

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Национальный исследовательский
Томский политехнический университет»



ISSN 1684-8519

ИЗВЕСТИЯ

ТОМСКОГО ПОЛИТЕХНИЧЕСКОГО УНИВЕРСИТЕТА

Том 324, № 5, 2014

Информационные технологии

ИЗДАТЕЛЬСТВО  ТПУ

г. Томск

ИЗВЕСТИЯ ТОМСКОГО ПОЛИТЕХНИЧЕСКОГО УНИВЕРСИТЕТА

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Сонькин М.А., гл. редактор, д-р техн. наук (Россия)
Глазырин А.С., выпуск. редактор, канд. техн. наук (Россия)
Кириянова Л.Г., выпуск. редактор, канд. филос. наук (Россия)

Ресурсы Планеты

Савичев О.Г., редактор серии, д-р географ. наук (Россия)
Покровский О.С., канд. геол.-минерал. наук (Франция)
Старостенко В.И., д-р физ.-мат. наук (Украина)
Конторович А.Э., д-р геол.-минерал. наук (Россия)
Шварцев С.Л., д-р геол.-минерал. наук (Россия)

Математика, физика и механика

Никитенков Н.Н., редактор серии, д-р физ.-мат. наук (Россия)
Силкин В.М., канд. физ.-мат. наук (Испания)
Коротеев Ю.М., д-р физ.-мат. наук (Россия)
Уленев О.Н., д-р физ.-мат. наук (Россия)
Борисов А.М., д-р физ.-мат. наук (Россия)
Рождкова С.В., д-р физ.-мат. наук (Россия)

Химия и химические технологии

Коршунов А.В., редактор серии, д-р хим. наук (Россия)
Пестряков А.Н., д-р хим. наук (Россия)
Тойпель У., Dsc (Германия)
Джин-Чун Ким, Dsc (Южная Корея)
Ильин А.П., д-р физ.-мат. наук (Россия)

Техника и технологии в энергетике

Заворин А.С., редактор серии, д-р техн. наук (Россия)
Ханьялич К., Dsc. (Нидерланды)
Кабов О.А., д-р физ.-мат. наук (Россия)
Маренго М., PhD (Италия)
Алексеев С.В., д-р физ.-мат. наук (Россия)
Воропай Н.И., д-р техн. наук (Россия)

Информационные технологии

Кочегуров А.И., редактор серии, канд. техн. наук (Россия)
Хамкер Ф.Х., Dsc (Германия)
Руи Д., PhD (Португалия)
Муравьев С.В., д-р техн. наук (Россия)
Спицын В.Г., д-р техн. наук (Россия)

Социально-гуманитарные технологии

Барышева Г.А., редактор серии, д-р экон. наук (Россия)
Завадскас Э.К., PhD, Dsc (Литва)
Касати Ф., PhD (Италия)
Грюнвальд А., PhD (Германия)
Бутовская М.Л., д-р ист. наук (Россия)
Головаха Е.И., д-р филос. наук (Украина)

Входит в Перечень ВАК РФ – ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней доктора и кандидата наук.

Подписной индекс по каталогу
агентства «Роспечать» – 18054

УВАЖАЕМЫЕ ЧИТАТЕЛИ!

Журнал «Известия Томского политехнического университета» (Известия ТПУ) – рецензируемый мультидисциплинарный журнал, предназначенный для ученых, исследующих специфику развития техники и технологии в различных научных сферах.

Журнал издаётся с 1903 года. Учредителем является Томский политехнический университет. Полнотекстовый доступ к электронной версии журнала возможен на сайтах www.elibrary.ru, scholar.google.com

Журнал зарегистрирован Министерством Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций – Свидетельство ПИ № 77-16615 от 24 октября 2003 г.

ISSN 1684-8519

Импакт-фактор РИНЦ 2012 – 0,174

Особое внимание в журнале уделяется ключевым проблемам развития современной техники и технологии, что выражается в тематических разделах журнала, охватывающих следующие аспекты науки и техники:

- Ресурсы Планеты
- Математика, физика и механика
- Химия и химические технологии
- Техника и технологии в энергетике
- Информационные технологии
- Социально-гуманитарные технологии

К публикации принимаются статьи, ранее нигде не опубликованные и не представленные к печати в других изданиях. Не принимаются литературные обзоры, комментарии, рецензии и т. д.

