-manual.pdf> (дата обращения: 25.05.2016).

DOI: 10.15827/0236-235X.115.175-180

Received 28.06.16

## SOFTWARE FOR SPATIAL NAVIGATION AND MONITORING BASED ON VISUAL ODOMETRY ALGORITHM

(Acknowledgements. This work was financially supported by RFBR, grants no. 16-01-00342, 16-29-04268, and the grant of the President of the Russian Federation no. HHI-6831.2016.8)

<sup>1,2</sup>Kochkarov A.A., Ph.D. (Engineering), Deputy Director R&D centre, [AKochkarov@oaorti.ru](mailto:AKochkarov@oaorti.ru)

<sup>1,3,4</sup>Kalinov I.A., Professional Employee, Postgraduate Student, Graduate Student

<sup>1</sup>OJSC "RTI", March 8 St. 10, Bld. 1, Moscow, 127083

<sup>2</sup>Financial University under the Government of the Russian Federation, Leningradskiy Ave. 49, Moscow, 125993, Russian Federation

<sup>3</sup>Moscow Institute of Physics and Technology, Institutsky Line 9, Dolgoprudny, 141700, Russian Federation

<sup>4</sup>Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Bld. 3, Moscow, 143026, Russian Federation

**Abstract.** The paper presents a software package for a multirotor UAV. This software package is focused on the monitoring territory and the search for human faces. The software package is based on the half implicitly direct visual odometry algorithm. The half-line approach eliminates the difficulties associated with the resource-intensive extraction of control points and at the same time it is a reliable method in terms of motion estimation. The algorithm operates with the directly intensity pixels, which results in a sub-pixel precision in the processing of the large volumes of frames. The probabilistic imaging technique explicitly recognizes an abnormal measurement to assess the three-dimensional coordinates of the point reduces the amount of abnormal measurements, and the measurements improve both quality and reliability. Such approach enhances the reliability of the estimation of movement in the case of the monotonous, repetitive or high privately texture. This algorithm is half implicitly direct visual odometry. The paper provides a comparison of this algorithm with other popular visual odometry algorithm in terms of reliability, operation speed and positioning mistakes. In addition, the paper discusses different approaches to the territory monitoring and the results of the computer implementation for these approaches.

**Keywords:** visual odometry, software system, monitoring, UAV, patrols, quadrotor, SLAM, PTAM, depth filters, algorithm.

### References

1. Karimov A.Kh. Objects and purposes that solved by new generation Unmanned Aircraft Complexes novogo pokoleniya. *Trudy MAI [Proc. MAI]*. 2011, no. 47, pp. 1–9 (in Russ.).
2. Kalinov I.A. Creation of an independent mini UAV based on a multirotor using the half implicitly direct visual odometry algorithm. *Intellektualnye sistemy, upravlenie i mekhatronika-2016: mater. Vseross. nauch.-tekhn. konf. [Proc. All-Russian Science and Practice Conf. "Intelligent Systems, Management and Mechatronics 2016"]*. Sevastopol, SevGU Publ., 2016, pp. 150–154 (in Russ.).
3. Malinetsky G.G., Kochkarov A.A. The future of Russian weapon and interdisciplinary approaches. *Intellekt i tekhnologii [Intellect and Technology]*. 2014, no. 1 (7), pp. 48–51 (in Russ.).
4. Forster C., Lynen S., Kneip L., Scaramuzza D. Collaborative Monocular SLAM with Multiple Micro Aerial Vehicles. *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. 2013, pp. 3962–3970.
5. Kerl C., Sturm J., Cremers D. Robust odometry estimation for RGB-D cameras. *Proc. IEEE ICRA*. 2013, pp. 3748–3754.
6. Meilland M., Comport A.I. On unifying key-frame and voxelbased dense visual SLAM at large scales. *Proc. IEEE/RSJ Int. Conf. on IROS*. 2013, pp. 3677–3683.
7. Klein G., Murray D. Parallel tracking and mapping for small AR workspaces. *IEEE and ACM Int. Sympos. on Mixed and Augmented Reality*. Nov. 2007, pp. 1–10.
8. Wu M., Ramakrishnan N., Lam S.-K., Srikanthan T. Low-complexity pruning for accelerating corner detection. *Proc. Conf. ISCAS-2012, IEEE Intern. Sympos.* 2012, pp. 1684–1687.
9. Weiss S., Achtelek M.W., Lynen S., Achtelek M.C., Kneip L., Chli M., Siegwart R. Monocular vision for long-term micro aerial vehicle state estimation: a compendium. *Journ. of FieldRobotics*. 2013, vol. 30, no. 5.
10. Odroid U2 user manual. Available at: <https://astataruf.files.wordpress.com/2015/11/odroid-x2-user-manual.pdf> (accessed May 25, 2016).

УДК 519.687.7

Дата подачи статьи: 23.05.16

DOI: 10.15827/0236-235X.115.181-186

## **ТИПОВЫЕ ФУНКЦИИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ РАБОЧЕГО МЕСТА ИНСТРУКТОРА АВИАЦИОННОГО ТРЕНАЖЕРА**

*Л.А. Чижикова, ведущий специалист, ludmilachizhikova@yahoo.com*

*(Компания «Гражданские самолеты Сухого», ул. Поликарпова, 23б, стр. 2, г. Москва, 125284, Россия)*

Разработкой авиационных симуляторов, имитирующих работу борта самолета и окружающей среды для обучения пилотированию летных составов, чаще всего называемых в Российской Федерации тренажерами, занимаются во всем мире. Исследования технических составляющих тренажеров показывают, что тренажерный комплекс можно разделить на несколько частей, одной из которых является программно-аппаратный комплекс рабочего места инструктора. Поэтому, как правило, в авиакосмической отрасли перед каждым разработчиком тренажера возникает задача создания ПО рабочего места инструктора.

Данное ПО позволяет инструктору, обучающему пилотированию воздушного судна, задавать параметры обучения полету, контролировать и управлять подсистемами симулятора. Очень часто для каждого разработчика тренажера данное ПО является уникальным, но, тем не менее, общий и основной функционалы прослеживаются у всех производителей авиационных тренажеров. В статье рассматриваются типовые функции ПО рабочего места инструктора, а также логика взаимодействия данного ПО с подсистемами тренажера.

**Ключевые слова:** авиационные тренажеры, математическое моделирование, ПО тренажеров, обучение пилотированию.

Все современные авиационные тренажеры включают в себя ПО *рабочего места инструктора* (РМИ). Данная работа посвящена технологиям создания компонента авиационных симуляторов, а именно РМИ, также в ней детально рассмотрены подходы к его построению.

Производители авиационных тренажеров зачастую вынуждены индивидуально проектировать данный компонент под свою вычислительную платформу. Однако выявление типовых функций РМИ может обеспечить возможность быстрого прототипирования ПО, что значительно упростит и сократит время разработки ПО такого типа. Функционал данного ПО направлен на задание параметрической информации бортового оборудования и управление тренажером при обучении пилотов [1, 2]. Цель исследования, освещенного в статье, – найти типовые и основные функции ПО РМИ, используемые при обучении пилотированию.

Предметом изучения были рекомендации авиационных стандартов, применяемых при разработке тренажеров, литература по данной предметной области, а также консультации с практикующими обучением пилотами-инструкторами.

Основными пользователями данного ПО являются пилоты-инструкторы. Они должны разрабатывать учебные программы подготовки пилотов, методики ведения летной подготовки, уметь применять на практике методику выполнения маневров и схем полета, по которой намерены проводить летную подготовку.

В соответствии с методическими рекомендациями по подготовке пилотов обладатель квалификационной отметки «пилот-инструктор» должен иметь знания о возможностях человека применительно к летной подготовке, а также об опасностях, связанных с имитацией отказов систем на воздушном судне, уметь контролировать самостоятельные

полеты пилотов-курсантов, проводить летную подготовку, необходимую для выдачи свидетельств частного пилота, коммерческого пилота, пилота многочленного экипажа и для внесения квалификационной отметки о праве на полеты по приборам и квалификационной отметки пилота-инструктора.

Основным средством при подготовке пилотов на авиационном тренажере во время летной сессии для пилота-инструктора является ПО РМИ, посредством которого он может задавать параметры обучения и возмущающее воздействие на подсистемы, контролировать работу тренажера.

В гражданской авиации в полнопилотажных тренажерах РМИ размещено в кабине тренажера, в авиационных симуляторах военного назначения – за пределами тренажера, что никак не влияет на типовой набор функций ПО.

ПО РМИ является прикладной программой с использованием графического пользовательского интерфейса, с системным модулем инициализации и ввода/вывода данных от подсистем тренажера.

Общими принципами построения ПО РМИ являются комплексность, которая реализуется путем создания системы, гибко настраиваемой для работы с различным сочетанием программных и аппаратных средств, и модульность, реализуемая последовательным делением структуры всей программы на программные модули, что должно обеспечивать наращивание функциональности [3].

Для предоставления инструктору возможности обучать пилотов самолетовождению по ранее разработанным методикам данное ПО имеет управляющие элементы графического пользовательского интерфейса, посредством нажатия на которые инструктор может задавать возмущающее воздействие на подсистемы тренажера, менять числовые значения параметров подсистем тренажера, а также элементы индикации необходимых параметров,

чтобы была возможность следить за корректным обучением пилотированию.

Структура программы РМИ может быть подразделена на следующие составные части:

- программный модуль общего функционала, в том числе обработка управляющих элементов графического пользовательского интерфейса приложения;
- интерфейсный программный модуль обмена информацией приложения РМИ с подсистемами тренажера.

В свою очередь, программный модуль общего функционала включает в себя следующие составляющие:

- программный модуль графического пользовательского интерфейса;
- модуль общих настроек приложения, взаимодействующий с элементами графического пользовательского интерфейса приложения;
- модель взаимодействия с интерфейсным уровнем приложения для изменения и индикации параметрической информации посредством графического пользовательского интерфейса.

На интерфейсном уровне программный модуль РМИ взаимодействует с математическими моделями, имитирующими работу бортовых систем самолета.

Так как ключевым фактором в создании ПО РМИ является его функциональность, чаще всего для разработки такого вида программ выбирают методологию FDD Agile (Feature driven development Agile), основанную на модификации гибкой (Agile) методологии, – подход к разработке ПО, управляемый функциональностью, который включает пять основных видов деятельности:

- разработка общей модели;
- составление списка необходимых функций системы;
- планирование работы над каждой функцией;
- проектирование функции;
- реализация функции [4].

Функционал ПО РМИ схож с функционалом всего авиационного тренажера как комплекса вычислительных систем с математическими моделями, так как затрагивает управление всеми подсистемами тренажера при обучении пилотированию.

Математические модели самолета, имитирующие работу подсистем, являются основными составляющими для любого тренажера [5].

Согласно рекомендациям стандарта ARINC 610 ревизии С – Руководству по проектированию авиационного оборудования и ПО для использования в учебных устройствах (Guidance for design of aircraft equipment and software for use in training devices), нестандартные для самолета процедуры (функции) в симуляторе могут быть разбиты на три основные категории: начальные условия, репозиция, параметры заморозки, что, как правило, отражено в графическом пользовательском интерфейсе про-

граммы. ПО РМИ взаимодействует с математическими моделями самолета и позволяет задавать параметры обучения полету – начальные условия, такие как географическое местоположение самолета, высота, скорость, заполнение топливом самолета, метеоусловия и параметры внешних воздействий, таких как температура, давление и ветер.

В более ранней версии стандарта ARINC 610 руководства по проектированию и интеграции авиационного и радиоэлектронного оборудования в тренажерах упоминаются рекомендации IATA по применению семнадцати основных функций симулятора (табл. 1), которые находят отражение в ПО РМИ.

Функции ПО РМИ, как и функции тренажера, могут быть разделены на следующие категории: контроль симуляции, сценарий настроек обучения, оптимизация, установки технического обслуживания, специфические функции для военного применения.

Функции управления и взаимодействия ПО РМИ сгруппированы в графическом пользовательском интерфейсе с учетом вышеупомянутых категорий.

Таблица 2 демонстрирует принадлежность функций к категориям.

Во время тренировочной сессии может понадобиться так называемая заморозка состояния самолета для того, чтобы приостановить обучение, дать экипажу дополнительные инструкции или обсудить его поведение. Эти функции могут потребовать полной заморозки процесса моделирования или остановки некоторых выбранных параметров самолета. Во время замораживания полета движение самолета приостанавливается, в то время как самолетные системы (математические модели, имитирующие работу самолетных систем) продолжают функционировать нормально и информационный обмен сохраняется.

Для этого в функциях ПО РМИ предусмотрены кнопки графического пользовательского интерфейса: заморозка полета, заморозка местоположения, заморозка высоты, заморозка топлива.

Типичными примерами использования инструктором функции заморозки во время обучения пилотированию могут служить следующие ситуации.

- Если экипаж вылетает за пределы полигона или допускает ошибку в процессе пилотирования, инструктор приостанавливает моделирование для того, чтобы обсудить с ним эту ситуацию.

- При переходе к следующей учебной задаче замораживание моделирования может позволить обучаемым установить измененную конфигурацию самолета без выполнения полета с расширенным маневром. Например, замораживание положения самолета при заходе на посадку дает время для снижения скорости приземления и выбора правильного положения закрылков.

Таблица 1

## Основные функции авиационного симулятора по версии IATA

Table 1

## Basic functions of a Flight Simulation Device (IATA version)

Номер функции	Функция	Эффект
1	Заморозка полета	Такие параметры, как широта, долгота, воздушная скорость, курс и высота, заморожены (то есть переменные при информационном обмене являются константами)
2	Заморозка местоположения	Широта и долгота местоположения самолета заморожены, все остальные параметры полета остаются динамическими
3	Заморозка высоты	Высота самолета остается неизменной
4	Заморозка топлива	Количество топлива остается неизменным
5	Настройка местоположения	Предусматривает возможность пошагового изменения местоположения самолета по широте и долготе
6	Изменение широты/долготы	Предусматривает возможность изменения местоположения самолета по широте и долготе посредством выбора из предоставленного диапазона
7	Изменение высоты	Предусматривает возможность изменения высоты самолета (больше или меньше) посредством выбора из предоставленного диапазона
8	Изменение скорости	Предусматривает возможность изменения скорости движения самолета (больше или меньше) посредством выбора из предоставленного диапазона
9	Изменение курса	Предусматривает возможность изменения курса (направления движения) самолета посредством выбора из предоставленного диапазона
10	Функция «сделать снимок» (состояния системы тренажера)	Позволяет сохранить все параметры (все переменные в виде констант) для записи необходимых состояний тренажера в момент снимка
11	Вызов (сохраненного) снимка (состояния системы тренажера)	Передаёт необходимые параметры тренажеру, чтобы вернуть его в состояние, когда был сделан снимок
12	Ускорение в $N$ раз	Вызывает скорость изменения положения, высоты и скорости изменения количества топлива, что модифицируется мультипликативным множителем $N$ (скорость самолета вдоль траектории полета изменена)
13	Нулевое значение топлива	Предусматривает изменение настроек самолета нулевого значения веса топлива на земле или в воздухе
14	Установка веса топлива	Предусматривает изменение веса топлива на земле или в воздухе
15	Настройка температуры/давления	Предусматривает изменение уровня моря, земли или температуры воздуха за бортом и давления
16	Репозиция (мгновенное перемещение) самолета	Моментальное изменение широты и долготы, которые могут включать в себя любую комбинацию изменений курса, высоты, воздушной скорости, угла атаки, тяги, угла наклона и конфигурации, с помощью которых может быть достигнуто гладкое перемещение. Выполнение данной функции может быть прекращено при заморозке полета
17	Настройка ветра	Предусматривает настройки направления ветра и скорости относительно земной поверхности или в воздухе

- В конце полета может возникнуть необходимость заморозить моделирование перед передачей тренажера другому пилоту для продолжения полета в условиях, которые только что были установлены.

- Во время выполнения маневра инструктор может заморозить моделирование для того, чтобы объяснить функцию системы или состояние, которое является переходным и пропадает в конце выполнения маневра.

В процессе работы тренажера замороженные параметры (переменные, участвующие в информационном обмене) остаются константами, а информационный обмен сохраняется с заданной циклограммой.

В течение тренировочной сессии летного состава может понадобиться перестановка место-

положения самолета (репозиция) или внешних воздействий, соответствующих задаче обучения. Использование данной функции ПО РМИ ведет к изменению географического местоположения самолета, высоты, курса, скорости.

Данная функция изменения местоположения самолета касается установки условий максимального взлетного веса самолета и условий видимости III категории или условий посадки при заходе на посадку с боковым ветром. Изменение параметров может выражаться в шаговом изменении на новое значение или поворотом как непрерывное динамическое изменение от начального значения до конечного. Например, инструктор может повернуть положение самолета. В число изменяемых параметров в соответствии со справочником пилота входят широта, долгота, курс, высота и воздушная

Таблица 2

## Сводная таблица основных функций программы РМИ

Table 2

## A summary table of the Instructor Station software main functions

Функции ПО РМИ/авиационного тренажера	Контроль симуляции	Настройки параметров обучения	Оптимизация	Установки технического обслуживания	Специфические функции для военного применения
Полная заморозка (полный останов)	x				
Заморозка полета (останов полета)	x				
Заморозка топлива	x				
Заморозка широты/долготы	x				
Заморозка высоты	x				
Изменение массы самолета		x			
Изменение широты/долготы		x			
Изменение высоты		x			
Изменение воздушной скорости		x			
Изменение курса		x			
Изменение температуры/давления		x			
Изменение направления ветра		x			
Переход к конкретным условиям		x			
Сброс внутреннего состояния системы		x			
Поддержка балансировки самолета		x			
Ускорение в $n$ раз			x		
Создание снимка			x		
Вызов снимка			x		
Сохранение снимка			x		
Загрузка снимка			x		
Сброс памяти ошибок				x	
Установка параметров регистрации сбоев				x	
Сохранение памяти ошибок				x	
Загрузка памяти ошибок				x	
Заморозка снаряжения (боеприпасов)					x
Заморозка времени выполнения миссии					x
Заморозка (останов) миссии					x
Изменение снаряжения (смена расходных материалов)					x
Изменение содержания подвесок (смена груза/грузоподъемности)					x
Изменение времени выполнения миссии					x
Приостановка хранения снаряжения (разгрузки)					x
Прекращение хранения снаряжения (разгрузки)					x

скорость; остаток топлива, ресурсы и полезная нагрузка; внешние воздействия, такие как температура, давление и состояние ветра.

Во время летной сессии (сеанса обучения пилотированию) может возникнуть необходимость сохранить конкретную установку моделирования и позже сделать ее доступной для повторного вызова во время сеанса обучения. Инструктор может использовать функцию вызова кадра для сохранения параметров самолета в любое время, не прерывая обучения. Функция повторного вызова кадра может быть полезной, если инструктор хочет иметь возможность вернуться к конкретной точке в сценарии обучения студентов. Это может быть использовано и для настройки начальных установок состояния. Кадры также могут быть использованы

для сохранения состояния конкретных самолетных систем. Кадр плана полета системы самолетовождения является наиболее общим примером.

Для каждой конкретной модели самолета зафиксирован набор отказов подсистем. Главный и типичный функционал программы включает в себя возможность симулировать, задавать отказы подсистем авиационного симулятора во время тренировочной сессии.

Основное назначение пилотажного учебного оборудования состоит в том, чтобы безопасно обучать ненормальным и срочным процедурам. Инструктор должен иметь возможность устанавливать и снимать отказы при симуляции пилотажного оборудования тренажера таким образом, чтобы все аварийные и экстренные случаи в учебных про-

граммах могли должным образом усваиваться, но без трудоемкой реконфигурации в конце каждого учебного сценария.

Сбой может быть определен как единственный физический элемент воздушного судна, который вышел из строя. Влияние отказа будет распространяться через все моделирование таким же образом, как на самолете, с физическим влиянием, индикацией систем кабины экипажа, предупреждающими сообщениями и сигналами тревоги.

Во всех этих случаях ПО должно предусматривать возможность использования других функций моделирования, таких как замораживание или перемещение, и переустановки индикаторов кабины и самолетных систем в состоянии отсутствия отказов при завершении сценария обучения.

Обучение пилотированию военных самолетов и обучение пилотированию гражданских воздушных судов имеют много общего, но для пилотирования военных самолетов необходимо освоение дополнительных функций, которые часто требуются для обучения тактическим операциям [6].

Часто встречающимися дополнительными функциями ПО РМИ являются функции сообщений метеосводок (ATIS – Automatic Terminal Information Service – служба автоматической передачи информации в районе аэродрома), отслеживание траектории полета по навигационным картам.

Таким образом, для обеспечения гибкости настройки для любого уровня тренажера и его вычислительной сети, наращивания функционала, а также возможности использования в тренажерах, требующих вычислений в реальном времени, структура ПО РМИ должна характеризоваться модульностью, под которой понимается логическое разделение составляющих программы на подпрограммы [3, 5].

Для предоставления инструктору возможности обучать пилотов самолетовождению по ранее разработанным методикам задаются параметры посредством нажатия на кнопки графического пользовательского интерфейса программы РМИ, таким образом задается возмущающее воздействие на подсистемы тренажера. Объектами автоматизации являются математические модели подсистем самолета, имитаторы реальных приборов бортового радиоэлектронного оборудования и внешней окружающей среды для визуализации закабинного пространства.

Симуляция полета реализована математическими моделями аэродинамики самолета, аэроупругости, системы автоматического управления, силовой установки, авионики, атмосферы и других [6, 7].

Математические модели, используемые для симуляции современных воздушных судов, состоят из массивного набора нелинейных дифференциальных уравнений с большим количеством данных (табличного представления, в коде программы

представлены как  $n$ -мерные массивы), аэродинамических функций, иногда зависящих от 4-5 переменных, произвольных дискретных и непрерывных функций, а также множества ограничений, наложенных на состояние различных систем самолета.

Из полученных сил и моментов, создаваемых этими системами, вычислительным комплексом создаются и решаются уравнения движения с использованием адекватного алгоритма интеграции. Это позволяет моделировать полный спектр статических и динамических условий эксплуатации воздушных судов, в том числе посадки и взлета (landing/take-off), наземного обслуживания и экстренных ситуаций, таких как отказ двигателя, потеря скорости, отказ подсистем [5].

Если уровень тренажера требует использования систем реального масштаба времени имитации полета, компьютер должен выполнять все вычисления в математической модели (моделях), в том числе обмен данными (ввод/вывод) при такте, как на реальном борту самолета, для достижения самой высокой точности исполнения [5].