Статьи, отбираемые для публикации в журнале, проходят закрытое (слепое) рецензирование.

Автор статьи имеет право предложить двух рецензентов по научному направлению своего исследования.

Окончательное решение по публикации статьи принимает главный редактор журнала.

Все материалы размещаются в журнале «Известия ТПУ» на бесплатной основе.

Журнал издаётся ежемесячно.

К публикации в серию «Информационные технологии» принимаются оригинальные по содержанию и ранее не опубликованные статьи, содержащие теоретические и экспериментальные результаты решения фундаментальных и прикладных задач по следующим основным направлениям: управление техническими, социальными и информационными системами; математическое моделирование технологических процессов и автоматизация производств; оптимизация процессов принятия решений; интеллектуальные информационно-телекоммуникационные системы и технологии передачи данных; прикладная математика; математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей; цифровая обработка сигналов и изображений; когнитивные программно-аппаратные комплексы с высокопроизводительной обработкой данных и беспроводные телекоммуникационные системы и технологии.

Ministry of Education and Science of the Russian Federation
National Research Tomsk Polytechnic University



ISSN 1684-8519

BULLETIN

OF TOMSK POLYTECHNIC UNIVERSITY

Volume 324, № 5, 2014

IT Technologies

TPU  PUBLISHING

Tomsk

BULLETIN OF TOMSK POLYTECHNIC UNIVERSITY

EDITORIAL BOARD

Sonkin M.A., editor in chief, Dr. Sc. (Russia)
Glazyrin A.S., managing editor, Cand. Sc. (Russia)
Kiryanova L.G., managing editor, Cand. Sc. (Russia)

The Planet Resources

Savichev O.G., *section editor*, Dr. Sc. (Russia)
Pokrovsky O.S., Cand. Sc. (France)
Starostenko V.I., Dr. Sc. (Ukraine)
Kontorovich A.Eh., Dr. Sc. (Russia)
Shvartsev S.L., Dr. Sc. (Russia)

Mathematics, Physics and Mechanics

Nikitenkov N.N., *section editor*, Dr. Sc. (Russia)
Silkin V.M., Cand. Sc. (Spain)
Koroteev Yu.M., Dr. Sc. (Russia)
Ulenekov O.N., Dr. Sc. (Russia)
Borisov A.M., Dr. Sc. (Russia)
Rozhkova S.V., Dr. Sc. (Russia)

Chemistry and Chemical Technology

Korshunov A.V., *section editor*, Cand. Sc. (Russia)
Pestryakov A.N., Dr. Sc. (Russia)
Teipel U., Dsc (Germany)
Jin-Chun Kim, Dsc (South Korea)
Ilyin A.P., Dr. Sc. (Russia)

Power Engineering

Zavorin A.S., *section editor*, Dr. Sc. (Russia)
Hanjalic K., Dsc. (Netherlands)
Kabov O.A., Dr. Sc. (Russia)
Marengo M., PhD (Italy)
Alekseenko S.V., Dr. Sc. (Russia)
Voropai N.I., Dr. Sc. (Russia)

IT Technologies

Kochegurov A.I., *section editor*, Cand. Sc. (Russia)
Hamker F.H., Dsc (Germany)
Rui D., PhD (Portugal)
Muravyov S.V., Dr. Sc. (Russia)
Spitsyn V.G., Dr. Sc. (Russia)

Humanities and Social Sciences Technologies

Barysheva G.A., *section editor*, Dr. Sc. (Russia)
Zavadskas E.K., PhD, DSc (Lithuania)
Casati F., PhD (Italy)
Grunwald A., PhD (Germany)
Butovskaya M.L., Dr. Sc. (Russia)
Golovakha E.I., Dr. Sc. (Ukraine)

AIMS AND SCOPE

Bulletin of Tomsk Polytechnic University (TPU Bulletin) is peer-reviewed multidisciplinary journal owned by Tomsk Polytechnic University. The Journal focuses on the key problems of the development of modern technologies and engineering.

The journal was founded in 1903. The journal is on open access on www.elibrary.ru, scholar.google.com.

The journal is registered internationally (ISSN 1684-8519) and nationally (Certificate PE № 77-166615, October 24, 2003 from the RF Ministry of Press, Broadcasting and Mass Communications).