Во время вызова инструктором и выполнения всех упомянутых функций информационный обмен системного модуля ПО РМИ с подсистемами тренажера сохраняется.

В ходе исследований, проводимых во время летных сессий инструкторами-пилотами, было установлено, что существующие программные решения РМИ гражданской тематики, как правило, имеют пользовательский интерфейс на английском языке. Для новых разработок такого типа ПО также было высказано пожелание использовать английский язык в графическом пользовательском интерфейсе. Данное требование обусловлено тем, что все сообщения пилотам в кабине, предупреждения об отказах оборудования выдаются на международном языке – английском.

Также в ходе консультаций с пилотами-инструкторами было выявлено, что наиболее часто используемыми функциями при обучении пилотированию являются замораживание полета, управление подачей питания на борт, создание снимка – мгновенное сохранение состояний параметров тренажера для разбора поведения пилотов после проведения обучения.

При обучении пилотированию инструкторы, как правило, используют следующий алгоритм.

1. Установка начального местоположения самолета (репозиция).
2. Установка веса самолета.
3. Установка метеоусловий.
4. Установка дополнительных параметров (выставка инерциальной системы и пр.).
5. Выставление отказов оборудования, если этого требует упражнение.

Данные факторы нашли отражение в проектировании графического пользовательского интерфейса программы РМИ, где наиболее часто исполь-

зуемые функции заданы управляющими элементами – кнопками на экране.

Чаще всего в графическом пользовательском интерфейсе данного типа программ представлена интеграция вышеописанных функций с выделением информационных страниц программы с разделением их по категориям и дополнением навигационными картами, функцией метеосводок для конкретного аэродрома, функцией предупреждения столкновений, набором отказов для каждой конкретной модели самолета [8].

На основании вышесказанного выделим типичный для авиационных тренажеров набор функций ПО РМИ, достаточный для проектирования программы с базовыми функциями для дальнейшей модификации и добавления индивидуальных параметров под определенную модель самолета:

- настройка состояния системы и внешних условий для построения сценария обучения;
- заморозка;
- мгновенное изменение местоположения (позиции) самолета;
- симуляция отказов подсистем воздушного судна;
- оптимизация выполнения учебных задач и служебные функции программы.

Рекомендациями и основными направлениями для дальнейших исследований оптимальных методов такого типа программ будут следующие:

- наиболее часто используемый пилотами-инструкторами функционал при обучении пилотированию летных составов ПО РМИ и возможные ал-

горитмы использования элементов управления ПО РМИ во время обучения пилотированию для нахождения оптимальной структуры такого типа программ;

- вопросы эргономики при проектировании графического пользовательского интерфейса такого вида программ.

Кроме того, необходимо исследовать возможность построения модели системы управления авиационным симулятором посредством ПО РМИ.

### Литература

- 1 Красовский А.А. Основы теории авиационных тренажеров. М.: Машиностроение, 1995. 304 с.
- 2 Клюев Г.И. Авиационные приборы и системы: учеб. пособие. Ульяновск: Изд-во УлГТУ, 2000. 343 с.
- 3 Hoogstraten J.A., van de Moesdijk G.A.J. Modular programming structure applied to the simulation of non-linear aircraft models. TU Delft Publ., m-437, Delft, Netherlands, 1982, 9 p.; URL: <http://repository.tudelft.nl/islandora/object/uuid:0a0508b9-7e2a-4f61-b3ce-006644642937?collection=research> (дата обращения: 22.05.2016)
- 4 Palmer S.R., Felsing J.M. A practical guide to feature-driven development. Prentice Hall Publ. 2002, 276 p.
- 5 Baarspul M. Flight Simulation Techniques with emphasis on the generation of high fidelity 6 DOF motion cues. TU Delft Publ., m-533, Delft, Netherlands, 1986, 13 p.; URL: <http://repository.tudelft.nl/islandora/object/uuid:dfd19219-afd9-4faf-b2c8-03401bfc89c0?collection=research> (дата обращения: 22.05.2016).
- 6 Allerton D. Principles of Flight Simulation. Wiley & Sons, UK, 2009, 492 p.
- 7 Mulder J.A., Hollander J.G. Status of Dynamic Flight test Technology-Model Identification for Flight Simulation. SAE Techn. Pap. Series no. 810597, 1981, pp. 1–6.
- 8 Vazques A.A. Touch screen use on flight simulator instructor/operator stations. MS Thesis. US Naval Acad. Publ., Monterey, 1990, 78 p.

DOI: 10.15827/0236-235X.115.181-186

Received 23.05.16

### TYPICAL FUNCTIONS OF FLIGHT SIMULATION DEVICE INSTRUCTOR STATION SOFTWARE

<sup>1</sup>Chizhikova L.A., *Leading Specialist*

<sup>1</sup>JSC “Sukhoi Civil Aircraft”, Polikarpova St. 23B/2, Moscow, 125284, Russian Federation

**Abstract.** Companies all around the world are engaged in development of a flight simulator device that imitates aircraft operation and environment for flight crew piloting training. The research of the flight simulator device’s technical components shows that a training complex can be divided into several parts, one of which is the Instructor Station hardware-software complex. Therefore, usually in each simulator developer faces the problem of creating an instructor station software. This software allows the instructor, who teaches aircraft piloting, to set flight parameters, monitor and control training device subsystems. It is very often that this software is unique for every training device developer. Nevertheless the general and basic functions are discernible among all flight simulator device manufacturers. The article considers typical functions of the instructor station software, as well as the logic of software interaction with the simulator subsystems.

**Keywords:** flight simulation device, mathematical modeling, simulation device software, crew piloting training.

### References

- 1 Krasovsky A.A. *Osnovy teorii aviatsionnykh trenazherov* [The basics of Aviation Simulator Theory]. Moscow, Mashinostroenie Publ., 1995, 303 p.
- 2 Klyuev G.I. *Aviatsionnye pribory i sistemy* [Aviation Devices and Systems]. Ulyanovsk, UIGTU Publ., 2000, 343 p.
- 3 Hoogstraten J.A., van de Moesdijk G.A.J. *Modular programming structure applied to the simulation of non-linear aircraft models*. TU Delft Publ., m-437, Delft, Netherlands, 1982, 9 p.
- 4 Palmer S.R., Felsing J.M. *A practical guide to feature-driven development*. Prentice Hall Publ., 2002, 276 p.
- 5 Baarspul M. *Flight Simulation Techniques with emphasis on the generation of high fidelity 6 DOF motion cues*. TU Delft Publ., m-533, Delft, Netherlands, 1986, 13 p.
- 6 Allerton D. *Principles of Flight Simulation*. Wiley & Sons Publ., UK, 2009, 492 p.
- 7 Mulder J.A., Hollander J.G. *Status of Dynamic Flight test Technology-Model Identification for Flight Simulation*. SAE Techn. Pap. Series no. 810597, 1981, pp. 1–6.
- 8 Vazques A.A. *Touch screen use on flight simulator instructor/operator stations*. MS Thesis. US Naval Acad. Publ., Monterey, 1990, 78 p.



УДК 621.391.822

Дата подачи статьи: 29.05.16

DOI: 10.15827/0236-235X.115.187-191

## **ВЛИЯНИЕ ВНЕШНИХ ДЕСТАБИЛИЗИРУЮЩИХ ФАКТОРОВ НА ВЕЛИЧИНУ НИЗКОЧАСТОТНОГО ШУМА ИНТЕГРАЛЬНЫХ СХЕМ**

*М.И. Горлов, д.т.н., профессор, m-gorlov@inbox.ru;*

*Д.М. Жуков, аспирант, ddimochka@mail.ru*

*(Воронежский государственный технический университет,  
Московский просп., 14, г. Воронеж, 394026, Россия)*

В статье представлены результаты экспериментов по измерению низкочастотного шума напряжения на выводах интегральных схем в зависимости от температуры окружающей среды, а также до и после воздействия электростатическими разрядами. Полученные данные показывают, что низкочастотный шум может являться информативным параметром для диагностики и разбраковки интегральных схем по надежности.

Низкочастотный шум достаточно заметно изменяется под влиянием внешних воздействий. Величина этого изменения зависит от состояния интегральной схемы. Ряд экспериментов, проведенных авторами статьи, доказывает данный факт. Так, например, имеет место накопительное действие электростатических разрядов, то есть каждый последующий разряд влияет на величину шума сильнее предыдущего. Зависимость низкочастотного шума от температуры говорит о том, что изменение внутреннего состояния интегральной схемы приводит к изменению величины данного вида шума. Такая зависимость дает основания предполагать, что низкочастотный шум может выступать индикатором любого внешнего воздействия, вызывающего изменение состояния внутренней структуры интегральной схемы. О чувствительности низкочастотного шума как информативного параметра говорит также изменение его величины при отжиге электростатических дефектов, рассмотренном в статье. После отжига дефекты, внесенные электростатическими разрядами, частично или полностью устраняются, так что изменение шума после отжига говорит о прямой зависимости его величины от внутреннего состояния структуры интегральной схемы.

Таким образом, эксперименты, показывающие прямую связь величины шума и внутреннего состояния интегральной схемы, позволяют разрабатывать методы сравнительных испытаний и разбраковки интегральных схем по надежности. В статье представлен способ сравнительной оценки качества и надежности двух и более партий интегральных схем, разработанный на основе полученных данных.

**Ключевые слова:** низкочастотный шум, электростатический разряд, интегральная схема, способ сравнительной оценки.

В состав радиоэлектронной аппаратуры (РЭА) входит огромное количество полупроводниковых изделий различных типов. Наиболее сложными из них являются интегральные схемы (ИС), реализующие большинство функций РЭА. С расширением функционала РЭА повышается и сложность ИС.

Современные технологии позволяют разместить на кристалле площадью менее 1 см<sup>2</sup> миллионы компонентов, выполняющих различные задачи. Причем на одном кристалле могут быть расположены как аналоговые, так и цифровые устройства.

С ростом сложности РЭА растут и требования к ее надежности. Особенно это касается аппаратуры военного, космического и промышленного назначения. Отказ всего одного элемента в таких приборах может обойтись очень дорого, вплоть до гибели людей. Поэтому необходимо знать, насколько надежен каждый компонент до его монтажа на печатную плату. Поскольку ИС являются наиболее сложными как с технологической, так и с конструктивной точки зрения, предсказание их надежности – трудная научная и инженерная задача.

Сегодня для обеспечения качества и надежности ИС осуществляются контроль их параметров на каждом этапе производства, а также различные тренировки, такие как термотренировки, электротренировки и электротермотренировки. Суть испытаний с использованием тренировок в том, что ИС

подвергаются повышенным тепловым и/или электрическим нагрузкам в течение длительного времени (порядка 1 000 час.) с периодическим или непрерывным контролем работоспособности. Испытания продолжаются до наступления параметрического или катастрофического отказа. Таким способом можно выяснить, какова средняя наработка до отказа в заданных условиях эксплуатации любого типа ИС.

Описанный метод наиболее достоверен в плане определения качества и надежности ИС, но имеет два основных недостатка: высокие денежные и временные затраты на испытания и невозможность испытать каждую выпускаемую ИС.

Для преодоления этих трудностей необходимы альтернативные методы контроля качества и надежности ИС на всех этапах производства – от контроля на пластине до контроля перед монтажом на печатную плату.

Одним из таких способов может стать контроль качества и надежности ИС по параметрам низкочастотного шума (НЧШ).

НЧШ присутствует на выводах любого электронного компонента и представляет собой флуктуации тока и/или напряжения [1]. Анализ характера таких флуктуаций может дать информацию о внутреннем состоянии ИС [2]. Сам факт существования НЧШ свидетельствует о наличии деградиационных процессов в схеме [3].

### Исследование НЧШ ИС

Для диагностики надежности ИС по параметрам НЧШ необходимо знать, каким образом НЧШ реагирует на изменение внутреннего состояния испытуемого прибора. Такую информацию можно получить, исследуя поведение шума в зависимости от внешних воздействий [4].

Известно, что источником НЧШ в полупроводниковых приборах и ИС являются различные дефекты кристалла полупроводника [5]. К ним относятся как дефекты, расположенные в объеме кристалла, так и поверхностные или дефекты на границах раздела двух материалов [6, 7]. Дефекты выступают в роли ловушек для свободных носителей заряда. Захват и эмиссия электронов или дырок на ловушечные энергетические уровни, расположенные внутри запрещенной зоны полупроводникового материала, порождают флуктуации концентрации свободных носителей, что, в свою очередь, порождает флуктуации тока и напряжения.

Под внешним воздействием происходит изменение внутреннего состояния полупроводникового кристалла [8]. Общая картина данного изменения будет зависеть от количества и типа дефектов, присутствующих в кристалле. Индикатором изменения может являться величина НЧШ [9].

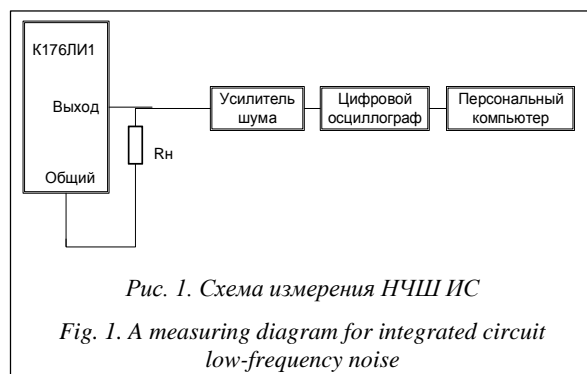
В работе исследовалось поведение шума под воздействием повышенной температуры и электростатических разрядов (ЭСР).

Для исследования случайным образом были выбраны 15 логических ИС типа К176ЛИ1.

Данная микросхема представляет собой логический элемент 9И-НЕ, выполненный по КМОП-технологии. НЧШ измерялся по схеме, приведенной на рисунке 1.

Номинал нагрузочного резистора составлял 600 Ом. При такой величине нагрузки ток, проходящий через ИС, был достаточно велик для измерения шума, но не вызывал перегрева схемы.

После оцифровки с помощью цифрового осциллографа полученный сигнал сохранялся в памяти компьютера в виде последовательности значений флуктуирующего напряжения. После этого вычислялось среднеквадратичное значение напряжения.



В начале эксперимента был измерен НЧШ каждой ИС в выборке в полосе частот 500–1 500 Гц при различных температурах окружающей среды. График изменения среднеквадратичного значения шумового напряжения приведен на рисунке 2. Температура изменялась от комнатной до 100 °С и обратно со скоростью не более 2 °С/мин.

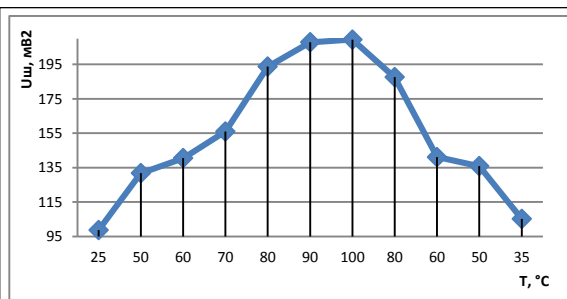


Рис. 2. График изменения дисперсии шумового напряжения в зависимости от температуры окружающей среды

Fig. 2. A graph of noise voltage dispersion behaviour depending on environment temperature

Такое поведение НЧШ было характерно для каждой исследуемой ИС – различались только значения.

Как видно из графика, уровень НЧШ увеличивается с температурой, что объясняется увеличением интенсивности процессов, являющихся причиной шума. К таким процессам можно отнести процессы захвата-эмиссии свободных носителей заряда ловушками в подзатворном окисле. Следует обратить внимание на то, что форма зависимости одинакова как при увеличении температуры, так и при ее последующем снижении (отсутствует гистерезис). Это говорит о том, что при кратковременном воздействии повышенной температуры в структуре ИС не происходит никаких необратимых процессов, которые могли бы повлиять на уровень шума.

Далее на каждую ИС в выборке воздействовали электростатическим разрядом потенциалом 200 В (максимально допустимым по техническим условиям) по модели тела человека. Воздействие проводилось по выводам «выход–общая точка», «выход–питание». По входам воздействия не проводилось из-за большой чувствительности подзатворного окисла к ЭСР. После воздействия измерялся НЧШ.

Далее проводились два цикла отжига при 100 °С длительностью 1 час каждый. График изменения уровня НЧШ показан на рисунке 3.

Из графика видно, что НЧШ ИС уменьшается после воздействия ЭСР. Далее имеет место нестабильность среднеквадратичного значения шумового напряжения после первого и второго отжигов. В итоге уровень шума стал ниже, чем в начале эксперимента.

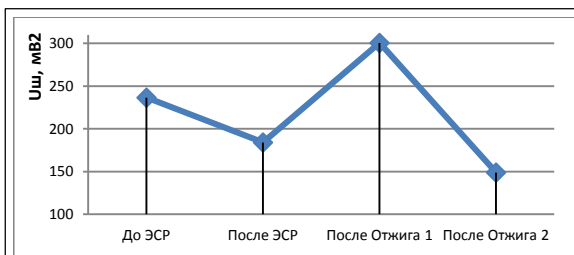


Рис. 3. График изменения уровня НЧШ

под воздействием ЭСР с последующим отжигом

Fig. 3. A graph of low-frequency noise level behaviour affected by an electrostatic discharge with follow annealing

Данное явление, по-видимому, связано с физическими свойствами дефектов структуры ИС, вносимых ЭСР.

Для исследования влияния ЭСР с потенциалами, значительно превышающими допустимые по ТУ, были случайным образом выбраны 10 логических ИС типа K176ЛП11.

Для исследования влияния величины потенциала ЭСР на величину НЧШ ИС был проведен следующий эксперимент. На выборке из 5 ИС измерялось среднеквадратичное значение шумового напряжения до воздействия ЭСР. Далее последовательно проводились воздействия ЭСР, каждый раз увеличивая величину его потенциала, начиная с 400 В. После каждого воздействия измерялся уровень НЧШ. Результаты эксперимента показаны на рисунке 4.

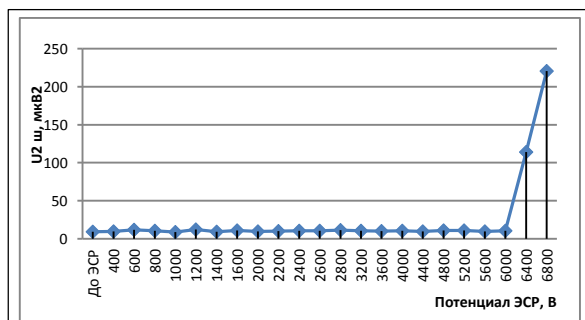


Рис. 4. Зависимость уровня НЧШ ИС типа K176ЛП11 от величины потенциала ЭСР

Fig. 4. A dependence of K176ЛП11 integrated circuit low-frequency noise level on a electrostatic discharge potential value

Очевидно, что уровень НЧШ ИС не менялся вплоть до потенциала ЭСР в 6 400 В. При достижении данной величины уровень шума резко возрастал на два порядка. При этом ИС оставалась работоспособной, хотя уровень логической единицы несколько отклонялся от первоначальной величины, оставаясь при этом в пределах нормы.

Для проверки критичности величины потенциала ЭСР в 6 400 В был проведен следующий эксперимент.

На выборке из 5 ИС было измерено среднеквадратичное значение шумового напряжения до воздействия ЭСР. Далее на ИС последовательно осуществлялись воздействия ЭСР потенциалом 6 400 В. После каждого воздействия измерялся уровень НЧШ. Всего было по 4 воздействия на каждую ИС. Результаты эксперимента представлены на рисунке 5.

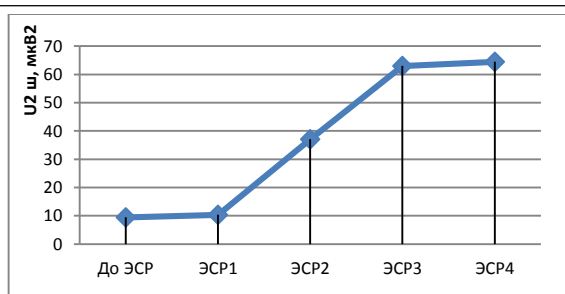


Рис. 5. Зависимость уровня НЧШ ИС типа K176ЛП11 от количества воздействий ЭСР потенциалом 6400 В

Fig. 5. A dependence of K176ЛП11 integrated circuit low-frequency noise level on a number of 6400 V potential electrostatic discharge influences

График показывает, что воздействие ЭСР потенциалом 6 400 В не так сильно изменяет уровень шума ИС, как в предыдущем эксперименте. Очевидно, что это происходит из-за того, что в предыдущем эксперименте на ИС предварительно проводилось 20 воздействий ЭСР со все более возрастающим потенциалом, то есть имеет место накопительный эффект. Так, из рисунка 5 видно, что НЧШ практически не изменился после первых двух воздействий ЭСР потенциалом 6400 В, но начал резко меняться уже после третьего.

Результаты всех проведенных экспериментов показывают, что ЭСР и повышенная температура влияют на уровень НЧШ ИС. Это влияние заключается в увеличении уровня НЧШ после воздействия ЭСР и во время воздействия повышенной температуры. Также установлено, что влияние имеют только ЭСР с достаточно высоким потенциалом. ЭСР с максимально допустимым по ТУ на ИС потенциалом никак не влияют на уровень НЧШ.

Установлено, что при воздействии ЭСР на ИС имеет место накопительный эффект, то есть величина, на которую изменяется уровень НЧШ после воздействия ЭСР, напрямую зависит от количества и потенциала предыдущих воздействий. Так, из третьего эксперимента видно, что при однократном воздействии ЭСР с потенциалом, превышающим максимально допустимый по ТУ более, чем в 30 раз, уровень НЧШ не изменяется. Изменения наступают при дальнейших воздействиях. Это говорит о том, что каждое воздействие вносит дополнительные дефекты в ИС, которые снижают стойкость ИС к последующим воздействиям.