Topical sections of the Journal focuses on the following aspect of science and engineering:

- The Planet Resources
- Mathematics, Physics and Mechanics
- Chemistry and Chemical Technology
- Power Engineering
- IT Technologies
- Humanities and Social Sciences Technologies

TPU Bulletin publishes only original research articles. It does not accept literature reviews, case reports, commentaries etc. All articles are peer reviewed by international experts. Both general and technical aspects of the submitted paper are reviewed before publication. Authors are advised to suggest 2 potential reviewers who are familiar with the research focus of the article. Final decision on any paper is made by the Editor In Chief.

The Bulletin of Tomsk Polytechnic University is published monthly.

The publication of manuscripts is free of charge.

IT Technologies. The Journal publishes only original articles containing theoretical and experimental results of solving fundamental and applied problems in the following key areas: engineering, social and information systems management; process mathematical simulation and production automation; decision making optimization; intelligent information telecommunication systems and procedures of data transmission; applied mathematics; software for computers, systems and computer networks; digital signal and image processing; cognitive hardware-software complexes with high-efficiency data processing and wireless telecommunication systems and technologies.

СОДЕРЖАНИЕ

CONTENTS

УПРАВЛЕНИЕ ТЕХНИЧЕСКИМИ СИСТЕМАМИ		CONTROL OF ENGINEERING SYSTEMS	
Определение временных интервалов в алгоритмах управления Ларкин Е.В., Ивутин А.Н.	6	Determination of time gaps in control algorithms Larkin E.V., Ivutin A.N.	
Робастная система управления с повышенным потенциалом Рустамов Г.А.	13	Robust control system with high potential Rustamov G.A.	
Автоматизированная система контроля параметров климатического и термовакuumного оборудования для изделий космической техники Андросов В.Я., Плотников А.В., Вагайцев С.А., Будников Е.А.	21	Automated control system for climatic and thermovacuum equipment parameters to test space industry products Androsov V.Ya., Plotnikov A.V., Vagaytsev S.A., Budnikov E.A.	
Оценка рисков физической безопасности линейной части магистрального нефтепровода Кукало И.А., Гривцов С.Н.	30	Risk assessment of linear part physical safety in the main oil pipeline Kukalo I.A., Gritsov S.N.	
Метод адаптивной идентификации и диагностики нефтяных пластов в процессе нормальной эксплуатации скважин и кратковременных остановок Романова Е.В., Сергеев В.Л.	43	Method of adaptive identification and diagnosis of oil reservoirs in well normal operation and short stoppages Romanova E.V., Sergeev V.L.	
ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ		INTELLIGENCE SYSTEM	
Применение цифровых водяных знаков для разработки систем графического пароля Шокарев А.В.	49	Application of digital water marks for developing graphical password systems Shokarev A.V.	
Автоматическая система мета-обучения с поддержкой выбора оптимального алгоритма решения задачи и вычисления оптимальных параметров его функционирования Орлов А.А.	57	Automatic meta-learning system supporting selection of optimal algorithm for problem solving and calculation of optimal parameters of its functioning Orlov A.A.	
Интеграция семантических моделей в исследованиях проблем энергетической безопасности Массель А.Г., Тюрюмин В.О.	70	Semantic models integration in research of energy security problems Massel A.G., Tyuryumin V.O.	
Нейросетевой параллельный алгоритм слежения за объектом в реальном времени Тарков М.С., Дубынин С.В.	78	Neural network parallel algorithm for real-time object tracking Tarkov M.S., Dubynin S.V.	
Применение сверточных нейронных сетей для выделения и распознавания автомобильных номерных знаков на изображениях со сложным фоном Друки А.А.	85	Application of convolutional neural networks for extraction and recognition of car number plates on images with complex background Druki A.A.	
Применение многошаговой сегментации для распознавания нечетких дубликатов изображений Немировский В.Б., Стоянов А.К.	92	Application of multi-step image segmentation for near-duplicate image recognition Nemirovskiy V.B., Stoyanov A.K.	
АЛГОРИТМИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ		SOFTWARE AND ALGORITHMIC SUPPORT	
К вопросу восстановления учетных данных на химических предприятиях Волошко А.В., Бедерак Я.С., Лутчин Т.Н., Кудрицкий М.Ю.	101	The problem of accounting data recovery on chemical enterprise Voloshko A.V., Bederak Ya.S., Lutchyn T.N., Kudritskiy M.Yu.	
Организация дистанционной калибровки средств измерений электрических величин Величко О.Н., Гурин Р.В., Баранов П.Ф.	108	Organization of remote calibration for measuring instruments of electrical quantities Velichko O.N., Gurin R.V., Baranov P.F.	
Применение технологии NVIDIA CUDA для обучения и декодирования скрытых марковских моделей Зацепин П.М., Гейфе Д.А.	115	NVIDIA CUDA application to train and decode the hidden Markov models Zatsepin P.M., Gefke D.A.	
Применение средств open source для визуализации стандартизированных медицинских данных Копаница Г.Д., Тараник М.А.	122	Applying open source data visualization tools to standard based medical data Kopanitsa G.D., Taranik M.A.	
Программное обеспечение для эмуляции передачи мгновенных значений измерений в соответствии со стандартом МЭК 61850 Баранов П.Ф., Муравьев С.В., Сулайманов А.О., Леванова В.С.	131	Software for emulating the sampled values transmission in accordance with IEC 61850 standard Baranov P.F., Muravyov S.V., Sulaymanov A.O., Levanova V.S.	