### Способ сравнительных испытаний по надежности партий ИС

На основе полученных экспериментальных данных разработан и запатентован способ сравнительной оценки партий ИС по надежности [10]. Он основан на измерении информативного параметра в исходном состоянии, после воздействия на ИС пяти импульсов ЭСР обеих полярностей потенциалом, предельно допустимым по ТУ, после хранения в течение 72 часов при нормальных условиях (атмосферном давлении, температуре  $T=22\pm 5$  °С), после термического отжига при температуре  $T_{отж}=100$  °С в течение 2 часов и сравнении трех величин:  $\Delta_1=X_{нач}-X_{ЭСР}$ ,  $\Delta_2=X_{ЭСР}-X_{хр}$ ,  $\Delta_3=X_{хр}-X_{отж}$ , где  $X_{нач}$  – значение информативного параметра в начале измерений (в исходном состоянии);  $X_{ЭСР}$  – значение информативного параметра после воздействия ЭСР;  $X_{хр}$  – значение информативного параметра после хранения в течение 72 часов;  $X_{отж}$  – значение информативного параметра после термического отжига при температуре  $T=100$  °С.

Способ был опробован на выборках из двух партий ИС типа К155ЛЕ1 (четыре логических элемента 2ИЛИ-НЕ, выполненных по ТТЛ-технологии с окисной изоляцией карманов). Из каждой партии методом случайной выборки было отобрано по 10 ИС. В качестве информативного пара-

метра взято среднеквадратичное значение напряжения НЧШ  $\overline{U_{ш}^2}$ . На каждой ИС измерено значение шумового напряжения  $\overline{U_{шнач}^2}$  по выводам *питание-общая точка* (выводы 14–7). Измерение НЧШ проводилось при токе 7,5 мА, с полосой частот 200 Гц, при центральной частоте 1 000 Гц. Затем ИС подвергались воздействию ЭСР потенциалом 200 В. Воздействие осуществлялось следующим образом. ЭСР подавался на выводы ИС: *питание-общая точка* (выводы 14–7), *вход-питание* (выводы 2–14), *выход-питание* (выводы 1–14), *вход-выход* (выводы 1–2). Сначала подавались пять ЭСР одной полярности, потом пять ЭСР другой полярности. Затем замерялось среднеквадратичное значение напряжения НЧШ после ЭСР  $\overline{U_{шЭСР}^2}$ . Далее ИС хранились в нормальных условиях в течение 72 часов. Замерялось среднеквадратичное значение напряжения НЧШ после хранения  $\overline{U_{шхр}^2}$ . Осуществлялся термический отжиг всех ИС при температуре  $T=100$  °С. Замерялось среднеквадратичное значение напряжения НЧШ после ЭСР  $\overline{U_{шотж}^2}$ . Далее для каждой ИС определялись значения величин  $\Delta_1 = \overline{U_{шнач}^2} - \overline{U_{шЭСР}^2}$ ,  $\Delta_2 = \overline{U_{шЭСР}^2} - \overline{U_{шхр}^2}$ ,  $\Delta_3 = \overline{U_{шхр}^2} - \overline{U_{шотж}^2}$ .

Результаты испытаний выборки из партии 1

Таблица 1

Table 1

Test results of a lot sample 1

№ ИС	$\overline{U_{шнач}^2}$ , мкВ <sup>2</sup>	$\overline{U_{шЭСР}^2}$ , мкВ <sup>2</sup>	$\overline{U_{шхр}^2}$ , мкВ <sup>2</sup>	$\overline{U_{шотж}^2}$ , мкВ <sup>2</sup>	$\Delta_1$	$\Delta_2$	$\Delta_3$
1	14,48	14,61	14,46	14,48	-0,13	0,15	-0,02
2	14,19	14,30	14,29	14,26	-0,11	0,01	0,03
3	14,81	14,9	14,88	14,86	-0,09	0,02	0,02
4	14,68	14,76	14,79	14,79	-0,08	-0,03	0
5	14,4	14,65	14,41	14,34	-0,25	0,24	0,07
6	14,68	14,76	14,7	14,72	-0,08	0,06	-0,02
7	14,27	14,34	14,34	14,39	-0,07	0	-0,05
8	14,43	14,69	14,38	14,47	-0,26	0,31	-0,09
9	14,52	14,60	14,58	14,56	-0,08	0,02	0,02
10	14,48	14,64	14,64	14,66	-0,16	0	-0,02

Таблица 2

Результаты испытаний выборки из партии 2

Table 2

Test results of a lot sample 2

№ ИС	$\overline{U_{шнач}^2}$ , мкВ <sup>2</sup>	$\overline{U_{шЭСР}^2}$ , мкВ <sup>2</sup>	$\overline{U_{шхр}^2}$ , мкВ <sup>2</sup>	$\overline{U_{шотж}^2}$ , мкВ <sup>2</sup>	$\Delta_1$	$\Delta_2$	$\Delta_3$
1	14,23	14,48	14,32	14,33	-0,25	0,16	-0,01
2	14,28	15,07	14,99	14,45	-0,79	0,08	-0,54
3	14,15	14,29	14,30	14,26	-0,14	-0,01	0,04
4	14,11	14,48	14,48	14,22	-0,37	0	0,26
5	14,18	14,31	14,22	14,24	-0,13	0,09	-0,02
6	14,22	14,37	14,30	14,29	-0,15	0,07	0,01
7	14,4	14,65	14,66	14,59	-0,25	-0,01	0,07
8	14,29	14,32	14,30	14,29	-0,03	0,02	0,01
9	14,46	14,59	14,46	14,40	-0,13	0,13	0,06
10	14,21	14,35	14,34	14,36	-0,14	0,01	-0,02

Результаты измерения величин среднеквадратичного значения напряжения НЧШ, а также значения величин  $\Delta_1$ ,  $\Delta_2$ ,  $\Delta_3$  представлены в таблице 1 (для выборки из партии 1) и в таблице 2 (для выборки из партии 2).

Из таблиц видно, что для каждой ИС величины  $\Delta_1$ ,  $\Delta_2$ ,  $\Delta_3$  могут иметь как отрицательные, так и положительные значения, причем  $\Delta_1$  всегда отрицательная. Величины  $\Delta_2$ ,  $\Delta_3$  могут быть как положительными, так и отрицательными. В первой выборке общее число отрицательных значений величин  $\Delta_2$  и  $\Delta_3$  девять (ноль считаем за отрицательное значение). Во второй выборке общее число отрицательных значений величин  $\Delta_2$  и  $\Delta_3$  семь. Таким образом, общее число отрицательных значений величин  $\Delta_1$ ,  $\Delta_2$ ,  $\Delta_3$  в первой выборке больше, чем во второй. На основании этого делаем вывод, что вторая партия значительно надежнее первой.

#### Литература

1. Ван дер Зил А. Шум (источники, описание, измерение); [пер. с англ.; под ред. А.К. Нарышкина]. М.: Сов. радио, 1973. 228 с.

2. Разуменко Д. Низкочастотные шумы электронных компонентов как инструмент для диагностики внутренних дефектов // Компоненты и технологии. 2008. № 9. С. 168–174.

3. Врачев А.С. Низкочастотный шум – свойство диссипативных систем // Шумовые и деградационные процессы в полупроводниковых приборах: матер. докл. науч.-технич. сем. М., 1996. С. 191–197.

4. Горлов М.И., Сергеев В.А. Современные диагностические методы контроля качества и надежности полупроводниковых изделий. Ульяновск: Изд-во УлГТУ, 2014. 406 с.

5. Лукьянчикова Н.Б. Флуктуационные явления в полупроводниках и полупроводниковых приборах. М.: Радио и связь, 1990. 296 с.

6. Нарышкин А.К., Врачев А.С. Теория низкочастотных шумов. М.: Энергия, 1972. 152 с.

7. Mohammadi S., Pavlidis D. A Nonfundamental Theory of Low-Frequency Noise in Semiconductor Devices. IEEE: Transactions on electron devices, vol. 47, no. 11, November 2000.

8. Avery L.R. Electrostatic discharge: mechanism, protection techniques and effects on integrated circuit reliability. RCA Rev, 1984, no. 2, pp. 291–302.

9. Карба Л.П., Ульман Н.Н. О выборе шумовых параметров для прогнозирования отказов транзисторов // Электронная техника. 1978. Сер. 8. Вып. 7. С. 14–19.

10. Горлов М.И., Жуков Д.М., Клюкин А.А. Способ сравнительных испытаний по надежности партий интегральных схем: пат. 2546998 Российская Федерация. № 2012116017/28.

DOI: 10.15827/0236-235X.115.187-191

Received 29.05.16

#### EXTERNAL DESTABILIZING FACTORS INFLUENCE ON INTEGRATED CIRCUIT LOW-FREQUENCY NOISE

<sup>1</sup>Gorlov M.I., Dr.Sc. (Engineering), Professor, m-gorlov@inbox.ru; <sup>1</sup>Zhukov D.M. Postgraduate Student, ddimochka@mail.ru  
<sup>1</sup>Voronezh State Technical University, Moskovsky Ave. 14, Voronezh, 394026, Russian Federation

**Abstract.** The article presents test results of low-frequency noise measurement at chip pins depending on environment temperature, as well as with an electrostatic discharge impact. The data shows that low-frequency noise may be an informative parameter for integrated circuit diagnostics and grading by their reliability.

A low-frequency noise changes under the influence of external factors. The variation value depends on an integrated circuit state. The number of experiments conducted by the authors prove the fact. For example there is a cumulative effect of electrostatic discharge, i.e. each subsequent discharge affects a noise value stronger than the previous one. The dependence of low-frequency noise on a temperature says that the change in an integrated circuit internal state leads noise value changing. Such dependence gives reasons to believe that low-frequency noise may be an indicator of any external influence, which causes changing of integrated circuit internal state. In addition, the sensitivity of low-frequency noise as the informative parameter suggests its value changing during electrostatic defect annealing. As known from literature, after annealing the defects of electrostatic discharges are partially or completely eliminated. The noise change after annealing proves a direct correlation between its magnitude and an integrated circuit structure internal state.

Thus, the experiments, which show a direct bond between a noise value and an integrated circuit internal state, allow developing techniques for comparative tests and integrated circuit grading by reliability. The article presents a method for comparative assessment of quality and reliability of two or more integrated circuit lots.

**Keywords:** low-frequency noise, electrostatic discharge, integrated circuit, method of comparative assessment.

#### References

1. Van der Ziel A. *Noise*. Prentice-Hall Publ., 1954 (Russ.ed.: A.K. Naryshkin (Ed.), Moscow, Sov. radio Publ., 1973, 228 p.).
2. Razumenko D. Electronic components' low-frequency noise as a tool for internal fault diagnosis. *Komponenty i tekhnologii* [Components & Technologies]. 2008, no. 9, pp. 168–174 (in Russ.).
3. Vrachev A.S. Low-frequency noise is a property of dissipative systems. *Mat. dokl. nauch.-tekhn. sem. "Shumovye i degradatsionnye protsessy v poluprovodnikovyykh priborakh"* [Proc. Science and Technical Workshop "Noise and Degradation in Semiconductor Devices"]. Moscow, 1996, pp. 191–197.
4. Gorlov M.I., Sergeev V.A. *Sovremennye diagnosticheskie metody kontrolya kachestva i nadezhnosti poluprovodnikovyykh izdely* [Modern Diagnosis Methods for Quality and Reliability Control of Semiconductor Devices]. Ulyanovsk, UIGTU Publ., 2014, 406 p.
5. Lukyanchikova N.B. *Fluktuatsionnye yavleniya v poluprovodnikovyykh priborakh* [Fluctuating in Semiconductor Devices]. Moscow, Radio i svyaz Publ., 1990, 296 p.
6. Naryshkin A.K., Vrachev A.S. *Teoriya nizkochastotnykh шумов* [Theory of Low-Frequency Noise]. Moscow, Energiya Publ., 1972, 152 p.
7. Mohammadi S., Pavlidis D. A Nonfundamental Theory of Low-Frequency Noise in Semiconductor Devices. *IEEE: Transactions on Electron Devices*. 2000, vol. 47, no. 11.
8. Avery L.R. Electrostatic discharge: mechanism, protection techniques and effects on integrated circuit reliability. *RCA Rev*. 1984, no. 2, pp. 291–302.
9. Karba L.P., Ulman N.N. On the choice of noise parameters to predict transistor failure. *Elektronnaya tekhnika* [Electronic Engineering]. 1978, ser. 8, no. 7, pp. 14–19 (in Russ.).
10. Gorlov M.I., Zhukov D.M., Klyukin A.A. *Sposob sravnitelnykh ispytany po nadezhnosti party integralnykh skhem* [Reliability Comparative Tests for Integrated Circuit Lots]. Patent 2546998 of the Russian Federation no. 2012116017/28, 2015.

УДК 519.676: 681.51

Дата подачи статьи: 25.12.15

DOI: 10.15827/0236-235X.115.192-197

## ОБНАРУЖЕНИЕ НАРУШЕНИЙ ПРИ МНОГОМЕРНОМ СТАТИСТИЧЕСКОМ КОНТРОЛЕ ТЕХНОЛОГИЧЕСКОГО ПРОЦЕССА

(Работа выполнена в рамках задания Минобрнауки России № 2014/232 и гранта РФФИ № 15-48-02038)

В.Н. Клячкин, д.т.н., профессор, v\_kl@mail.ru;

Ю.А. Кравцов, аспирант

(Ульяновский государственный технический университет,  
ул. Северный Венец, 32, г. Ульяновск, 432027, Россия)

Статистический контроль многопараметрического технологического процесса основан на использовании карты Хотеллинга. Как правило, критерием нарушения стабильности процесса является выход точки за контрольную границу карты. При этом иногда некоторые нарушения процесса карта не обнаруживает. Для повышения эффективности многомерного статистического контроля предлагаются два подхода. Первый основан на использовании специальных структур на карте Хотеллинга: вероятность случайного появления такой структуры соизмерима с вероятностью ложной тревоги, поэтому ее наличие на карте свидетельствует о нарушении процесса. В качестве структур специального вида рассмотрены тренд, цикличность, резкие скачки, приближение множества точек к контрольной границе или к оси абсцисс. Второй подход базируется на применении предупреждающей границы: попадание нескольких точек подряд между предупреждающей и контрольной границами также свидетельствует о нарушении процесса. Разработана программа, реализующая соответствующие алгоритмы поиска. Исходные данные включают количество контролируемых параметров процесса, количество наблюдений в выборке, количество выборок, результаты измерений. В зависимости от степени коррелированности параметров их можно объединить в группы и строить карты для каждой группы. Рассмотрен численный пример: контроль десяти параметров – диаметров крышки датчика аэродинамических углов. С целью выявления реакции различных модификаций карты Хотеллинга на возможные нарушения были смоделированы три типа нарушений: смещение среднего уровня процесса по одному из показателей, тренд процесса по всему подмножеству показателей и увеличение рассеяния данных по одному из контролируемых показателей. Показано, что нарушения, которые «не видит» обычная карта Хотеллинга, обнаруживаются предложенными методами.

**Ключевые слова:** статистический контроль, карта Хотеллинга, диагностика нарушений, специальные структуры, предупреждающая граница.

Одним из направлений обеспечения качества продукции при серийном производстве является статистический контроль технологического процесса. При контроле независимых параметров исследуемого процесса стандарты предусматривают применение карт Шухарта и кумулятивных сумм. В случае контроля нескольких коррелированных параметров могут использоваться карты Хотеллинга и многомерных экспоненциально взвешенных скользящих средних, а также их модификации.

Многие работы зарубежных и отечественных авторов [1–3] рассматривают применение статистических методов для проведения контроля технологических процессов. Обнаружение нарушений процесса, как правило, проводится на основе выхода контролируемой статистики за заданные границы карты. Вместе с тем для одномерного контроля широко используются и другие критерии, в частности, поиск структур специального вида на карте Шухарта: если на карте имеет место структура (набор точек), вероятность появления которой близка к вероятности ложной тревоги, это свидетельствует о нарушении процесса.

Используются и карты с предупреждающими границами: попадание нескольких точек подряд между границами указывает на нарушение процесса. Современная компьютерная техника и ПО позволяют применить аналогичные подходы при многомерном статистическом контроле процесса.

### Постановка задачи

Контролируются  $p$  параметров технологического процесса  $X_1, X_2, \dots, X_p$ . Предполагается, что  $p$ -мерный параметр имеет многомерное нормальное распределение  $X \sim N_p(\mu, \Sigma)$  ( $\mu$  – вектор средних,  $\Sigma$  – ковариационная матрица). Результаты наблюдений анализируются с применением контрольной карты Хотеллинга [3–5]: для каждой  $t$ -й мгновенной выборки ( $t = 1, \dots, m$ ) рассчитывается статистика:

$$T_t^2 = n(\bar{X}_t - \mu_0)^T S^{-1}(\bar{X}_t - \mu_0), \quad (1)$$

здесь  $\bar{X}_t$  – вектор средних в мгновенных выборках,  $\bar{X}_t = (\bar{x}_{t1}, \dots, \bar{x}_{tp})^T$ ;  $\bar{x}_{tj}$  – среднее значение в  $t$ -й мгновенной выборке по  $j$ -му показателю ( $j = 1, \dots, p$ );  $\mu_0$  – вектор общих средних,  $\mu_0 = (\mu_1, \dots, \mu_p)^T$ ,  $\mu_j = \frac{1}{mn} \sum_{t=1}^m \sum_{i=1}^n x_{ijt}$ ;  $(2)$

$S$  – оценка ковариационной матрицы  $\Sigma$ ; компоненты ковариационной матрицы, определяющие рассеяние показателей качества и степень тесноты их связи, оцениваются по формуле

$$s_{jk} = \frac{1}{m(n-1)} \sum_{t=1}^m \sum_{i=1}^n (x_{ijt} - \mu_j)(x_{ikt} - \mu_k), \quad (3)$$

$j, k = 1, \dots, p$ .

Положение контрольной границы карты Хотеллинга определяется по формуле

$$T_{kp}^2 = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{1-\alpha}(p, mn-m-p+1), \quad (4)$$

где  $\alpha$  – уровень значимости (вероятность ложной тревоги);  $F_{1-\alpha}(k_1, k_2)$  – квантиль  $F$ -распределения Фишера с числами степеней свободы в числителе  $k_1$ , в знаменателе  $k_2$ .

По аналогии с широко распространенными картами Шухарта процесс считается стабильным, если значения статистики (1) не превышают критическое значение (4); Шухартом было рекомендовано в качестве признаков вероятной разладки процесса рассматривать попадание точки, соответствующей контролируемому параметру, в область, находящуюся вне контрольных границ. На практике по мере расширения использования этих карт все чаще возникали случаи, когда все точки находились в зоне между контрольными границами, но по графику было видно, что процесс протекает нестабильно. Например, исследуемый процесс имеет явную тенденцию к росту или снижению, носит периодический характер, присутствует размещение множества последовательных точек по одну сторону от центральной линии и т.д. Вследствие этого возникла необходимость расширить перечень критериев, с помощью которых можно визуально оценить стабильность процесса. С этой целью было предложено выявлять структуры специального вида, или так называемые неслучайные структуры [6–8], то есть совокупности точек, вероятность образования которых сопоставима с вероятностью ложной тревоги (таких, появление которых при нормальном ходе процесса маловероятно).

Еще один подход для повышения эффективности контрольных карт может быть основан на использовании на карте предупреждающей границы. Для одномерных карт средних значений используется стандарт, предусматривающий возможность применения предупреждающей границы в трех вариантах принятия решения: процесс считается нестабильным, если в зоне между предупреждающей и контрольной границами находятся две, три или четыре точки подряд. Аналогичный подход возможен и при многомерном контроле [9].

### Подсистема поиска специальных структур на карте Хотеллинга

Опыт использования контрольных карт Хотеллинга пока недостаточен, чтобы можно было четко сформулировать основные критерии нестабильности процесса. Тем не менее, очевидно, что некоторые из критериев, используемых для интерпретации карт Шухарта, применимы и для карты Хотеллинга. По аналогии с рекомендуемой стандартом методикой обнаружения структур специального вида, свидетельствующих о нарушении процесса, на карте Шухарта область между горизонтальной осью  $t$  на карте Хотеллинга и контрольной границей (Upper Control Limit)  $UCL = T_{kp}^2$  разделена на три зоны – А, В и С (рис. 1).

Существенный практический интерес может представлять анализ следующих структур специального вида, показанных на этом рисунке:

а) выход точки, соответствующей расчетному значению статистики Хотеллинга, за контрольную границу  $UCL$ ;

б) тренд (сколько точек подряд на возрастание или на убывание в зависимости от количества контролируемых параметров могут рассматриваться как специальная структура, то есть вероятность появления такого количества точек соизмерима с вероятностью ложной тревоги);

в) приближение к оси абсцисс (зона С: в диапазоне до  $1/3 T_{kp}^2$ ) (сколько точек, расположенных

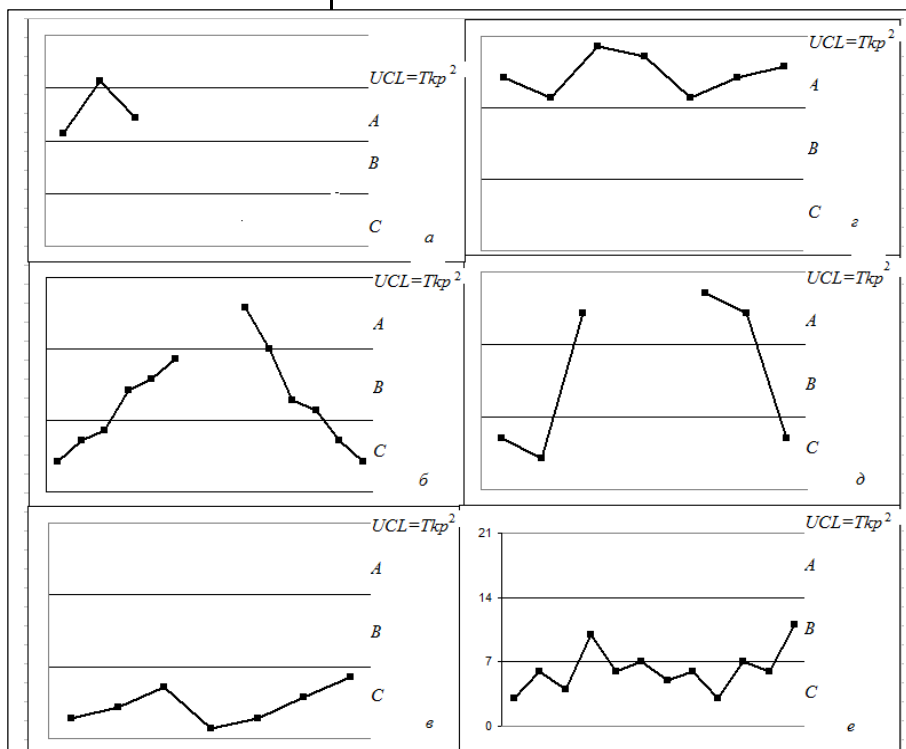


Рис. 1. Структуры специального вида на карте Хотеллинга

Fig. 1. A special form structure on the Hotelling's chart



Таблица 1

## Результаты расчета структур специального вида

Table 1

## Calculation data for special form structures

Структура	Количество контролируемых параметров						
	2	3	4	5	6	7	8
Тренд (кол-во точек)	6	6	6	6	7	7	7
Приближение к оси абсцисс (кол-во точек)	-	-	15	12	10	9	8
Приближение к контрольной границе (кол-во точек)	2 из 5	2 из 3	2 из 3 или 3 из 10	2 из 3 или 3 из 9	3 из 8 или 4 из 15	3 из 7 или 4 из 14	3 из 6 или 4 из 13
Резкие скачки, $k$	0,75	0,75	0,70	0,70	0,70	0,70	0,65
Цикличность (кол-во точек)	8	9	9	9	9	10	10

подряд в этой зоне, в зависимости от количества контролируемых параметров могут рассматриваться как структура специального вида);

г) приближение к контрольной границе (зона А);

д) резкие скачки на карте – на величину  $kT_{кр}^2$ , (при каком  $k$  при заданном количестве контролируемых показателей  $p$  соответствующий скачок на карте может рассматриваться как специальная структура);

е) цикличность (сколько точек, расположенных в шахматном порядке, могут рассматриваться как структура специального вида).

Результаты расчета структур, свидетельствующих о возможном нарушении процесса [9], при контроле от двух до восьми параметров представлены в таблице 1.

Из данной таблицы следует, что, например, при контроле шести параметров в качестве структур специального вида могут рассматриваться тренд из 7 точек на карте Хотеллинга, или 10 точек подряд вблизи оси абсцисс, или 9 циклически расположенных точек и т.п.

Поиск соответствующих структур проводится с применением специальных алгоритмов. На рисунке 2 в качестве примера показана блок-схема алгоритма обнаружения приближения точки на карте Хотеллинга к оси абсцисс.

### Подсистема построения карты с предупреждающей границей

Эффективность контрольных карт также может быть повышена путем использования предупреждающей границы, которое подразумевает разбиение области значений контролируемого параметра на три подобласти. Кроме обычных зон, характерных для стандартной карты Хотеллинга, при помощи предупреждающей границы выделяется дополнительная область для переходного состояния процесса (зона W). Расположение нескольких точек подряд (от 2 до 4) в этой области может свидетельствовать о возможном нарушении процесса.

Определим положение контрольной границы UCL и предупреждающей UWL таким образом,

чтобы попадание  $k$  точек подряд ( $k = 2, 3, 4$ ) в область W свидетельствовало о наличии заданного

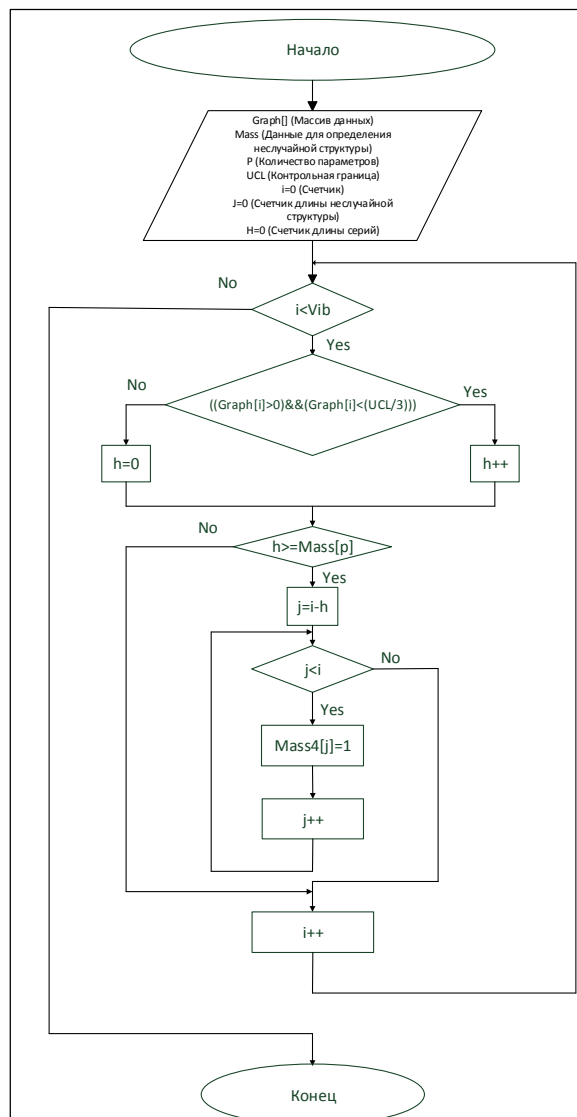


Рис. 2. Блок-схема алгоритма обнаружения приближения точки на карте Хотеллинга к оси абсцисс

Fig. 2. A flow chart of the approaching point detecting algorithm on the Hotelling's chart to the x-axis

опасного смещения среднего уровня процесса. Результат решения этой задачи методами теории цепей Маркова представлен в таблице 2, где дано положение контрольной UCL и предупреждающей UWL границ карты в зависимости от количества контролируемых показателей  $p$  и количества точек  $k$ , расположение которых подряд между границами карты свидетельствует о нарушении процесса на уровне значимости  $\alpha = 0,005$ , наиболее распространенном при многомерном контроле [9].

Таблица 2

**Положение контрольной и предупреждающей границ при 2, 3 и 4 точках подряд между границами**

Table 2

**The location of a control and warning limit with 2, 3 and 4 points in a row between limits**

Количество параметров	2 точки		3 точки		4 точки	
	UCL	UWL	UCL	UWL	UCL	UWL
2	12,39	5,70	13,29	3,60	14,10	2,55
3	14,89	7,44	15,35	5,12	15,60	3,87
4	17,82	8,84	18,29	6,31	18,63	5,95
5	19,48	10,52	20,79	7,70	21,19	7,64
6	21,77	11,79	22,69	8,87	23,00	8,79
7	23,74	13,27	24,58	10,15	24,98	9,91
8	25,72	14,74	26,26	11,43	26,83	10,83

### Программа обнаружения нарушений процесса

Программа «Анализ специальных структур на карте Хотеллинга» (рис. 3) предназначена для диагностики нарушений при многомерном статистическом контроле технологического процесса. Исходные данные включают количество контролируемых параметров, количество наблюдений в выборке, количество выборок, результаты измерений. В зависимости от степени коррелированности

параметров их можно объединить в группы и строить карты для каждой группы.

Программа включает две подсистемы: для обнаружения структур специального вида на карте Хотеллинга; для построения карты с предупреждающей границей с поиском последовательности заданного числа точек, расположенных между предупреждающей и контрольной границами. Программа обеспечивает диагностику процесса путем выявления соответствующих нарушений.

Разработанный программный продукт имеет следующие характеристики: тип операционной системы – Windows XP и выше; среда разработки – C#; использованы стандартные библиотеки MS Visual Studio 2013; размер программы – 3,42 Мб.

### Численный эксперимент

Осуществлялся контроль десяти параметров  $X_1, X_2, \dots, X_{10}$  – диаметры крышки датчика аэродинамических углов. Заданы допуски на изготовление:  $X_1 = 116^{+0,036}_{-0,071}$  ( $X_{1min}=115,929$  мм;  $X_{1max}=115,964$  мм),  $X_2 = 113^{+0,35}$  ( $X_{2min} = 113$  мм;  $X_{2max} = 113,35$  мм),  $X_3 = 110e8$  ( $X_{3min} = 109,928$  мм;  $X_{3max} = 109,982$  мм) и т.п. Для контроля предполагалось использовать многомерную карту Хотеллинга. Основная задача – проследить, как реагирует эта карта на возможные нарушения процесса.

На этапе анализа отлаженного процесса через равные промежутки времени делались замеры одной крышки (контроль по индивидуальным наблюдениям:  $n = 1$ ). По результатам контроля 30 крышек построена корреляционная матрица (на рис. 4 жирным шрифтом выделены значимые корреляции), из которой следует, что все множество параметров может быть разбито на три подмножества по наличию корреляций: 1)  $X_1, X_3, X_{10}$ ; 2)  $X_2, X_4, X_5, X_6$ ;

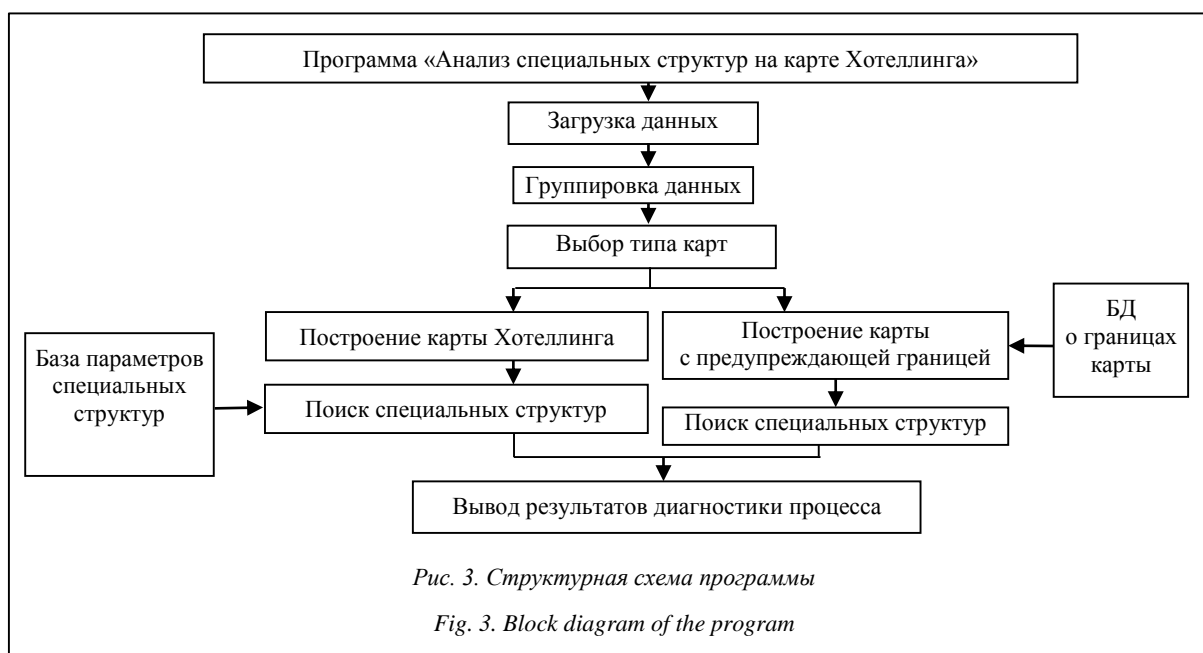


Рис. 3. Структурная схема программы

Fig. 3. Block diagram of the program

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
X1	1,00000	-0,00025	<b>0,87553</b>	0,18678	0,10196	0,17610	-0,10134	-0,11754	-0,09513	<b>0,81069</b>
X2	-0,00025	1,00000	-0,09148	<b>0,54980</b>	<b>0,37185</b>	<b>0,55407</b>	-0,23063	-0,27002	-0,25800	0,10967
X3	<b>0,87553</b>	-0,09148	1,00000	0,18343	0,14762	0,19114	-0,09289	0,03120	0,00991	<b>0,69853</b>
X4	0,18678	<b>0,54980</b>	0,18343	1,00000	<b>0,81090</b>	<b>0,78474</b>	-0,15765	-0,16410	-0,07633	0,20818
X5	0,10196	<b>0,37185</b>	0,14762	<b>0,81090</b>	1,00000	<b>0,77601</b>	-0,06318	-0,02507	0,03892	0,13071
X6	0,17610	<b>0,55407</b>	0,19114	<b>0,78474</b>	<b>0,77601</b>	1,00000	-0,17052	-0,25069	-0,00278	0,18427
X7	-0,10134	-0,23063	-0,09289	-0,15765	-0,06318	-0,17052	1,00000	<b>0,74280</b>	<b>0,61656</b>	0,04218
X8	-0,11754	-0,27002	0,03120	-0,16410	-0,02507	-0,25069	<b>0,74280</b>	1,00000	<b>0,63250</b>	-0,13864
X9	-0,09513	-0,25800	0,00991	-0,07633	0,03892	-0,00278	<b>0,61656</b>	<b>0,63250</b>	1,00000	-0,07224
X10	<b>0,81069</b>	0,10967	<b>0,69853</b>	0,20818	0,13071	0,18427	0,04218	-0,13864	-0,07224	1,00000

Рис. 4. Корреляционная матрица

Fig. 4. Correlation matrix

3)  $X_7$ ,  $X_8$ ,  $X_9$  (см. [2, 3]). По полученным данным были построены три карты Хотеллинга по каждому из рассматриваемых подмножеств: все три карты подтвердили стабильность процесса изготовления крышки. Основной целью этого этапа было определение положения контрольных границ карт для каждого подмножества параметров.

При последующем мониторинге процесса с помощью карт Хотеллинга с найденными на предыдущем этапе контрольными границами по результатам наблюдений за 60 деталями нарушений процесса обнаружено не было. При этом для диагностики возможных нарушений процесса были использованы как карты с предупреждающей границей [9], так и обычные карты Хотеллинга с анализом неслучайных структур [8].

С целью выявления реакции различных модификаций карты Хотеллинга на возможные нарушения были смоделированы три типа нарушений:

1) смещение среднего уровня процесса по одному из показателей (с 10-го наблюдения до 20-го все значения  $X_1$  увеличивались на некоторую постоянную величину, задаваемую в долях от стандартного отклонения:  $X'_{1,k} = X_{1,k} + a\sigma_1$ ;  $k = 10, 11, \dots, 20$ );

2) тренд процесса по всему подмножеству показателей; для моделирования этого нарушения каждое из значений  $X_1$ ,  $X_3$ ,  $X_{10}$ , начиная с 10-го наблюдения, последовательно увеличивалось на некоторую величину, например, для  $X_3$ :  $X'_{3,10} = X_{3,10} + 0,2\sigma_3$ ,  $X'_{3,11} = X_{3,11} + 0,4\sigma_3$  и т.п.;

3) увеличение рассеяния данных по одному из контролируемых показателей.

Первое из этих нарушений (смещение среднего уровня процесса) было выявлено с помощью карты Хотеллинга с предупреждающей границей (рис. 5), второе (тренд), как и ожидалось, проявилось в виде тренда на обычной карте Хотеллинга, третье (увеличение рассеяния) проявилось сразу в виде трех неслучайных структур: цикличности, резкого скачка и приближения значений статистики Хотеллинга к оси абсцисс (рис. 6). Заметим, что ни одно из смоделированных нарушений не проявилось бы, если бы была использована стандартная методика оценки статистической управляемости процесса, при которой нарушением процесса считается выход статистики Хотеллинга за контрольную границу. Как видно из результатов проведенных испытаний, выходов за контрольную границу при заданных нарушениях процесса нет.

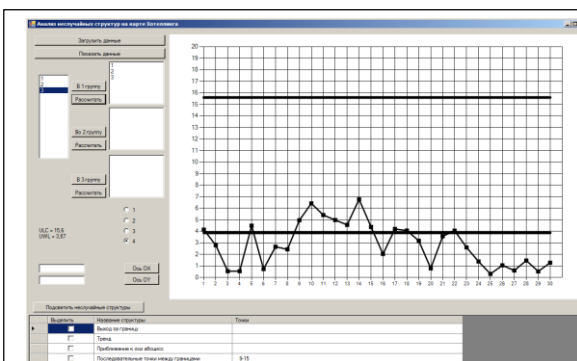


Рис. 5. Карта с предупреждающей границей обнаружила нарушение: точки 9–15 расположены между предупреждающей и контрольной границами

Fig. 5. A chart with warning limit has found violation: points 9–15 are located between warning and control limits

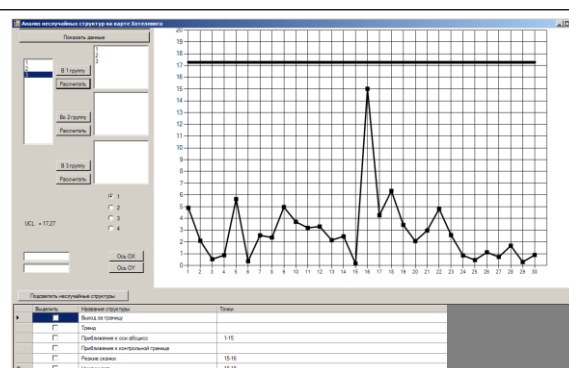


Рис. 6. Карта Хотеллинга обнаружила нарушение сразу по трем критериям

Fig. 6. The Hotelling's chart has detected the breach on the three criteria at once

Итак, в работе предложена программа для интерпретации контрольных карт, реализующая методы, повышающие чувствительность многомерного контроля к возможным нарушениям технологического процесса. Результаты численного эксперимента на примере контроля десяти параметров датчика аэродинамических углов с использованием разработанного ПО свидетельствуют об эффективности этих методов.

### Литература

1. Уилер Д., Чамберс Д. Статистическое управление процессами. Оптимизация бизнеса с использованием контрольных карт Шухарта. М.: Альпина Бизнес Букс, 2009. 409 с.
2. Айвазян С.А., Бухштабер В.М., Енюков И.С. Прикладная статистика // Классификация и снижение размерности. М.: Финансы и статистика, 1989. 607 с.

3. Montgomery D.C. Introduction to statistical quality control. NY: John Wiley and Sons, 2009, 754 p.
4. Ryan T.P. Statistical methods for quality improvement. NY, John Wiley and Sons, 2011, 687 p.
5. Bersimis S., Psarakis S., Panaretos J. Multivariate statistical process control charts: an overview. *Quality and Reliability Eng. Intern.*, 2007, vol. 23, pp. 517–543.
6. Robert F., Marilyn K. Criteria for lack of control. *Quality Management in Health Care*, 2003, vol. 12, iss. 1, pp. 5–19.
7. Юнак Г.Л., Годлевский А.Н., Плотников А.Н. Об интерпретации серий на контрольных картах // Методы менеджмента качества. 2005. № 4. С. 41–48.
8. Клячкин В.Н., Кравцов Ю.А. Диагностика состояния объекта по наличию неслучайных структур на контрольной карте // Приборы и системы. Управление, контроль, диагностика. 2013. № 5. С. 44–50.
9. Клячкин В.Н., Кравцов Ю.А. Повышение эффективности статистического контроля многопараметрического технологического процесса на основе карты Хотеллинга с предупреждающей границей // Автоматизация и современные технологии. 2013. № 10. С. 35–37.

DOI: 10.15827/0236-235X.115.192-197

Received 25.12.15

### IRREGULARITIES IN MULTIVARIATE STATISTICAL CONTROL OF A TECHNOLOGICAL PROCESS

(Acknowledgements. Work has been done within the framework of the task of the Ministry of Education and Science of the Russian Federation no. 2014/232 and the Russian Foundation for Basic Research grant no. 15-48-02038)

<sup>1</sup>Klyachkin V.N., Dr.Sc. (Engineering), Professor, v\_kl@mail.ru

<sup>1</sup>Kravtsov Yu.A., Postgraduate Student

<sup>1</sup>Ulyanovsk State Technical University,

Severnny Venec St. 32, Ulyanovsk, 432027, Russian Federation

**Abstract.** Statistical control of multiparameter technological process is based on using Hotelling's charts. Generally, a criterion for process stability violation is when a point exits a chart's control limit. Sometimes a card does not detect some violations of a process. The authors offer two approaches to improve the efficiency of multivariate statistical control. The first is based on using special structures on the Hotelling's chart. The probability of occurrence of such structure is commensurable with the probability of false alarm, so its presence on the chart reveals a violation of process. A trend, cyclicity, spikes, an approximation of a set of points to a control limit or to the X-axis is considered as a special type of structure.

The second approach is based on the warning limit. Several points between warning and control limits also constitute a violation of the process. The developed software package implements appropriate search algorithms. Source data include a number of controlled process parameters, a number of observations in a sample, a number of samples, measurement results. Depending on the degree of correlativeness of the parameters it is possible to combine them into groups and build charts for each group. As a numerical example, the authors examined 10 control parameters, which are aerodynamic angle sensor's cap diameters. In order to identify the reactions of Hotelling's charts various modifications on possible violations the authors modeled three types of violations: process middle-tier offset for one of the indicators, a process trend across a subset of indicators and scattering data increasing on any of the monitored targets. The article shows that violations which are not identified by the Hotelling's chart can be detected by the proposed methods.

**Keywords:** statistical control, Hotelling's chart, diagnosis of violation, special structures, the warning limit.

### References

1. Wheeler D.J., Chambers D.S. *Understanding Statistical Process Control*. SPC Press, Knoxville, Tennessee, 2009, 409 p.
2. Ayvazyan S.A., Bukhshtaber V.M., Enyukov I.S. *Prikladnaya statistika. Klassifikatsiya i snizhenie razmernosti* [Applied Statistics. Classification and Reduction in Dimensionality]. Moscow, Finansy i statistika, 1989, 607 p. (in Russ.).
3. Montgomery D.C. *Introduction to statistical quality control*. NY, John Wiley and Sons Publ., 2009, 754 p.
4. Ryan T.P. *Statistical methods for quality improvement*. NY, John Wiley and Sons Publ., 2011, 687 p.
5. Bersimis S., Psarakis S., Panaretos J. Multivariate Statistical Process Control Charts: An Overview. *Quality and reliability Engineering Int.* 2007, vol. 23, pp. 517–543.
6. Robert F., Marilyn K., Criteria for Lack of Control. *Quality Management in Health Care*. 2003, vol. 12, iss. 1, pp. 5–19.
7. Yunak G.L., Godlevsky A.N., Plotnikov A.N. On interpretation of series on control charts. *Metody menedzhmenta kachestva* [Quality management methods]. 2005, no. 4, pp. 41–48 (in Russ.).
8. Klyachkin V.N., Kravtsov Yu.A. Object's condition diagnostics using nonrandom structures on a control chart. *Pribory i sistemy. Upravlenie, kontrol, diagnostika* [Instruments and Systems: Monitoring, Control, and Diagnostics]. 2013, no. 5, pp. 44–50 (in Russ.).
9. Klyachkin V.N., Kravtsov Yu.A. Increase of efficiency of multivariable technological process statistical control on the basis of Hotelling's chart with a warning limit. *Avtomatizatsiya i sovremennye tekhnologii* [Automation and modern technologies]. 2013, no. 10, pp. 35–37 (in Russ.).