Управление техническими системами

УДК 004.382

ОПРЕДЕЛЕНИЕ ВРЕМЕННЫХ ИНТЕРВАЛОВ В АЛГОРИТМАХ УПРАВЛЕНИЯ

Ларкин Евгений Васильевич,

д-р техн. наук, профессор, заведующий кафедрой «Робототехника и автоматизация производства» Тульского государственного университета, Россия, 300012, Тула, пр. Ленина, 92. E-mail: elarkin@mail.ru

Ивутин Алексей Николаевич,

канд. техн. наук, доцент, доцент кафедры «Вычислительная техника» Тульского государственного университета, Россия, 300012, Тула, пр. Ленина, 92. E-mail: alexey.ivutin@gmail.com

Актуальность работы обусловлена необходимостью предоставить простые и эффективные методы оценивания характеристик управляющих алгоритмов, что позволит обеспечить достижение качественных параметров регулирования при управлении сложными многоконтурными объектами.

Цель работы: определить с точностью до плотностей распределения временные интервалы между опросами одного и того же периферийного устройства и опросами различных устройств на основе модели формирования транзакций в цифровых контроллерах как результата блужданий по состояниям полумарковского процесса, каждое из которых моделирует интерпретацию соответствующего оператора управляющего алгоритма.

Методы исследования: для достижения заявленного научного результата использованы фундаментальные научные теории: марковских процессов, теории вероятностей, теории управления, системного анализа.

Результаты: получены временные и вероятностные характеристики транзакций для управляющих алгоритмов и объектов управления самого общего вида, что позволяет производить оценивание качественных параметров регулирования при управлении сложными многоконтурными объектами. С точностью до плотностей распределения определены временные интервалы между опросами одного и того же периферийного устройства и опросами различных устройств.

Ключевые слова:

Цифровая система, алгоритм, транзакция, полумарковский процесс, эргодический процесс.

Введение

Включение ЭВМ фон-неймановского типа в контур цифрового управления объектами различной сложности порождает весьма высокие требования к временной вычислительной сложности алгоритмов управления. Под вычислительной сложностью в данном случае понимается промежуток времени между появлением на объекте ситуации, требующей адекватного реагирования со стороны управляющей ЭВМ, и выдачей соответствующего воздействия на объект. Программный цифровой контроллер, кроме того, что алгоритмически реализует заданный закон управления как элемент системы управления, вносит чистое запаздывание по времени, в связи с чем временной фактор напрямую влияет на качественные характеристики переходных процессов в системе.

В свою очередь, временная вычислительная сложность реализаций управляющих алгоритмов определяется целым рядом факторов, к которым, кроме базисных архитектуры ЭВМ, частоты тактирования и структуры команд, относятся: операционная среда, дисциплина диспетчеризации, организация транзакций, математические модели процессов, по которым строятся управляющие алгоритмы, и т. п.