УДК 621.77

Дата подачи статьи: 13.11.15

DOI: 10.15827/0236-235X.115.198-201

## **МЕТОД ОРГАНИЗАЦИИ ИНФОРМАЦИОННОГО И ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ СИСТЕМЫ ТЕХНИЧЕСКОЙ ПОДГОТОВКИ ПРОИЗВОДСТВА ДЛЯ ПРОЦЕССОВ ГЛУБОКОЙ ВЫТЯЖКИ**

*(Работа выполнена при поддержке областного гранта «Разработка методологии технического перевооружения предприятий машиностроительного комплекса Оренбуржья на основе производственных систем нового поколения», соглашение № 36 от 23 июня 2015 г.)*

**А.И. Сердюк**, д.т.н., профессор, директор;

**М.В. Овечкин**, к.т.н., старший преподаватель, [tahov-1@mail.ru](mailto:tahov-1@mail.ru);

**М.А. Корнипаев**, к.т.н., доцент

*(Оренбургский государственный университет, аэрокосмический институт,  
просп. Победы, 13, г. Оренбург, 460018, Россия);*

**А.В. Щеголев**, генеральный директор

*(Механический завод, просп. Мира, 4, г. Орск, 462403, Россия)*

В статье описывается метод организации информационного и программного обеспечения автоматизированной системы технической подготовки производства для процессов глубокой вытяжки. Отмечено, что эффективность функционирования бизнес-процессов виртуальных предприятий, основанных на сложных процедурах управления и информационного обмена, обеспечивается весьма трудоемкой, выполняемой с использованием информационных и компьютерных технологий технической подготовкой производства, которая включает в себя конструкторскую, технологическую и организационно-экономическую подготовку производства. В этих условиях особую актуальность приобретают разработка и использование компьютерных средств для комплексной автоматизации системы технической подготовки производства – от разработки технологического процесса глубокой вытяжки, проектирования технологического инструмента, разработки управляющих программ для систем ЧПУ современного оборудования по его изготовлению до автоматизированного оформления конструкторской и технологической документации. На примере производства стальных баллонов высокого давления описаны компоненты, участвующие в автоматизированных процессах передачи информации о подготовке производства. Разработана и представлена схема эффективной организации информационного и программного обеспечения, состоящая из шести модулей, объединенных в четыре группы в зависимости от функционального назначения. Отмечено, что важным компонентом автоматизации производства вытяжных заготовок является программа DEFORM, использование которой на предприятии позволяет оценить форму заготовки на каждом шаге моделирования, определить возможность появления разрывов при свертке листовой заготовки, подобрать усилие вытяжки. Приведены формулы расчета количества вытяжек на основании степени деформации металла. Проиллюстрированы поэтапно сформированные модели вытяжных инструментов, экспортированные в программу КОМПАС-3D. Предложены пути дальнейшего развития метода, основанные на автоматизации изготовления вытяжного инструмента.

**Ключевые слова:** автоматизация, АСТПП, подготовка производства, информационное обеспечение, DEFORM.

Эффективность функционирования бизнес-процессов виртуальных предприятий, основанных на сложных процедурах управления и информационного обмена, обеспечивается весьма трудоемкой, выполняемой с использованием информационных и компьютерных технологий *технической подготовкой производства* (ТПП), которая включает в себя конструкторскую, технологическую и организационно-экономическую подготовку производства [1]. ТПП осуществляется в целях эффективного освоения нового или модернизированного изделия, внедрения новых сложных машин и оборудования, новых технологических приемов и изменений организации производства. Она предусматривает разработку технологических процессов, проектирование оснастки, приспособлений и специального инструмента, необходимых для обеспечения технологического процесса, расчеты производственных мощностей, подготовку нормативной документации [2].

С развитием компьютерных CAD/CAM/CAE/... САх-технологий системы ТПП все более интегри-

руются в единую *автоматизированную систему ТПП* (АСТПП), выполняющую роль верхнего уровня управления предприятием, обеспечивающего развитие ресурсов предприятия для выпуска все более совершенных изделий в нужных количествах в установленные сроки и с минимальной себестоимостью.

Требуется максимальная эффективность в работе подразделений, обеспечить которую возможно лишь с использованием АСТПП, связанных с разработкой конструкторской и технологической документации на необходимые изделия, а также быстрым перепрофилированием оборудования цехов и участков предприятия, занятых их изготовлением.

В этих условиях особую актуальность приобретают разработка и использование компьютерных средств для комплексной автоматизации АСТПП – от разработки технологического процесса глубокой вытяжки, проектирования технологического инструмента, разработки управляющих программ для систем ЧПУ современного оборудования по

его изготовлению до автоматизированного оформления конструкторской и технологической документации.

Организация информационного и программного обеспечения АСТПП формализованно может быть представлена в виде системы информационных потоков, то есть совокупности всех физических перемещений информации. Применительно к процессам глубокой вытяжки в качестве компонентов, участвующих в автоматизированных процессах передачи информации о подготовке производства, можно выделить следующие:

- модуль расчета параметров заготовки;
- модуль расчета параметров инструмента вырубki;
- банк данных;
- система моделирования процессов обработки металлов давлением DEFORM;
- модуль расчета параметров инструментов вытяжки;
- модуль расчета числа вытяжных операций.

Модель организации информационных потоков АСТПП процессов глубокой вытяжки на примере производства стальных баллонов высокого давления представлена на рисунке 1. Ее неотъемлемой

частью является система моделирования процессов обработки металлов давлением DEFORM, в которой составляется таблица с данными по параметрам вытяжных операций [3]. DEFORM – система конечно-элементного моделирования, предназначенная для анализа трехмерного течения металла при различных процессах обработки металла давлением, инструмент, позволяющий прогнозировать характер формообразования при операциях обработки металлов давлением без существенных затрат на экспериментальное исследование [4]. DEFORM является наиболее распространенной системой моделирования обработки металлов давлением. В России пользователями являются ГАЗ, Трубодеталь, Тяжпрессмаш и многие другие предприятия. Использование DEFORM позволяет оценить форму заготовки на каждом шаге моделирования, определить возможность появления разрывов при свертке листовой заготовки, подобрать усилие вытяжки. Результаты исследования, описанные в работе [5], показывают соответствие результатов, полученных с помощью математического моделирования в программе DEFORM и при проведении натурных испытаний. Кроме того, система позволяет провести моделирование с использованием инструментов, полученных на этапе автоматизированного расчета параметров вытяжных инструментов.

Одним из ключевых блоков представленной модели информационных потоков является банк данных. В его состав входят несколько БД, СУБД на базе MS Access, а также СУБД на Delphi с использованием механизмов ADO. Кроме того, в него могут входить данные в виде книжных и электронных справочников. Необходимо подготовить данные по механическим свойствам деформируемого металла заготовки. Их можно получить путем выборки из БД механических свойств используемых металлов посредством формирования SQL-запроса на основании указанного материала заготовки (либо его характеристик в виде удельного веса и коэффициента уточнения). При заполнении базы реко-



Рис. 1. Организация информационного и программного обеспечения системы ТПП баллонов

Fig. 1. Information support and software for a cylinder fitting-out system

мендуется использовать аппроксимацию для перевода графических данных по свойствам металлов в числовые характеристики. Кроме того, необходимо задать эвристику по подбору числовых параметров, точные значения которых отсутствуют в БД.

Банк данных, в свою очередь, осуществляет обмен информацией с блоком «Автоматизация расчета параметров технологического процесса вырубки», который состоит из модулей расчета параметров заготовки и инструмента вырубки. Основными составляющими алгоритма проектирования заготовки для глубокой вытяжки являются верное разбиение заготовки на составляющие элементарные объемы и четкий контроль механических свойств, опирающийся на сведения об изменении пределов прочности металла [6].

На выходе блока автоматизации расчета параметров технологического процесса вырубки, представленного внизу схемы, находится 3D-модель заготовки, вырубных матриц и пуансона, представляемых в формате КОМПАС-3D. Для повышения качества передачи 3D-модели из КОМПАС-3D предложено использовать систему Geomagic Studio, а для передачи данных о геометрии – формат STEP. Экспериментально установленные режимы настройки Geomagic Studio позволили экспортировать в формат STL в препроцессор DEFORM-3D геометрические модели пуансона, матрицы и прижима. Кроме того, блок связан с банком данных, представляющим собой автоматизированную информационную систему централизованного хранения и коллективного использования данных [7]. Выходными данными блока являются коэффициент зазора, диаметр матрицы, высота кромки и угол скалывания, 3D-модель инструментальной сборки в формате КОМПАС-3D, построенная по результатам автоматизации расчета.

Расчет инструментов вытяжки, являющийся частью функциональной составляющей блока автоматизации расчета параметров технологического процесса вырубки, состоит из расчета параметров инструментов и расчета количества вытяжек. При определении числа вытяжек следует учитывать, что с ростом степени холодной пластической деформации усиливаются прочностные свойства металла (увеличиваются пределы прочности и текучести, твердость), а пластические свойства ослабевают (уменьшаются относительное удлинение, сужение, ударная вязкость). Холодная пластическая деформация сопровождается искажением кристаллической решетки металла – образованием новых дислокаций, дроблением зерен, их сплющиванием и удлинением в направлении наибольшего течения металла. В результате искажений кристаллической решетки и появления остаточных напряжений изменяются физико-химические свойства металла, например, уменьшается электро- и теплопроводность. В результате холодной деформации в металле возникают также преимущественная ори-

ентировка (текстура) и анизотропия свойств, то есть их неоднородность в зависимости от направления преимущественного течения металла. При дискретизации данных, получаемых на основании эксперимента или аппроксимации на основе механических свойств материалов, следует проводить проверку порядка получаемых значений на соответствие формуле

$$n = \frac{\lg(1 - K_{\delta})}{\lg(1 - K_{\text{ср}})}, \quad (1)$$

где  $n$  – количество вытяжек;  $K_{\delta}$  – суммарная степень деформации;  $K_{\text{ср}}$  – средняя операционная степень деформации [8].

Кроме того, число вытяжек можно определить графическим путем на основании графических данных степени деформации при вытяжке, определяемой по формуле

$$K = \frac{F_0 - F_i}{F_0} \cdot 100\%, \quad (2)$$

где  $F_0$  – площадь поперечного сечения полуфабриката до вытяжки;  $F_i$  – площадь поперечного сечения полуфабриката после вытяжки [5].

Модули автоматизации расчета параметров вытяжных инструментов позволяют в автоматизированном режиме рассчитывать и строить вытяжные инструменты и на выходе получать поэтапные модели пуансонов и параметры матриц процесса вытяжки (рис. 2). Механические свойства материалов для данного этапа берутся из банка данных, параметры расчета могут приниматься на основе таблицы параметров расчетов, формируемой на основе результатов моделирования в DEFORM. Для автоматизированного определения числа вытяжек на основании деформации заготовки необходимо использовать БД, содержащую информацию по суммарному наклепу с определенной степенью дискретизации. Пользователь вводит параметры полуфабриката и выбирает материал из базы либо указывает коэффициент растяжения материала. Результатом работы модуля являются 3D-модели пуансонов, а также диаметры матриц и толщины стенок получаемых заготовок.

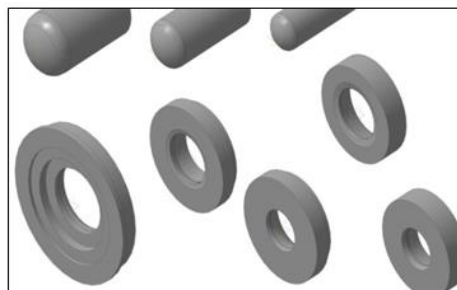


Рис. 2. Пример поэтапно сформированных моделей вытяжных инструментов

Fig. 2. The example of step-by-step formed models of drawing tools



В предложенной модели учтена возможность самостоятельного задания последовательности прохождения полуфабриката через матрицы в виде сценариев. Можно составить сценарий использования матриц «вручную», последовательно выбирая необходимые, а также загрузить его из файла.

Таким образом, в статье описана разработанная схема эффективной организации информационного и программного обеспечения АСПП баллонов. Выделено шесть модулей, объединенных в четыре группы по функциональному назначению: «Автоматизация расчета параметров технологического процесса вырубки», «Банк данных», «Система моделирования процессов обработки металлов давлением DEFORM», «Автоматизация расчета параметров вытяжных инструментов». Описаны потоки данных и даны рекомендации по связям между модулями и группами. Завершающим этапом автоматизации технической подготовки производства стальных баллонов для газов является автоматизация изготовления вытяжного инструмента. При его производстве рекомендуется использовать станки с ЧПУ, так как качество обработки напрямую влияет на процесс вытяжки.

## Литература

1. Схиртладзе А.Г., Скворцов А.В., Чмырь Д.А. Проектирование единого информационного пространства виртуальных предприятий. М.: Абрис, 2012. 615 с.
2. Медведева С.А. Основы технической подготовки производства: учеб. пособие. СПб: Изд-во СПбГУ ИТМО, 2010. 69 с.
3. Щеголев А.В., Овечкин М.В., Сергеев А.И. Обзор систем моделирования процессов обработки металлов давлением // V Междунар. науч.-практич. конф.: сб. матер. Краснодар, 2014. С. 119–121.
4. Паршин В.С., Кармышев А.П., Некрасов И.И., Пугин А.И., Федулов А.А. Практическое руководство к программному комплексу DEFORM-3D. Екатеринбург: Изд-во УрФУ, 2010. 266 с.
5. Смирнов-Аляев Г.А. Сопrotivление материалов пластическому деформированию. Л.: Машиностроение, 1978. 368 с.
6. Щеголев А.В., Овечкин М.В. Автоматизация расчетов кружка для операции глубокой вытяжки // Современные проблемы науки и образования. 2014. № 6. URL: <http://www.science-education.ru/120-15632> (дата обращения: 12.11.2015).
7. Наумов Д.М., Вальтер А.И. Методика применения баз данных в автоматизированных системах расчета процессов ОМД // Автоматизация: проблемы, идеи, решения: сб. тр. Тула: Изд-во ТулГУ, 2009. С. 27–30.
8. Валиев С.А. Комбинированная глубокая вытяжка листовых материалов. М.: Машиностроение, 1973. 176 с.
9. Щеголев А.В., Сердюк А.И., Сергеев А.И. Методика анализа процесса глубокой вытяжки // Науч.-технич. вестн. Поволжья. 2015. № 2. С. 232–234.

DOI: 10.15827/0236-235X.115.198-201

Received 13.11.15

## A METHOD OF ORGANIZING INFORMATION AND SOFTWARE FOR CAM PROCESSES DEEP DRAWING

(Acknowledgements. The work was done with the support of the regional grant "Development of a technical re-equipment methodology of Orenburg region machine-building enterprises on the basis of new generation production systems", agreement no. 36, June 23, 2015)

<sup>1</sup>Serdyuk A.I., Ph.D., Professor, Director; <sup>1</sup>Ovechkin M.V., Ph.D., Chief Lecturer, maxov-1@mail.ru;

<sup>1</sup>Kornipaev M.A., Ph.D., Associate Professor; <sup>2</sup>Shchegolev A.V., Director General

<sup>1</sup>Orenburg State University, Aerospace Institute, Pobedy Ave. 13, Orenburg, 460018, Russian Federation

<sup>2</sup>JSC «Mechanical plant», Mira Ave. 4, Orsk, 462403, Russian Federation

**Abstract.** This article describes a method of organizing information and software support for an automated fitting-out system of deep drawing processes. It is noted that the efficiency of business processes of virtual enterprises, which are based on complex management information exchange procedures, is provided by highly labour-intensive fitting-out system (CCI) using information and computer technologies. The fitting-out system includes design, technological, organizational and economic pre-production. In these conditions the development and use of computer facilities for complex automation of a fitting-out system is extremely important. The system includes a lot of procedures ranging from deep drawing technological process development, technological tool design, control programs development for CNC systems of modern equipment for its production to automated execution of design and technological documentation. On the example of steel high-pressure cylinders production the article describes the components involved in the automated processes of production information transfer. The authors developed and presented an effective organization chart of information and software support, which consists of six modules grouped into four groups by functional purpose.

It is noted that an important component of drawing pieces automated production is DEFORM software. It allows the company to assess the shape of a workpiece at each step of modeling, to determine the possibility of gaps occurrence at convolution the blanks, to choose a drawing force. The paper presents formulas for calculating a number of extracts on the basis of a metal deformation degree. It also illustrates stepwise generated model extraction tools that are exported to the program «КОМПАС-3D». The authors propose the ways of further development of the method based on automation of exhaust tool manufacturing.

**Keywords:** automation, CAM, preparat fitting-out, information supply, DEFORM.

## References

1. Skhirtladze A.G., Skvortsov V.A., Chmyr D.A. *Proektirovanie edinogo informatsionnogo prostranstva virtualnykh predpriyat* [Design of a Single Information Space of Virtual Enterprises]. Moscow, Outline Publ., 2012. 615 p.
2. Medvedeva S.A. *Osnovy tekhnicheskoy podgotovki proizvodstva* [Fundamentals of Technical Preparation of Production]. St. Petersburg, SPbSU ITMO Publ., 2010, 69 p.
3. Shchegolev A.V., Ovechkin M.V., Sergeev A.I. Overview of modeling systems of metal processing by pressure. *V Mezhdunar. nauch.-praktich. konf.: sb. mater.* [Proc. 5 Int. Science and Practice Conf.]. Krasnodar, 2014, pp. 119–121 (in Russ.).
4. Parshin V.S., Karmyshev A.P., Nekrasov I.I., Pugin A.I., Fedulov A.A. *Prakticheskoe rukovodstvo k programmnomu kompleksu DEFORM-3D* [A practical Guide to Program Complex DEFORM-3D]. Ekaterinburg, UrFU Publ., 2010, 266 p.
5. Smirnov-Alyayev G.A. *Soprotivlenie materialov plasticheskomu deformirovaniyu* [Resistance of Materials to Plastic Deformation]. Leningrad, Mashinostroenie Publ., 1978, 368 p.
6. Shchegolev A.V., Ovechkin M.V. Automation of circle calculations for deep drawing. *Sovremennye problemy nauki i obrazovaniya* [Modern Problems of Science and Education]. 2014, no. 6. Available at: <http://www.science-education.ru/120-15632> (accessed November 12, 2015).
7. Naumov D.M., Walter A.I. The method of databases application in automated systems for metal forming processes calculation. *Avtomatizatsiya: problemy, idei, resheniya* [Automation: Problems, Ideas, Solutions]. Tula, Tula State Univ. Publ., 2009, pp. 27–30 (in Russ.).
8. Valiev S.A. *Kombinirovannaya glubokaya vytyazhka listovykh materialov* [Combined Deep Drawing of Sheet Materials]. Moscow, Mashinostroenie Publ., 1973, 176 p.
9. Shchegolev A.V., Serdyuk A.I., Sergeev A.I. The method of deep drawing process analysis. *Nauchno-tekhnich. vestnik Povolzhya* [Scientific and Technical Volga region Bulletin]. 2015, no. 2, pp. 232–234 (in Russ.).

# FORECASTING OPTIMAL DURATION OF A BEER MAIN FERMENTATION PROCESS USING THE KALMAN FILTER

<sup>1</sup>Niyonsaba T., Master of Science, Student, [nitherence@mail.ru](mailto:nitherence@mail.ru);

<sup>1</sup>Pavlov V.A., Ph.D. (Military Science), Associate Professor

<sup>1</sup>Tver State Technical University, Nikitin Quay 22, Tver, 170026, Russian Federation

**Abstract.** One of the most important processes of beer production is the main process of fermentation. In this process, the wort transforms into beer. The quality of beer depends on the dynamics of wort parameters. The main fermentation process continues for 10 days and requires high costs.

Therefore, the main purpose of this article is to forecast the optimal duration of the beer main fermentation process and provide its optimal control. The Kalman filter can provide optimal control of the main fermentation process. It also estimates state variables of the main fermentation process taking into account the characteristics of random effects at the object's input and filters measuring noise. The initial values of raw materials and control parameters of the beer main fermentation process can be determined based on the predicted completion time of the main fermentation process.

**Keywords:** main fermentation, wort, Kalman filter, temperature, time, alcohol, extract.

Most technological processes are multifactorial, which is characterized by several parameters, and their occurrence depends on many factors. Dynamic changes in process parameters have an inconsistent nature. Moreover, random external influences are not always predictable and should be taken into account.

This article discusses the possibility of controlling the main fermentation process of beer (more precisely forecasting its optimal completion time) applying the Kalman filter.

**Statement of the problem.** The technological process of the main fermentation process in brewing manufacture is characterized by basic dynamic and threshold changes of parameters, such as the percentage of alcohol ( $A$ ), the extract ( $E$ ), temperature ( $T^0$ ) and ( $pH$ ) value [1]. The completion time of the product manufacturing process is the time when the basic parameters reach the specified (critical) values of ( $A < A_{cr}$ ,  $E > E_{cr}$ ,  $pH > pH_{cr}$ ). The temperature of the product must be in the interval of  $T_{cr \min}^0 \leq T^0 \leq T_{cr \max}^0$  ( $T_{cr \min}^0$ ,  $T_{cr \max}^0$  – are minimum and maximum product temperature values, at which the process is implemented).

Based on the experimental data analysis [2] the equations of relationship between the values of parameters and the process time is shown in formula (1) and the graphs of this relationships are shown in figure 1.

$$\begin{cases} -0,043t^3 + 0,35t^2 + 0,34t + 7,765 \leq 4, \\ -0,1725t + 5,793 \leq 4,3, \\ 0,0137t^3 - 0,1t^2 - 1,675t + 16,269 \leq 2,5, \\ -0,0116t^2 + 0,61t + 0,04 = 4,5, \end{cases} \quad (1)$$

where  $t \geq 0$ .

The process duration is defined as a maximum time to achieve specified (critical) values of all parameters. These critical values for the studied process are:  $T^0 = 4$ ;  $pH = 4,3$ ;  $E = 2,5$ ;  $A = 4,5$ .

It is possible to control the process duration only by changing the temperature. In the course of the process others parameters should change in accordance with the relevant rules and by the completion of the process to achieve their critical values.

The goal of an optimal process control is forecasting a minimum duration.

The basis of the optimal process control is forecasting the dynamics of the main parameters. It should be noted that the dynamics of the process parameters is affected by various random factors of natural origin (for example, the quality of product ingredients, deviation of ambient temperature from the set point, the accuracy observance of technological modes, etc.). Therefore, the process parameters' values  $A(t)$ ,  $E(t)$ ,  $pH(t)$  at each moment of time will be of a random nature. In this context, they are presented at each current time interval ( $T_i, t$ ) implementations of random functions in the process time interval  $T_i < t < T_f$ , where  $T_i$ ,  $T_f$  are initial and final, and  $t$  is the current time points of the process. Such functions are independent and non-stationary. It means that they change over time; their mathematical expectations are time-dependent and the correlation function is at the position of the interval  $[T_i, T_f]$  on the time axis. Therefore, forecasting the dynamics of the process parameters must be carried out

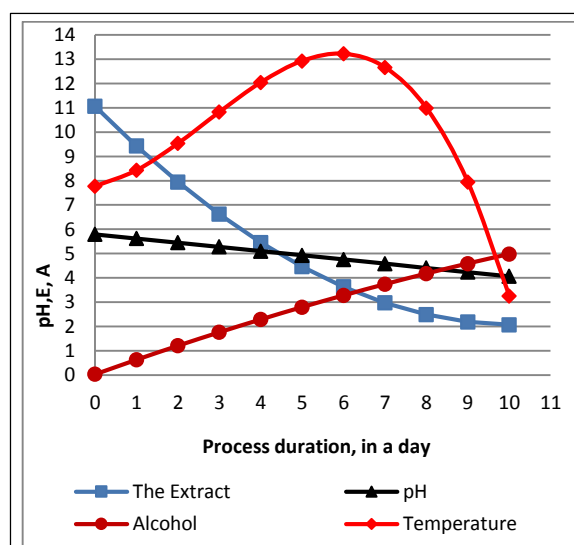


Fig. 1. The changes in parameters of the process

using the method of forecasting non-stationary random processes.

The forecasting parameters process  $A(t)$ ,  $E(t)$ ,  $pH(t)$  denote by  $x_i(t)$ , where  $i$  is a parameter index,  $i = 1, 2, 3$ , and present it in the form of a segment of a convergent series (in the sense of mean square convergence)

$$x_i(t) = \sum_{k=1}^N a_{ik} \frac{\varphi_{ik}(t)}{\sqrt{\lambda_{ik}}}, i = 1, 2, 3. \quad (2)$$

In an orthogonal basis of functions  $\varphi_{ik}(t)$ ,  $k=1, 2, 3, \dots, N$ , where  $a_k$  – are pair wise uncorrelated random variables with zero mathematical expectations,  $\varphi_k(t)$  – functions of own integral equations of the form;

$$\lambda_i \varphi_i(t) = \int_0^T k_i(t, \tau) \varphi_i(\tau) d\tau, \quad (3)$$

$\lambda_i$  – is its own values, independent and their correlation functions change relatively slowly in time. In this case for the evaluation of correlation functions it is possible to use the principle of moving smoothing the experimental data for each  $x_i(t)$  process parameter and write the calculated expression for them in the form

$$K_i(t, \tau) \triangleq K_i(t, t + \tau) = \int_{t-T/2}^{t+T/2} x_i(t') x_i(t' + \tau) dt'. \quad (4)$$

The best forecasting value of the corresponding parameter is calculated by the expression:

$$\hat{x}_i(t + \Delta t) = \sum_{k=1}^N a_{ik} \frac{\hat{\varphi}_{ik}(t + \Delta t)}{\sqrt{\lambda_{ik}}}, \quad (5)$$

where the sign  $\hat{\phantom{x}}$  refers to a continuation, which is an extrapolation  $x_i(t)$ ,  $\varphi_{ik}(t)$  to the time  $t + \Delta t$  – point-in-time extrapolation. Obviously, because of the deterministic basis functions  $\varphi_{ik}(t)$ ,  $k=1, 2, \dots, N$ , are written expressions for the derivatives of the predicted process.

However, there is no general method of solving the integral equation (3).

Therefore, we will consider the process of product manufacturing with its polynomial representation in non-orthogonal basis functions followed by the estimation of the coefficients of the polynomial non-stationary adaptive Kalman filter by “external” thermal management of product manufacturing.

It is known that the mathematical structure of a non-stationary filter is similar to the mathematical structure of the stationary Kalman filter taking into account external control [3, 4]. At the input of the filter there will be a constant measuring the parameters of the process values. For practical purposes it is important not to select an orthogonal basis set of functions  $1, t, t^2, \dots, t^n$  [5, 6].

On the basis of  $1, t, t^2, \dots, t^n$  extrapolation (forecasting) of the parameters for the process of product manufacturing we recorded process values and rate of changes. The results of the estimates obtained at the output of the Kalman filter are shown in the following expressions:

$$\hat{x}_{i3}(t + \Delta t) = \hat{x}_i(t) + \hat{\dot{x}}_i(t)(t + \Delta t) + \frac{1}{2} \hat{\ddot{x}}_i(t)(t + \Delta t)^2 + \dots$$

$$+ \frac{1}{n!} \hat{x}_i^{(n)}(t + \Delta t)^n, i = 1, 2, 3,$$

$$\hat{x}_{i3}(t + \Delta t) = \hat{x}_i(t) + \hat{\dot{x}}_i(t)(t + \Delta t) + \dots$$

$$+ \frac{1}{(n-1)!} \hat{x}_i^{(n-1)}(t + \Delta t)^{n-1}, i = 1, 2, 3, \quad (6)$$

$$x^{(n)}_{i3}(t + \Delta t) = \hat{x}_i^n(t), i = 1, 2, 3.$$

The sign  $\hat{\phantom{x}}$  denotes the estimation of appropriate values of the controlled process in the current time. At the same time the Kalman filter forms a variance-covariance matrix of extrapolation errors. Fundamentally, it is calculated by the expression:

$$\Psi_{i3} = M[\Delta \hat{x}_{i3} \Delta \hat{x}_{i3}^T], \quad (7)$$

where  $\Delta \hat{x}_{i3} = F_{i3} \Delta x_{i3}$  is error extrapolation  $t_{3i} = t + \Delta t$ ;

$$F_{i3} = \begin{bmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^n}{n!} \\ 0 & 1 & t & \dots & \frac{t^{n-1}}{(n-1)!} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

is the operator of extrapolation [7].

As a result we will obtain the necessary input data for determining the completion time of the manufacturing process of the product. Sufficient data must contain the necessary and valid threshold region. At the input the extrapolated values of the parameters (trajectories) set the completion time for the process of product manufacturing with specific parameters.

The Kalman filter is represented by the equations shown below.

The equation of state containing external control of Markov dynamics of the parameters of the process  $x(t)$ :

$$x(t+1) = \Phi(t+1, t)x(t) + w(t) + B(t)u(t). \quad (8)$$

The equation of observations. It measures the parameters of  $x(t)$  dynamics:

$$z(t) = x(t) + v(t). \quad (9)$$

The equation for evaluating extrapolation:

$$\hat{x}(t+1|t) = \Phi(t+1, t)\hat{x}(t) + \mu_w(t) + B(t)u(t) + k_{\mu}[z(t) - \mu_v(t) - \hat{x}(t)], \quad (10)$$

where  $\hat{x}(t)$  is the value of the dynamics of  $x(t)$  parameter at the next time of taking the measurement.

Assessment of the dynamics of  $x(t)$  parameters at the current point of time:

$$\hat{x}(t+1) = \hat{x}(t+1|t) + K(t+1)[z(t+1) - \mu_v(t) - \hat{x}(t+1|t)]. \quad (11)$$

The gains in the equation for evaluating extrapolation of the parameters for  $x(t)$  dynamics:

$$K_{\mu}(t) = V_{wv}(t)V_v^{-1}(t). \quad (12)$$

The gains in the equation for the estimation of the parameters of  $x(t)$  dynamic:

$$K_{\mu}(t+1) = V_{\bar{x}}(t+1|t)[V_{\bar{x}}(t+1|t) + V_x(t+1)]^{-1}. \quad (13)$$

The extrapolation error:

$$\hat{x}(t+1|t) \doteq \hat{x}(t+1) - \hat{x}(t+1|t). \quad (14)$$

The covariance error of extrapolation:

$$\begin{aligned} V_{\bar{x}}(t+1|t) &= [\Phi(t+1|t) - \\ &- K_{\mu}(t)]V_{\bar{x}}(t+1)[\Phi(t+1|t) - K_{\mu}(t)]^T + \\ &+ V_w(t) - K_{\mu}(t)V_v(t)K_{\mu}^T(t). \end{aligned} \quad (15)$$

Covariance of the estimation error:

$$V_{\bar{x}}(t+1) = [I - K(t+1)]V_{\bar{x}}(t+1). \quad (16)$$

Here we use the following notation:  $t=T^0, 1, 2, 3, \dots$ , for the  $i^{\text{th}}$  of the manufacturing process parameters of the product;

$$\begin{aligned} x &= \begin{pmatrix} r \\ \dot{r} \\ \ddot{r} \end{pmatrix}, \quad \Phi \equiv F = \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}, \quad V = \begin{pmatrix} r \\ \dot{r} \\ \ddot{r} \end{pmatrix}, \\ \mu_x &= \begin{pmatrix} M_r \\ M_{\dot{r}} \\ M_{\ddot{r}} \end{pmatrix}, \quad B = \begin{pmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & b \end{pmatrix}, \quad u = \begin{pmatrix} u_r \\ u_{\dot{r}} \\ u_{\ddot{r}} \end{pmatrix}, \quad Z \equiv z \quad (17) \end{aligned}$$

$w$  is Jamming effects in the manufacturing of product process (equation of state);  $v$  is Jamming effects in the equation of observation (the measurement of the  $x(t)$  dynamics parameters).

The filter equation in the notation introduced will be used to assess the parameter of  $x(t)$  dynamics for the given state equations as algebraic equations of the second order.

The Jamming effects  $w(t)$  and  $v(t)$  can't follow the normal Gaussian probability distribution. However, the recurrence relation of the Kalman filter remains valid and abnormal distributions  $w(t)$  and  $v(t)$  [9]:

$$\begin{aligned} \text{Cov}[w(t), w(k)] &= V_w(t)\delta(t-k), \\ \text{Cov}[v(t), v(k)] &= V_v(t)\delta(t-k), \\ \text{Cov}[w(t), v(k)] &= V_{wv}(t)\delta(t-k), \\ V_x(0) &= D[x(0)], \mu_v(t) = M[v(t)], \mu_w(t) = \\ &= M[w(t)] \end{aligned} \quad (18)$$

where  $M$  is the symbol of mathematical expectation

$$B = \begin{pmatrix} b_r & 0 & 0 \\ 0 & b_{\dot{r}} & 0 \\ 0 & 0 & b_{\ddot{r}} \end{pmatrix}.$$

Values are determined from physical considerations (if they are equal to 1, then there are not strengthened and not weakened).

Initial conditions:

$$u(0), \hat{x}(0) = \hat{x}(0|0) = \mu_x(0) = M[x(0)],$$

$$V_{\bar{x}}(0) = V_{\bar{x}}(0|0) = D[x(0)] = V_{\bar{x}}(0).$$

(Their values can be set based on the initial conditions of the process beginning; the result is a set of trajectories at each control point.)

The manufacturing process of the product is completed when all the observed parameters (in our case they are  $A$ ,  $E$  and  $pH$ ) reach the required values.

In that case under the optimal completion time of the product manufacturing process we will understand the moment in time at which the required probability of achieving all  $i^{\text{th}}$  process parameters required values, e.g. completion parameters. This probability is calculated according to the probability density [8]

$$\begin{aligned} P(\hat{x}_{i\beta}(t+\Delta t)|x_i(t+\Delta t)) &= \\ &= \frac{1}{\sqrt{2\pi\Psi_{\hat{x}_{i\beta}, \hat{x}_{i\beta}}}} \exp\left\{-\frac{1}{2}(\hat{x}_{i\beta}(t+\Delta t) - \right. \\ &- x_i(t+\Delta t))^T \Psi_{\hat{x}_{i\beta}, \hat{x}_{i\beta}}^{-1}(\hat{x}_{i\beta}(t+\Delta t) - x_i(t+\Delta t))), \\ &x_i(t+\Delta t) \in [\alpha_{i\beta}, \beta_{i\beta}], \end{aligned} \quad (19)$$

$\alpha_{i\beta}, \beta_{i\beta}$  are the extrapolated lower and upper boundaries of the interval values of the completion indicator of the  $i^{\text{th}}$  process parameter, calculated to a point in time  $t_3=t+\Delta t$  as the point of intersection of the extrapolated trajectory of the lower and upper boundary lines – the borders of the established region of the optimal values of the product completion indicator for the  $i^{\text{th}}$  parameter.  $\Psi_{\hat{x}_{i\beta}, \hat{x}_{i\beta}}$  is the variance of the extrapolated values of

$\hat{x}_{i\beta}(t+\Delta t)$  trajectory  $i^{\text{th}}$  product component in the process of its manufacturing, which suits the value of the first element of the first row in the matrix  $\Psi_{i\beta}$ . Normal density is adopted due to the fact that such density has maximum entropy compared to other densities with the same first two moments as normal density. In this case the obtained result will be the best guaranteed.

The estimated probability formula is written as:

$$\int_{\alpha_{i\beta}}^{\beta_{i\beta}} P(\hat{x}_{i\beta}(t+\Delta t)|x_i(t+\Delta t))dx_i(t+\Delta t). \quad (20)$$

The application of the Gaussian probability density is based on the linearity of the non-stationary Kalman filter, forming a trajectory of completion of the  $i^{\text{th}}$  process parameter. The completion intervals of the  $i^{\text{th}}$  parameter for a product set by solving inequalities [9].

$$\hat{x}_{i\beta}(t+\Delta t) \geq \alpha_{i\beta}, \hat{x}_{i\beta}(t+\Delta t) \leq \beta_{i\beta}, t_{i\alpha}, t_{i\beta}, \quad (21)$$

where  $t_{i\alpha}$  and  $t_{i\beta}$  are the points in time of reaching trajectory  $\hat{x}_{i\beta}(t+\Delta t)$  parameters of the product for the  $i^{\text{th}}$  parameter with the speed  $\hat{x}_{i\beta}(t+\Delta t)$ .

At the same time the expression  $\bigcap_{i=1}^3(t_{i\alpha}, t_{i\beta}) = \Delta t$  is

determined by the intersection of the intervals and sets the desired optimal time, as the time of completion of the product  $\tau_{opt} = \inf_{T_0 \leq t \leq T_n} \Delta T$  subject to achieve a desired probability of product availability.

However, it may be a situation in which  $\bigcap_{i=1}^3(t_{i\alpha}, t_{i\beta}) = \emptyset$ . In this case, the decision about the optimal completion time of the product manufacturing process obviously should be taken only on the basis of

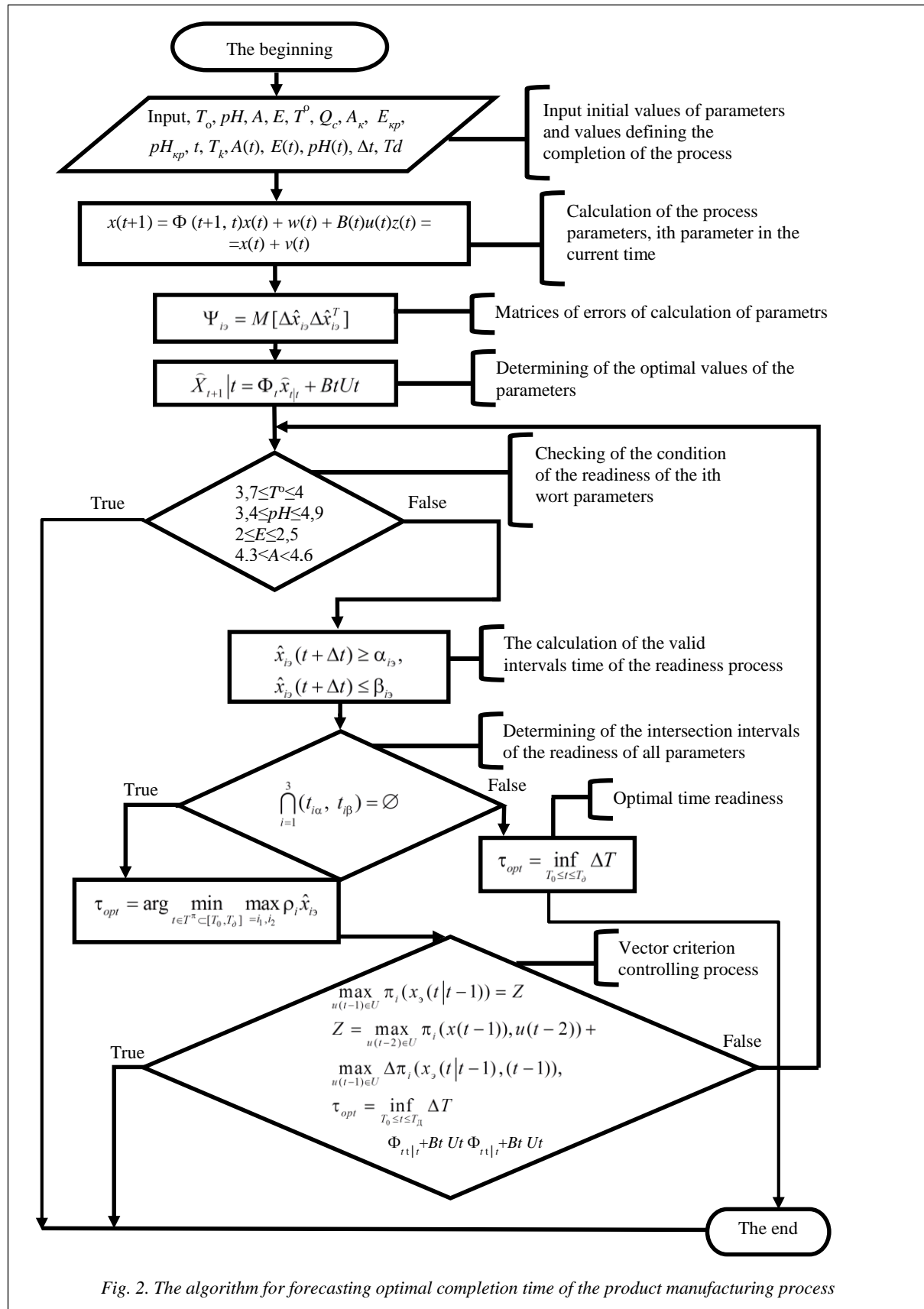


Fig. 2. The algorithm for forecasting optimal completion time of the product manufacturing process

the Pareto expression. The Pareto set is determined by the dynamics of the extrapolated trajectories of two dominant “conflicting” parameters of the manufac-

tured product at an acceptable time interval  $[T_0, T_D]$ . To calculate the product optimal completion time we will use the formula:

$$\tau_{opt} = \arg \min_{t \in T \subset [T_0, T_D]} \max_{i_1} \rho_{i_1} \hat{x}_{i_1}, \quad (22)$$

where  $\rho_i$  are the weights of trajectories' parameters of the product manufacturing process;  $\hat{x}_{i_1}$  are extrapolated values of the trajectories;  $T^\pi$  is the Pareto set of efficient points of completion time of the product manufacturing process;  $i_1, i_2$  are the dominant indexes "conflicting" parameters in the product manufacturing process.

On the basis of the physical nature of the product process control, a control criterion should be based on the requirement of product completion for the minimum time. The control criterion must be complex, e.g. vector, and consist of two individual criteria: one of them must reflect completion and the other must reflect the time spent on product manufacturing. For the criteria formation of the product completion we will introduce the utility function

$$\pi(\hat{x}_3(t|t-1), \hat{u}(t-1)) = \pi(\hat{x}(t-1), \hat{u}(t-2)) + \Delta\pi(\hat{x}_3(t|t-1), \hat{u}(t-1)). \quad (23)$$

The estimated Kalman filter of the trajectory of the product manufacturing  $i^{\text{th}}$  parameter to the current time and the extrapolated value of the trajectory at the current time depends on the decision maker. The choice of utility function depends on the decision maker with respect to the product completion. We proceed when the selected function from the terms of the decision maker disincline towards risk. In this case the utility function is nothing other than the desired criterion, which is the functional definition on the trajectory of the product manufacturing. Such a criterion is continuous, convex upwards, owing to the decision maker disinclination towards risk and, obviously, not additive. However, it is admitted by virtue of the Markov parameters of the trajectory of the product, its reduction to additive form at every current time  $t$  for each parameter. For this, we simply use the following expression:

$$\pi_i(\hat{x}_3(t|t-1), \hat{u}(t-1)) = \pi_i(\hat{x}(t-1), \hat{u}(t-2)) + \Delta\pi(\hat{x}_3(t|t-1), \hat{u}(t-1)),$$

$$\text{where } \Delta\pi_i(\hat{x}_3(t|t-1), \hat{u}(t-1)) = \pi_i(\hat{x}_3(t|t-1), \hat{u}(t-1)) - \pi_i(\hat{x}(t-1), \hat{u}(t-2)). \quad (24)$$

Now, taking into account the requirement to minimize the process time, it is possible to form a vector criterion process control of the product manufacturing as a maximum of the utility function of the product and the minimum time for its manufacturing can be written in the following form

$$\begin{aligned} \max_{u(t-1) \in U} \pi_i(\hat{x}_3(t|t-1), \hat{u}(t-1)) = \\ = \max_{u(t-2) \in U} \{ \pi_i(\hat{x}(t-1), \hat{u}(t-2)) + \\ + \max_{u(t-2) \in U} \Delta\pi_i(\hat{x}_3(t|t-1), (t-1)) \}, \\ \tau_{opt} = \inf_{T_0 \leq t \leq T_n} \Delta T. \\ \text{Or } \max_{u(t-1) \in U} \pi_i(\hat{x}_3(t|t-1), u(t-1)) = \end{aligned} \quad (25)$$

$$\begin{aligned} = \max_{u(t-2) \in U} \{ \pi_i(\hat{x}(t-1), u(t-2)) + \\ + \max_{u(t-1) \in U} \Delta\pi_i(\hat{x}_3(t|t-1), (t-1)) \}, \\ \tau_{opt} = \arg \min_{t \in T^\pi \subset [0, T_D]} \max_{i_1} \rho_{i_1} \hat{x}_{i_1}(t|t-1). \end{aligned} \quad (26)$$

It should be noted that, the first formed criteria is written as the solution of the Bellman equation and provides feedback to the Kalman filter for optimal control. The algorithm for forecasting optimal completion time of the product manufacturing process is shown in figure 2 [10].