Как правило, к началу разработки прикладного программного обеспечения большинство из названных факторов уже определено, поэтому реализуемый закон и последовательность программного опроса периферийных устройств являются единственными варьируемыми параметрами для оптимизации временных интервалов между транзакциями.

Алгоритмы цифрового управления объектами имеют ряд специфических особенностей, которые

были исследованы рядом авторов [1–5]. Алгоритмы являются циклическими, т. е. они имеют оператор начала, но не имеют оператора окончания, окончание вычислительного процесса производится через внешнее прерывание. Опрос периферийных устройств производится в цикле за счет включения в алгоритм специальных операторов управления транзакциями. Выбор ветви продолжения вычислительного процесса в местах ветвления алгоритма является случайным и определяется условиями, включенными в операторы принятия решения, и законами распределения обрабатываемых данных. Время выполнения операторов алгоритма является случайным, причем функции распределения времени выполнения операторов также определяются законами распределения обрабатываемых данных. В подобном аспекте управляющие алгоритмы рассматривались в [6–8].

Таким образом, процесс интерпретации детерминированного алгоритма управляющим контроллером для внешнего по отношению к системе наблюдателя рядом авторов [9–15] рассматривается как полумарковский процесс с непрерывным временем. Операторы алгоритма рассматриваются как состояния полумарковского процесса, выполнение оператора – как пребывание полумарковского процесса в соответствующем состоянии, а процесс интерпретации алгоритма в целом рассматривается как блуждание алгоритма по состояниям полумарковского процесса. При смене состояний полумарковского процесса он через случайные моменты времени попадает с определенной вероятностью в операторы транзакций, что и формирует поток опросов периферийных устройств. В указанном потоке существенными являются два типа временных интервалов:

- время между запросами к одному и тому же периферийному устройству (оно должно удовлетворять условиям Найквиста теоремы об отсчетах [16]);
- время между запросами к разным периферийным устройствам (оно должно обеспечивать требуемое качество управления, если это интервал между вводом и выводом данных в/из

ЭВМ, или требуемое рассогласование между временными решетчатыми функциями, если это интервал между вводом (выводом) данных по разным каналам).

Задача заключается в том, чтобы определять указанные интервалы для произвольного количества произвольно расположенных в алгоритме операторов управления транзакциями.

Общая полумарковская модель алгоритма управления

Типовая функциональная схема включения цифрового управления объектами приведена на рис. 1, где показаны $U=\{u_1(t), \dots, u_k(t), \dots, u_K(t)\}$ – множество опорных сигналов; $V=\{v_1(t), \dots, v_l(t), \dots, v_L(t)\}$ – множество сигналов управления объектом; $W=\{w_1(t), \dots, w_m(t), \dots, w_M(t)\}$ – множество сигналов с датчиков обратной связи. Управляющий алгоритм в контроллере организует опрос контроллеров названных сигналов, при этом при проектировании вычислительного процесса требуется обеспечить интервал времени между опросами сигналов множеств U, V, W , а также сигналами разных множеств и/или разными сигналами одного и того же множества (показано пунктирными стрелками на рис. 1).

Управление производится по обобщенному алгоритму

$$M=\{A, R, h(t)\}, \quad (1)$$

где $A=\{a_1, \dots, a_j, \dots, a_J\}$ – множество операторов алгоритма; $R=[R_{jn}]$ – матрица смежности размером $J \times J$, отражающая структуру алгоритма; $h(t)$ – полумарковская матрица размером $J \times J$, задающая стохастические и временные параметры операторов алгоритма;

$$R_{jn} = \begin{cases} 1, & \text{если из оператора } a_j \text{ алгоритма} \\ & \text{можно попасть в оператор } a_n; \\ 0, & \text{если из оператора } a_j \text{ алгоритма} \\ & \text{нельзя попасть в оператор } a_n; \end{cases}$$

$$h(t) = p \otimes f(t) = [p_{jn} \cdot f_{jn}(t)];$$

$p(t)=[p_{jn}]$ – матрица вероятностей перехода к n -му оператору алгоритма после завершения выполнения j -го оператора; $f(t)=[f_{jn}(t)]$ – матрица плотно-