Input initial values of parameters and values defining the completion of the process (fig. 3).

The algorithm is implemented in a high-level programming language C#, Microsoft Visual Studio 2015.

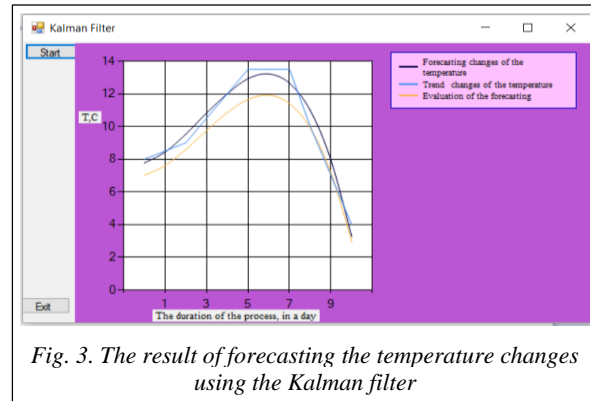


Fig. 3. The result of forecasting the temperature changes using the Kalman filter

**Conclusion.** The implementation of the above algorithm, the process control based on Kalman filter in an object-oriented automated control system can optimize the product manufacturing process at various combinations of the product ingredients and the parameters, under which the process is manufactured.

### References

1. Ermolaeva G.A., Kolchaeva R.A. *Tekhnologiya i oborudovanie proizvodstva piva i bezalkogolnykh napitkov* [Technology and equipment for beer and soft drinks production]. Study guide, Moscow, IRPO, Akademiya Publ., 2000, 416 p. (in Russ.).
2. Schuurman M., Brarudi S.A. *Sustainable Development Report 2010*. 2011, 16 p.
3. Baranov A.M., Osipov E.V. *Sintez filtra Kalmana s ispolzovaniem MatLab* [Synthesis of Kalman Filter Using Matlab]. Tambov, 2010 (in Russ.).
4. Welch G., Bishop G. *An Introduction to the Kalman Filter*. Univ. of North Carolina at Chapel Hill, NC, 2001.
5. Ogarkov M.A. *Metody statisticheskogo otsenivaniya parametrov sluchaynykh protsessov* [The Statistical Estimation Methods for Random Process Parameters]. Moscow, Energoatomizdat Publ., 1990, 208 p. (in Russ.).
6. Berezin I.S., Zhidkov N.P. *Metody vychisleniy* [Calculation Methods]. Moscow, FML Publ., 1962, vol. 1, 464 p. (in Russ.).
7. Kuzmin S.Z. *Osnovy teorii tsifrovoy obrabotki radiolokatsionnoy informatsii* [The Basics of the Radar Data Digital Processing Theory]. Moscow, Sov. radio Publ., 1974, 432 p. (in Russ.).
8. Sage A.P., Melsa J.L. *Estimation Theory with Applications to Communications and Control*. McGraw-Hill Publ., US, 1971, 752 p. (Russ.ed.: B.R. Levin, Moscow, Svyaz Publ., 1976, 496 p.).
9. Albert A. *Regressiya, psevdoinversiya i rekurrentnoe otsenivanie* [Regression, Pseudo-Inverse and Recurrent Estimation]. Moscow, Nauka Publ., 1977, 224 p.

УДК 658.5

DOI: 10.15827/0236-235X.115.202-207

Received 17.06.16

**ПРОГНОЗИРОВАНИЕ ОПТИМАЛЬНОЙ ДЛИТЕЛЬНОСТИ ПРОЦЕССА ГЛАВНОГО БРОЖЕНИЯ ПИВА  
НА ОСНОВЕ ФИЛЬТРА КАЛМАНА****Нийонсаба Т., магистр, nitherence@mail.ru;****Павлов В.А., к.в.н., доцент***(Тверской государственный технический университет, наб. Аф. Никитина, 22, г. Тверь, 170026, Россия)*

Одним из важнейших этапов в производстве пива является процесс главного брожения.

На этой стадии пивное сусло превращается в молодое пиво, параметры пивного сусла изменяются, и качество молодого пива зависит от динамики изменения этих параметров. Процесс главного брожения продолжается 10 дней, что сопряжено со многими затратами.

Основной целью данной статьи является исследование оптимального управления процессом главного брожения и обеспечение его минимальной продолжительности без ухудшения качества пива. Оптимальное управление процессом главного брожения выполняется фильтром Калмана.

Фильтр Калмана оценивает переменные состояния процесса главного брожения с учетом характеристик случайных воздействий на входе объекта и фильтрует измерительный шум.

На основе прогнозированного времени окончания процесса главного брожения определяются начальные параметры сырья и управление параметрами процесса главного брожения.

**Ключевые слова:** главное брожение, сусло, фильтр Калмана, температура, спирт, экстрактивность, продолжительность.

**Литература**

1. Ермолаева Г.А., Колчаева Р.А. Технология и оборудование производства пива и безалкогольных напитков. М.: Академия, 2000. 416 с.
2. Schuurman M., Brarudi S.A. Sustainable Development Report 2010. 2011, 16 p.
3. Баранов А.М., Осипов Е.В. Синтез фильтра Калмана с использованием MatLab // Сб. статей Тамбовского гос. техн. ун-та. 2006. Вып. 5. С. 115–118.
4. Welch G., Bishop G. An Introduction to the Kalman Filter. Univ. of North Carolina at Chapel Hill, NC, 2001.
5. Огарков М.А. Методы статистического оценивания параметров случайных процессов. М.: Энергоатомиздат, 1990. 208 с.
6. Березин И.С., Жидков Н.П. Методы вычислений. М.: Физматлит, 1962. Т. 1. 464 с.
7. Кузьмин С.З. Основы теории цифровой обработки радиолокационной информации. М.: Сов. радио, 1974. 432 с.
8. Сейдж Э., Мелс Дж. Теория оценивания и ее применение в связи и управлении; [пер. с англ.; под ред. проф. Б.Р. Левина]. М.: Связь, 1976. 496 с.
9. Алберт А. Регрессия, псевдоинверсия и рекуррентное оценивание. М.: Наука, 1977. 224 с.

---

---

**НИИ «Центрпрограммсистем»  
Разработка тренажеров, тренажерных комплексов  
и автоматизированных систем обучения**

Предлагаем комплексные решения создания систем теоретического обучения и тренажерной подготовки:

- автоматизированные системы обучения;
- специализированные тренажеры для индивидуальной и групповой подготовки специалистов по техническому обслуживанию и боевому использованию образцов вооружения, военной и специальной техники;
- комплексные тренажеры для подготовки расчетов боевых постов и командных пунктов различных уровней по решению тактических и оперативных задач;
- тренажеры по подготовке к управлению техническими средствами и борьбой за живучесть корабля;
- аппаратно-программный комплекс руководства обучением и комплексирования.

Предприятие имеет 25-летний опыт разработки учебно-тренировочных средств для Вооруженных сил, достаточную материально-техническую базу и необходимые лицензии.

Подробности на сайте [www.cps.tver.ru](http://www.cps.tver.ru)

Телефон: (4822) 44-32-44

Электронная почта: [info@cps.tver.ru](mailto:info@cps.tver.ru)



## **СИСТЕМА ДИАГНОСТИКИ И ОЦЕНКИ РИСКА ОСТЕОПОРОТИЧЕСКОГО ПЕРЕЛОМА НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ**

*Г.А. Дмитриев, д.т.н., профессор;*

*Аль-Факих Али Салех Али, аспирант, alfakih.ali@mail.ru*

*(Тверской государственный технический университет,*

*наб. Аф. Никитина, 22, г. Тверь, 170026, Россия)*

Использование информационных технологий в медицине при диагностике различного рода заболеваний требует совершенствования методов хранения и обработки данных. Для оценки риска остеопоротического перелома используется вычислительная модель, основанная на использовании схемы байесовского вывода. Задача прогнозирования рассматривается как задача классификации, то есть как задача нахождения апостериорной вероятности принадлежности пациента к одному из двух классов исходной классификации. Фактор, определяющий возможность остеопоротического перелома, является многомерной случайной величиной, оценка характеристик которой требует хранения и обработки больших объемов информации. В статье описываются состав и архитектура программного комплекса для диагностики остеопороза и оценки риска остеопоротического перелома. Комплекс включает в себя информационную и вычислительную компоненты. Вычислительная компонента содержит методы интеллектуального анализа данных, направленного на обнаружение закономерностей и тенденций, а также на выявление взаимосвязей существующих в многомерных массивах клинических данных. Информационная компонента содержит модели выборочных данных в виде многомерных кубов, которые формируются на базе OLAP-технологий и таблиц сопряженности. Объединение информационной и вычислительной компонент образует единую модель системы. Выходные данные информационной компоненты используются в качестве входа для вычислительной компоненты, которая служит для вывода суждений в условиях неопределенности и неполной информации. Ее основу составляют методы исчисления вероятностей, байесовы и нейронные сети. Рассмотрены вопросы использования служб Analysis Services: SQL Server Data Tools (SSDT) и SQL Server Management Studio в качестве платформы для создания и анализа многомерных моделей на основе технологий Data Mining.

**Ключевые слова:** информационная система, медицинская диагностика, остеопороз, байесовские сети.

Одним из самых распространенных хронических прогрессирующих метаболических заболеваний костной системы, которое характеризуется нарушением микроархитектоники костной ткани и усилением хрупкости по причине нарушения метаболизма костной ткани с преобладанием катаболизма над процессами костеобразования, является остеопороз [1–3]. Это социально-значимая патология, связанная не только с широкой распространенностью у населения старше 50 лет, но и с высокой частотой его тяжелых осложнений, среди которых наиболее неблагоприятными в прогностическом плане являются *остеопоротические переломы* (ОП). Ранняя диагностика заболевания и определение риска перелома позволяют выявлять лиц, имеющих высокий риск ОП, и выбрать наиболее подходящий препарат и стратегию лечения.

Развитие остеопороза прямо связано со снижением костной массы в единице объема и, соответственно, *минеральной плотности костной ткани* (МПКТ), определяющей прочность кости и ее устойчивость к чрезмерному физическому воздействию. МПКТ является важнейшим фактором, определяющим риск ОП. Для оценки МПКТ используется такой показатель, как Т-критерий:

$$T = \frac{y_i - m_y}{\sigma_y}, \text{ где } y_i - \text{значение МПКТ у } i\text{-го индивида; } m_y - \text{среднее значение МПКТ в популяции; } \sigma_y - \text{стандартное отклонение. Т-критерий означает}$$

количество стандартных отклонений выше или ниже среднего показателя пика костной массы. Т-критерий уменьшается при снижении костной массы и возрастает при ее увеличении. МПКТ определяется в абсолютных единицах как значение числа стандартных отклонений между МПКТ пациента и возрастной нормой для здоровых людей такого же возраста и пола.

Связь между низкой костной массой и повышенным риском переломов позволяет проводить инструментальную диагностику остеопороза и выявлять риск перелома, основываясь на МПКТ [1, 2]. Низкая МПКТ довольно эффективно может быть оценена с помощью целого ряда технологий, однако, как было показано, например, в [2, 3], МПКТ не в состоянии идентифицировать всех пациентов, у которых в будущем произойдет перелом. Это лишь один из факторов риска перелома, а для его более точной оценки необходимо использовать и другие, хорошо измеряемые факторы, которые давали бы дополнительную к МПКТ информацию. Кроме того, массовый скрининг на остеопороз, выполняемый путем измерения плотности костной ткани, из-за высокой стоимости исследования не является эффективным. Для выявления больных с высоким риском ОП или для предварительного скрининга тех, кому может потребоваться измерение плотности костной ткани, были разработаны такие методики, как «Вопросник для самооценки остеопороза», «Индекс риска остеопороза» [3–5].

В последнее время широкое применение нашел способ оценки риска перелома, получивший название FRAX (Fracture risk assessment tool) [6, 7]. Алгоритм FRAX дает обобщенную оценку риска перелома на основе имеющихся у пациента клинических факторов риска в сочетании (или без) с данными денситометрии. При прогнозировании десятилетней вероятности перелома по методике FRAX используются следующие клинические факторы риска: возраст, пол, индекс массы тела, предшествующий перелом, перелом бедра у родителей, семейный анамнез переломов, курение, прием глюкокортикоидов, ревматоидный артрит, вторичный остеопороз, злоупотребление алкоголем. Кроме того, может быть добавлен показатель минеральной плотности кости шейки бедра.

Тем не менее, инструмент FRAX, как и перечисленные выше, имеет определенные ограничения. Многие из факторов риска, используемых в инструменте FRAX, такие как курение сигарет, потребление алкоголя, прием глюкокортикоидов, перенесенные ранее переломы, оказывают дозозависимое влияние на риск перелома. Для этих факторов в инструменте FRAX используют относительные риски, основанные на усредненном влиянии фактора, что снижает достоверность выводов. Также эти инструменты не содержат целый ряд других показателей костной ткани, которые дают ценную информацию о риске перелома [8]. К ним относятся биохимические показатели метаболизма костной ткани, количественные показатели оценки кости с помощью ультразвукового метода и компьютерной томографии, а также измерения МПК в других областях скелета, уровни биохимических маркеров метаболизма костной ткани, риск падений, предшествующее медикаментозное лечение и ряд других. Указанные выше недостатки делают актуальными разработки новых методов и информационных технологий оценки риска ОП, позволяющие проводить диагностику на основе мониторинговых исследований, учитывающих как общую тенденцию, так и местную специфику.

### Модель оценки риска ОП

В системе используется модель оценки риска ОП, основанная на использовании схемы байесовского вывода [8–11]. Байесовский вывод – один из методов статистического вывода, в котором для уточнения вероятностных оценок на истинность гипотез при поступлении свидетельств используется формула Байеса:

$$p(y_j | X_1, X_2, \dots, X_n) = \frac{p(X_1, X_2, \dots, X_n | y_j) \times p(y_j)}{\sum_{j=1}^2 p(X_1, X_2, \dots, X_n | y_j) \times p(y_j)},$$

где случайная величина  $Y = (y_1, y_2)$  описывает состояние пациента – наличие или отсутствие перелома. Набор факторов риска определяет  $n$ -мерную случайную величину  $X$  с компонентами  $(X_1, \dots, X_n)$ .

Как уже было отмечено, развитие ОП прямо связано со снижением МПКТ, определяющей прочность кости. Кроме факторов риска переломов, связанных с чрезмерной потерей костной массы ( $X_1$ ), существуют факторы, не связанные с этим состоянием ( $X_2$ ). На основе проведенного анализа были отобраны факторы, определяющие риск ОП. С учетом разделения факторов на связанные ( $X_1$ ) и не связанные ( $X_2$ ) с чрезмерной потерей костной массы механизм их воздействия на частоту переломов  $Y$  можно представить в виде орграфа (рис. 1).

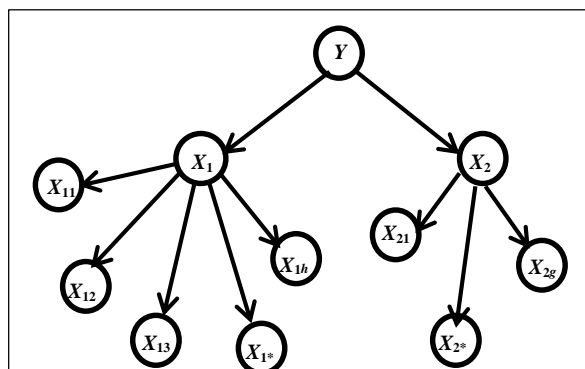


Рис. 1. Байесовская сеть для оценки риска ОП

Fig. 1. Bayesian network for OF risk assessment

В изображенной на рисунке 1 байесовской сети вершины представляют собой случайные величины, а дуги – вероятностные зависимости, которые определяются через таблицы условных вероятностей. Таблица условных вероятностей каждой вершины содержит вероятности состояний этой вершины при условии состояний ее родителей.

Задачу оценки риска перелома будем рассматривать как задачу нахождения апостериорной вероятности принадлежности пациента к одному из классов классификации –  $y_1$  или  $y_2$ . Апостериорная вероятность того, что пациент принадлежит к классу  $y_1$  (то есть существует вероятность перелома), имеет вид  $p(y_1 | X) = \frac{p(y_1)p(X | y_1)}{\sum_{j=1}^2 p(y_j)p(X | y_j)}$ .

Байесовская сеть строится на основе как экспертных оценок, так и статистических данных. Эти данные использовались и для установления взаимосвязей между случайными элементами, и для получения оценок условных вероятностей.

В общем случае вид плотности распределения и его параметры не известны, поэтому используются данные случайных независимых выборок. На основе этих данных строятся таблицы априорных вероятностей для вершины графа  $Y$  и условных вероятностей для вершин  $X_1$  и  $X_2$ :

$Y$	$p(y_i)$		$p(X_1 y_i)$	$p(X_2 y_i)$
$y_1$	$p(y_1)$		$p(X_1 y_1)$	$p(X_2 y_1)$
$y_2$	$p(y_2)$		$p(X_1 y_2)$	$p(X_2 y_2)$

Законы распределения случайных величин  $p(X_1)$  и  $p(X_2)$  вычислялись на основе дискретной логики модели множественного выбора:  $x_1^* = \sum \beta_j x_{1j} + \varepsilon$ ;  $x_2^* = \sum \beta_j x_{2j} + \varepsilon$ , где  $x_1^*$  и  $x_2^*$  – значения латентных переменных;  $x_{1j}$ ,  $x_{2j}$  – наблюдаемые значения факторных переменных. Используя  $p(X_1)$  и  $p(X_2)$  в качестве априорных вероятностей, находим их апостериорные значения  $p(X_1|X_{11}, X_{12}, \dots, X_{1h})$  и  $p(X_2|X_{21}, X_{22}, \dots, X_{2g})$ . Окончательно модель для вычисления вероятности ОП имеет вид  $p(Y|X_1|X_{11}, X_{12}, \dots, X_{1h}), (X_2|X_{21}, X_{22}, \dots, X_{2g}))$ .

Модель основана на анализе многомерной системы дискретных случайных величин (многомерной дискретной случайной величины) с неизвестным законом распределения. Закон распределения такой случайной величины задается в виде многомерной таблицы, в ячейках которой расположены вероятности совместного проявления событий. На основании этой таблицы могут быть вычислены все условные и маргинальные вероятности. Для оценки вероятностей неизвестного многомерного закона распределения на основе выборочных данных использовались таблицы сопряженности, имеющие структуру, аналогичную многомерной таблице, задающей закон распределения случайной величины. В качестве оценок вероятностей в таблице сопряженности использовались эмпирические частоты совместного проявления событий. Последующая обработка данных проводилась на основе логарифмической модели и методов измерения связей многовходовых таблиц сопряженности.

#### Организация вычислений в системе диагностики и оценки риска ОП

Система состоит из совокупности модулей, обеспечивающих хранение данных и их обработку. Данные анамнеза представляются в виде совокупности векторов  $(x_1, x_2, \dots, x_n)$ , содержащих значения различных количественных и качественных признаков по каждому пациенту. Массивы данных о состоянии  $M$  пациентов проходят предварительную обработку, которая заключается в их группировке и бинаризации. В результате этой обработки формируется таблица (см. таблицу), содержащая исходные амнестические данные по каждому пациенту для анализа.

На основе исходной таблицы (см. таблицу) формируется многовходовая таблица сопряженности в виде многомерного куба. Это позволяет применить технологию OLAP (online analytical processing) для аналитической обработки информации в различных срезах. Результатом обработки многомерных кубов являются, в частности, оценки вероятностей многомерного распределения. На основе этих данных и алгоритмов настройки байесовской сети вычисляются параметры модели оценки риска ОП. Обученная таким образом модель сохраняется в

библиотеке моделей и используется для оценки риска ОП для пациента в процессе дифференциальной диагностики.

Массив данных  
Data array

	Бинарная матрица							
	1 Пол	2 Куре- ние	3 Алко- голь	4 Пере- ломы	5 Арт- рит	6 Диа- бет2	7 Осте- опороз2	8 Зре- ние
1	1	0	0	0	1	1	0	1
2	1	0	0	0	1	0	0	0
3	1	0	0	0	1	0	0	0
4	0	1	0	1	0	0	1	0
5	1	0	0	0	1	0	0	0
6	1	0	0	0	1	1	0	0
7	0	1	1	1	0	0	1	0
8	1	1	0	0	1	1	0	1
9	1	1	0	0	0	1	1	1

В качестве инструмента разработки системы диагностики и оценки риска ОП была использована СУБД MS SQL Server, поскольку на ее базе имеется возможность хранить данные в виде многомерных кубов, создавать модели Data Mining (интеллектуального анализа данных) и применять их для поддержки принятия решений [12]. Среда MS SQL Server включает подсистему службы Analysis Service, которая содержит алгоритмы интеллектуального анализа данных и средства создания запросов к данным. Обширная библиотека алгоритмов службы Analysis Service содержит алгоритмы классификации, дерева принятия решений, регрессионного анализа, логистической регрессии, кластеризации, анализа взаимосвязей между различными атрибутами в наборе данных, анализа временных рядов нейронных сетей. Кроме того, в библиотеку алгоритмов можно включать собственные алгоритмы и связывать их с данными. Эта возможность была использована при создании программного комплекса для диагностики остеопороза и оценки риска ОП. Для этого разработанные на основе методов исчисления вероятностей и байесовых сетей алгоритмы, кратко описанные выше, были добавлены в экземпляр сервера в качестве новой функции интеллектуального анализа данных. Регистрация алгоритмов производилась путем добавления необходимых метаданных в INI-файл экземпляра службы Analysis Services.

Логическая архитектура программного комплекса для диагностики остеопороза и оценки риска ОП на платформе MS SQL Server приведена на рисунке 2.

Программный комплекс состоит из трех компонент: набора моделей интеллектуального анализа данных, структур интеллектуального анализа данных и клиентского приложения.

Модели интеллектуального анализа данных создаются на основе алгоритмов из библиотеки экземпляра службы Analysis Service. Для создания модели сначала описывается ее структура, а затем

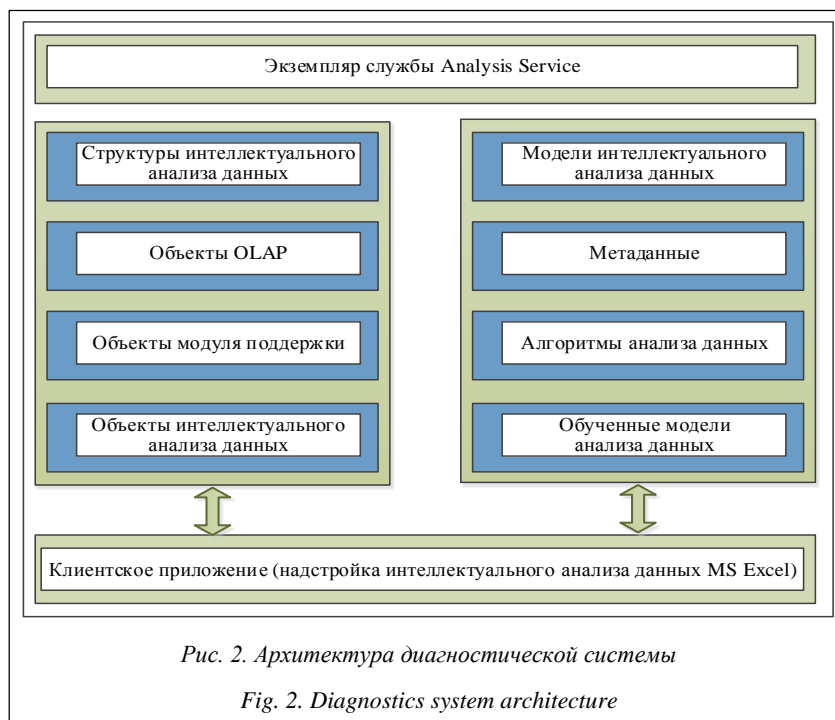


Рис. 2. Архитектура диагностической системы

Fig. 2. Diagnostics system architecture

определяются ее параметры на основе имеющихся данных. Этот процесс называется обучением. После обучения модель интеллектуального анализа данных содержит метаданные о модели, ссылку на алгоритм, который использовался для анализа данных, и результат анализа. Метаданные определяют имя модели и сервер, где она хранится, а также описание модели, включая данные, которые использовались для построения модели с привязкой к структуре интеллектуального анализа данных. Сама модель представляется стандартной структурой независимо от применяемого алгоритма и содержит свое полное описание. Обученная модель используется для вывода закономерностей и взаимосвязей, формирования прогнозов при получении новых данных и хранится в библиотеке моделей экземпляра службы Analysis Service. Источником данных как для обучения, так и для обученных моделей являются структуры интеллектуального анализа данных.

Структура интеллектуального анализа данных является хранилищем данных, на основе которых строятся модели интеллектуального анализа. Несколько моделей могут использовать одну и ту же структуру интеллектуального анализа данных. Структура и модель интеллектуального анализа данных являются отдельными объектами службы Analysis Service. В структуре интеллектуального анализа данных хранятся сведения, определяющие источник данных. Процесс создания структуры данных включает построение модели на основе OLAP, привязку модели к БД, загрузку в нее данных и присвоение разрешений на доступ к данным. Многомерные кубы являются базовыми объектами запросов. Администрирование БД осуществляется

с помощью среды SQL Server Management Studio. С ее помощью можно вносить изменения в БД.

Пользовательский интерфейс системы диагностики (рис. 2) реализован в среде MS Excel, выполняющей роль клиентского приложения. Использование надстроек интеллектуального анализа для пакета Excel позволяет проводить интеллектуальный анализ данных средствами Microsoft SQL Server. Чтобы использовать средства интеллектуального анализа таблиц для Excel, необходимо создать соединение с экземпляром служб Analysis Services. Это соединение обеспечивает доступ к алгоритмам интеллектуального анализа данных и БД. При подключении к БД служб Analysis

Services посредством клиентского приложения происходит подключение к многомерному кубу внутри этой БД.

Описанный подход в медицинской практике для поддержки принятия решений на стадии анализа медицинских данных позволяет использовать технологию клиент-сервер при формировании БД, необходимых для разработки моделей заболевания. Клиент интеллектуального анализа данных для Excel поддерживает активное соединение с сервером, обеспечивая масштабируемость системы. Таким образом, модель интеллектуального анализа может быть размещена как на локальном компьютере, так и на удаленных серверах БД. Это важно для организации работ по компьютерной диагностике остеопороза и прогнозированию ОП в медицинских учреждениях, поскольку источниками данных могут быть как отдельные медицинские учреждения, так и группы подобных учреждений, диагностические центры. Данные также могут поступать из других источников в ходе проведения мониторинговых исследований, например, из БД Министерства здравоохранения, Всемирной организации здравоохранения, Фонда медицинского страхования и т.п. На основе рассмотренной архитектуры можно реализовывать различные варианты системы диагностики – локальные, региональные и т.д., используя единую технологию их создания.

### Литература

1. Cadarette S.M. et al. Development and validation of the Osteoporosis Risk Assessment Instrument to facilitate selection of women for bone densitometry. CMA Journ., 2000, vol. 162, no. 9, pp. 1289–1294.

2. Cadarette S.M. et al. Evaluation of decision rules for referring women for bone densitometry by dual-energy x-ray absorptiometry. *Journ. AMA*, 2001, vol. 286, no. 1, pp. 57–63.
3. Cosman F., de Beur S.J., LeBoff M.S. et al. Clinician's guide to prevention and treatment of osteoporosis. *Osteoporos Int.*, 2014, vol. 25, no. 8, pp. 2359–2381.
4. Михайлов Е.Е., Беневоленская Л.И. Эпидемиология остеопороза и переломов: руководство по остеопорозу. М.: Бином. Лаборатория знаний, 2003. С. 10–55.
5. Richey F. et al. Validation and comparative evaluation of the osteoporosis self-assessment tool (OST) in a Caucasian population from Belgium. *QJM*, 2004, vol. 97, no. 1, pp. 39–46.
6. Канис Дж.А., Оден А., Йохансон Г., Боргстром Ф., Стром О., Макклоски И.В. FRAX – новый инструмент для оценки риска перелома: применение в клинической практике и пороговые уровни для вмешательства // Остеопороз и остеопатии. 2012. № 2. С. 38–44.
7. Лесняк О.М. Новая парадигма в диагностике и лечении остеопороза: прогнозирование 10-летнего абсолютного риска перелома (калькулятор FRAX) // Остеопороз и остеопатии, 2012. № 1. С. 23–28.
8. Рассел С., Норвиг П. Искусственный интеллект: современный подход. М.: Вильямс, 2006. 1408 с.
9. Терехов С.А. Введение в байесовы сети // Нейроинформатика-2003: науч. сессия МИФИ-2003. V Всерос. науч.-технич. конф. М.: Изд-во МИФИ, 2003. Ч. 1. С. 149–187.
10. Прокопчина С.В., Ветров А.Н., Нестеров А.О. Управление инвестиционными рисками строительных организаций на основе байесовских информационных технологий // Программные продукты и системы. 2014. № 1 (105). С. 212–216.
11. Еремеев А.П., Хазиев Р.Р., Зуева М.В., Цепенко И.В. Прототип диагностической системы поддержки принятия решений на основе интеграции байесовских сетей доверия и метода Демпстера–Шефера // Программные продукты и системы. 2013. № 1 (101). С. 11–16.
12. Службы Analysis Services. URL: [https://msdn.microsoft.com/ru-ru/library/ms174949\(v=sql.120\).aspx](https://msdn.microsoft.com/ru-ru/library/ms174949(v=sql.120).aspx) (дата обращения: 10.04.2016).
13. Шилин Д.Е., Шилин А.Д., Адамян Л.В. Существует ли у населения России связь между риском переломов по шкале FRAX (ВОЗ 2008) и потреблением кальция? // Остеопороз и остеопатии. 2010. № 1. С. 53–54.

DOI: 10.15827/0236-235X.115.208-212

Received 11.04.16

### THE ARCHITECTURE OF THE SYSTEM FOR OSTEOPOROTIC FRACTURE DIAGNOSTICS AND RISK ASSESSMENT

<sup>1</sup>Dmitriev G.A., Dr.Sc. (Engineering), Professor; <sup>1</sup>Al-Fakih Ali Saleh Ali, Postgraduate Student, [alfakih.ali@mail.ru](mailto:alfakih.ali@mail.ru)  
<sup>1</sup>Tver State Technical University, Nikitin Quay 22, Tver, 170026, Russian Federation

**Abstract.** The use of information technologies in medicine for diagnosis of various diseases needs improvements in data storage and processing. To assess the osteoporotic fracture risk the specialists use a computational model based on the Bayesian inference scheme. A prediction task is considered as a classification task, i.e. a task of finding the posterior probability of patient belonging to one of two original classification classes. The factor that determines the possibility of osteoporotic fracture is a multidimensional random variable. To evaluate its performance it is required to store and process large information volumes. The complex includes informational and computational components. The computational component contains methods of data mining aimed at detecting patterns and trends, as well as at identifying existing relationships in multidimensional arrays of clinical data. The informational component contains sample data models in the form of multidimensional cubes that are formed on the base of OLAP-technologies and contingency tables. Information and computer components are combined into a single system model. Data output of the information component is used as input for the computer component, which is used to display the statements under conditions of uncertainty and incomplete information. It is based on probability calculation methods and Bayesian networks. The article examines the use of Microsoft SQL Server Analysis Services as a platform to create and analyze multivariate models based on Data Mining technology.

**Keywords:** informational system, medical diagnostics, osteoporosis, Bayesian networks.

#### References

1. Cadarette S.M. Development and validation of the Osteoporosis Risk Assessment Instrument to facilitate selection of women for bone densitometry. *Canadian Medical Association Journ.* 2000, no. 162 (9), pp. 1289–1294.
2. Cadarette S.M. Evaluation of decision rules for referring women for bone densitometry by dual-energy x-ray absorptiometry. *Journ. of the American Medical Association*. 2001, vol. 286, no. 1, pp. 57–63.
3. Cosman F., de Beur S.J., LeBoff M.S. Clinician's Guide to Prevention and Treatment of Osteoporosis. *Osteoporosis international*. 2014, vol. 25, no. 8.
4. Mikhaylov E.E., Benevolenskaya L.I. *Epidemiologiya osteoporoza i perelomov. Rukovodstvo po osteoporozu* [Epidemiology of Osteoporosis and Fractures. Guidelines of Osteoporosis]. Moscow, Binom, Laboratoriya znany Publ., 2003.
5. Richey F. Validation and comparative evaluation of the osteoporosis self-assessment tool (OST) in a Caucasian population from Belgium. *QJM*. 2004, vol. 97, no. 1, pp. 39–46.
6. Kanis J.A., McCloskey E.V., Johansson H., Oden A., Ström O., Borgstrom F. FRAX – a new tool for the assessment of fracture risk: application in clinical practice and thresholds for intervention. *Osteoporoz i osteopatii* [Osteoporosis and Osteopathy]. 2012, no. 2, pp. 38–44.
7. Lesnyak O.M. The new paradigm in diagnosis and treatment of osteoporosis: prediction of a 10-year absolute risk of fracture (calculator frax). *Osteoporoz i osteopatii* [Osteoporosis and Osteopathy]. 2012, no. 1, pp. 23–28.
8. Russell S., Norvig P. *Iskusstvennyy intellekt: sovremennyy podkhod* [Artificial Intelligence: Modern Approach]. Vilyams Publ., 2006, 1408 p.
9. Terexhov S.A. Introduction to Bayesian networks. *Neyroinformatika-2003. Nauchnaya sessiya MIFI-2003. V Vseross. nauch.-tehnich. konf.* [Proc. 5th All-Russian Science and Technical Conf. "Neuroinformatics-2003". Scientific Session of MIFI]. Moscow, MIFI Publ., 2003, p. 1.
10. Prokopchina S.V., Vetrov A.N., Nesterov A.O. Investment risks management for construction organizations based on Bayesian information technologies. *Programmnye produkty i sistemy* [Software & Systems]. 2014, no. 1 (105), pp. 212–216 (in Russ.).
11. Eremeev A.P., Khaziev R.R., Zueva M.V., Tsapenko I.V. Prototype of the diagnostic decision support system on the basis of integration bayesian belief networks and the Dempster–Shafer method. *Programmnye produkty i sistemy* [Software & Systems]. 2013, no. 1 (103), pp. 11–16 (in Russ.).
12. Analysis Services. Available at: [https://msdn.microsoft.com/ru-ru/library/ms174949\(v=sql.120\).aspx](https://msdn.microsoft.com/ru-ru/library/ms174949(v=sql.120).aspx) (accessed April 10, 2016).
13. Shilin D.E., Shilin A.D., Adamyan L.V. Is there a link between a risk of fracture according to FRAX scale (ВОЗ 2008) and calcium consumption among the population of Russia? *Osteoporoz i osteopatii* [Osteoporosis and Osteopathy]. 2010, no. 1, pp. 53–54 (in Russ.).

# СОДЕРЖАНИЕ

<b>Посевкин Р.В., Бессмертный И.А.</b> Естественно-языковой пользовательский интерфейс диалоговой системы .....	5
<b>Полещук Е.А., Платонов А.В.</b> Использование вероятностного вывода в слабоформализованных базах знаний .....	10
<b>Сорокин В.Е.</b> Об эффективности наследования таблиц в СУБД PostgreSQL .....	15
<b>Тиханьичев О.В.</b> О некоторых проблемах предметной области поддержки принятия решений .....	24
<b>Фомин В.В., Мальцев С.В.</b> Вычислительные процедуры обработки неточной временной информации .....	29
<b>Полтавцева М.А., Хабаров А.Р.</b> Безопасность баз данных: проблемы и перспективы .....	36
<b>Дроботун Е.Б., Цветков О.В.</b> Построение модели угроз безопасности информации в автоматизированной системе управления критически важными объектами на основе сценариев действий нарушителя .....	42
<b>Дроботун Е.Б.</b> Синтез систем защиты автоматизированных систем управления от разрушающих программных воздействий .....	51
<b>Кочкаров А.А., Яцкин Д.В.</b> Алгоритм поиска оптимального расположения сенсоров для решения задачи мониторинга пространства .....	60
<b>Скаков Е.С., Малыш В.Н.</b> Пчелиный алгоритм оптимизации для решения задачи планирования беспроводной сети .....	67
<b>Письмак А.Е., Харитонова А.Е., Цопа Е.А., Клименков С.В.</b> Метод автоматического формирования семантической сети из слабоструктурированных источников .....	74
<b>Павлов А.И., Столбов А.Б.</b> Прототип системы поддержки проектирования агентов для имитационных моделей сложных систем .....	79
<b>Глазкова А.В.</b> Интеллектуальная система автоматического определения категории потенциальных адресатов текста .....	85
<b>Trinh Quang Duc, Nguyen Hoai Giang</b> A numerical simulation of multi-spectral photo-acoustic imaging toward hematological analysis .....	90
<b>Бутенко Д.В., Бутенко Л.Н., Большаков А.Л.</b> Поддержка принятия решений при выборе лекарственных препаратов на основе метода анализа иерархии .....	96
<b>Яровой С.В.</b> Агентный подход при моделировании лесопожарных ситуаций .....	101
<b>Зотин А.Г., Пахирка А.И., Дамов М.В., Савчина Е.И.</b> Улучшение визуального качества изображений, полученных в сложных условиях освещенности на основе инфракрасных данных .....	109

<b>Потапов А.Н., Абу-Абед Ф.Н., Мартынов Д.В., Угловский Е.П.</b> Реализация субъективных методов вероятностной оценки адекватности тренажеров на основе комплекса программ с привлечением экспертов и анкетированием обучаемых.....	121
<b>Потапов А.Н., Абу-Абед Ф.Н., Мартынов Д.В., Талалаев А.Б.</b> Комплекс программных средств для анализа причинно-следственных факторов конфликта применения тренажеров эрготехнических радиоэлектронных средств.....	129
<b>Вильвер П.Ю., Юрин А.Ю.</b> Специализированный программный комплекс «АСТ» для создания тренажеров.....	136
<b>Беляева М.А., Безотосова О.К.</b> Этапы разработки интегрированной информационно-аналитической системы обработки экономической информации в организациях .....	142
<b>Гончаров А.А., Семенов Н.А.</b> Аспектно-ориентированное программирование в контексте решения вопросов повышения эффективности экономических показателей IT-проектов.....	149
<b>Дмитриев В.М., Ганджа Т.В., Панов С.А.</b> Система виртуальных инструментов и приборов для автоматизации учебных и научных экспериментов .....	154
<b>Извозчикова В.В., Меженин А.В.</b> 3D-моделирование методов съемки мобильными видеосистемами .....	163
<b>Батчев С.А., Зайцев А.Г., Талалаев А.Б., Тимаков Д.А.</b> Метод обнаружения и сопровождения воздушных объектов по отраженным радиосигналам сторонних источников в пассивно-активных системах радиолокации .....	168
<b>Кочкаров А.А., Калинов И.А.</b> Программный комплекс пространственной навигации и мониторинга на основе алгоритма визуальной одометрии.....	175
<b>Чижикова Л.А.</b> Типовые функции программного обеспечения рабочего места инструктора авиационного тренажера .....	181
<b>Горлов М.И., Жуков Д.М.</b> Влияние внешних дестабилизирующих факторов на величину низкочастотного шума интегральных схем .....	187
<b>Клячкин В.Н., Кравцов Ю.А.</b> Обнаружение нарушений при многомерном статистическом контроле технологического процесса .....	192
<b>Сердюк А.И., Овечкин М.В., Корнипаев М.А., Щеголев А.В.</b> Метод организации информационного и программного обеспечения системы технической подготовки производства для процессов глубокой вытяжки.....	198
<b>Niyonsaba T., Pavlov V.A.</b> Forecasting optimal duration of the main fermentation process of beer using the Kalman filter .....	202
<b>Дмитриев Г.А., Аль-Факих Али Салех Али</b> Система диагностики и оценки риска остеопоротического перелома на основе интеллектуального анализа данных.....	208



## C O N T E N T

<b>Posevkin R.V., Bessmertny I.A.</b> Natural language user interface of a dialogue system .....	5
<b>Poleschuk E.A., Platonov A.V.</b> Probabilistic inference in weakly formalized knowledge bases .....	10
<b>Sorokin V.E.</b> About table inheritance effectiveness in PostgreSQL DBMS.....	15
<b>Tikhanychev O.V.</b> On some problems in the decision support domain .....	24
<b>Fomin V.V., Maltsev S.V.</b> Imprecise temporal information processing .....	29
<b>Poltavtseva M.A., Khabarov A.R.</b> Database security: problems and prospects .....	36
<b>Drobotun E.B., Tsvetkov O.V.</b> Modeling information security threats in the automated control system for crucial objects on the basis of attack scenarios .....	42
<b>Drobotun E.B.</b> Synthesis of protection systems of automated control systems against destroying program influence .....	51
<b>Kochkarov A.A., Yatskin D.V.</b> Search algorithm of optimal sensors location for solving the space monitoring problem .....	60
<b>Skakov E.S., Malysh V.N.</b> Bee colony optimization algorithm for wireless network planning problem solving .....	67
<b>Pismak A.E., Kharitonova A.E., Tsopa E.A., Klimenkov S.V.</b> Method of automatic generation of semantic network from semi-structured sources .....	74
<b>Pavlov A.I., Stolbov A.B.</b> A prototype of an agents design support system for complex system simulation models.....	79
<b>Glazkova A.V.</b> Intelligent system for automatic identification of text addressee category .....	85
<b>Trinh Quang Duc, Nguyen Hoai Giang</b> A numerical simulation of multi-spectral photo-acoustic imaging toward hematological analysis .....	90
<b>Butenko D.V., Butenko L.N., Bolshakov A.L.</b> Decision support when choosing medications based on the hierarchy analysis method.....	96
<b>Yarovoy S.V.</b> Forest fire simulation using an agent-based model .....	101
<b>Zotin A.G., Pakhirka A.I., Damov M.V., Savchina E.I.</b> Visual quality enhancement of images obtained in adverse illumination conditions based on infrared data .....	109

<b>Potapov A.N., Abu-Abed F.N., Martynov D.V., Uglovsky E.P.</b> Implementation of subjective methods of simulator adequacy probabilistic assessment based on a software system involving experts and questioning trainees .....	121
<b>Potapov A.N., Abu-Abed F.N., Martynov D.V., Talalayev A.B.</b> A software suite to analyze cause-effect factors of a conflict of ergotechnical radio-electronic equipment simulators .....	129
<b>Wilwer P.Yu., Yurin A.Yu.</b> "AST" software suite for creating training simulators.....	136
<b>Belyaeva M.A., Bezotosova O.K.</b> Development stages of integrated information analysis system of processing economic information in organizations .....	142
<b>Goncharov A.A., Semenov N.A.</b> Aspect-oriented programming as an approach to improve the efficiency of IT projects economic performance.....	149
<b>Dmitriev V.M., Gandzha T.V., Panov S.A.</b> VIP. The system of virtual instruments and devices for education and scientific experiment automation.....	154
<b>Izvozhikova V.V., Mezhenin A.V.</b> 3D modeling for mobile video systems .....	163
<b>Batchev S.A., Zaycev A.G., Talalaev A.B., Timakov D.A.</b> Method of detecting and tracking aerial objects by reflected radio signals from third-party sources in a passive-active radar system.....	168
<b>Kochkarov A.A., Kalinov I.A.</b> Software for spatial navigation and monitoring based on visual odometry algorithm.....	175
<b>Chizhikova L.A.</b> Typical functions of flight simulation device instructor station software .....	181
<b>Gorlov M.I., Zhukov D.M.</b> External destabilizing factors influence on integrated circuit low-frequency noise .....	187
<b>Klyachkin V.N., Kravtsov Yu.A.</b> Irregularities in multivariate statistical control of a technological process .....	192
<b>Serdyuk A.I., Ovechkin M.V., Kornipaev M.A., Shegolev A.V.</b> A method of organizing information and software for CAM processes deep drawing .....	198
<b>Niyonsaba T., Pavlov V.A.</b> Forecasting optimal duration of a beer main fermentation process using the Kalman filter .....	202
<b>Dmitriev G.A., Al-Fakih Ali Saleh Ali</b> The architecture of the system for osteoporotic fracture diagnostics and risk assessment .....	208