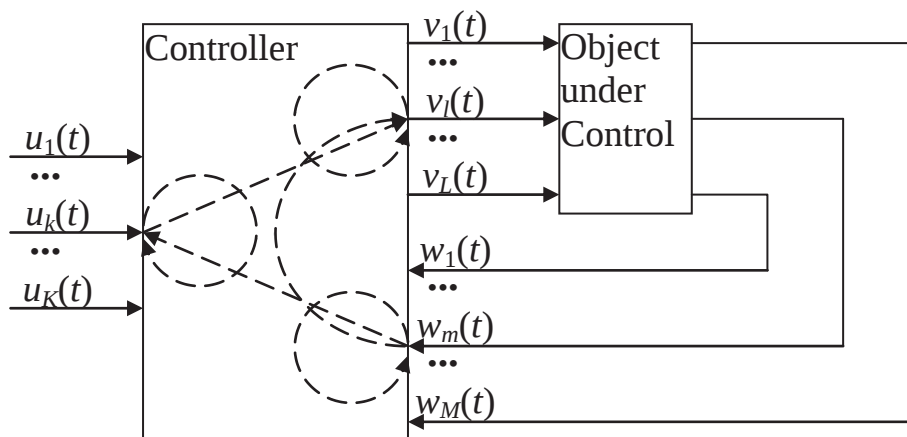


Рис. 1. Структурная схема следящей системы

стей распределения времени выполнения j -го оператора алгоритма при условии последующего перехода к выполнению n -го оператора.

Обобщенный управляющий алгоритм является циклическим, но в нем отсутствует эффект закливания, поэтому на структуру алгоритма, заданную множеством A и матрицей R , накладываются ограничения в виде сильной связности соответствующего графа. Кроме того, обобщенный управляющий алгоритм является циклическим, следовательно, на структуру R накладываются условия сильной связности. Вследствие цикличности управляющего алгоритма на вероятности и плотности распределения полумарковской матрицы $h(t)$ накладываются следующие ограничения:

$$p_{jn} = 0, \text{ если } r_{jn} = 0;$$

$$\sum_{n=1}^J p_{jn} = 1 \text{ для всех } 1 \leq j \leq J;$$

$$f_{jn}(t) = 0 \text{ при } t \leq 0;$$

$$f_{jn}(t) = \lim_{T_{jn} \rightarrow \infty} \delta(t - T_{jn}), \text{ если } r_{jn} = 0, \int_0^{\infty} f_{jn}(t) dt = 1. (2)$$

Таким образом, полумарковский процесс, моделирующий работу обобщенного алгоритма (1), является существенным и возвратным, т. к. все переключения описываются сильносвязным графом. В силу первого условия (2) полумарковский процесс является однородным. Если полумарковский процесс является однородным, все его состояния являются существенными, то он является эргодическим [17, 18].

Период опроса одного и того же устройства

Выделим в моделируемом циклическом алгоритме операторы опроса одного из задатчиков $u_k(t)$ ($1 \leq k \leq K$) или одного из управляющих воздействий $v_l(t)$ ($1 \leq l \leq L$), или одного из датчиков $w_m(t)$ ($1 \leq m \leq M$). Без нарушения общности рассуждений можно считать, что выделенные операторы алгоритма, а следовательно и состояния полумарковского процесса имеют индексы с наименьшими значениями, т. е.

$$A \supset A_s = \{a_1, \dots, a_s, \dots, a_S\}, S < J. (3)$$

Такая индексация всегда может быть обеспечена, поскольку A представляет собой неупорядоченное множество, т. е. список. Матрицы R и $h(t)$ при изменении индексов состояний могут быть получены из исходных путем соответствующей перестановки строк и столбцов.

Очевидно, что каждое переключение из состояния $a_s \in A_s$ в состояние $a_r \in A_s$ формирует поток опросов соответствующего периферийного оборудования.

Для определения времени переключения полумарковского процесса из $a_s \in A_s$ в $a_r \in A_s$ расценим каждое состояния подмножества A_s , например $a_s \in A_s$, в виде двух состояний: начального $a_{s(b)}$, отмеченного индексом $s(b)=s$, и конечного $a_{s(e)}$, отмеченного индексом $s(e)=J+s$. Таким образом, из исходного эргодического полумарковского процесса (1) формируется полумарковский процесс

$$M' = \{A', R', h'(t)\}, (4)$$

в котором

$$A' = \{a_{1(b)}, \dots, a_{s(b)}, \dots, a_{S(b)}, a_{S(b)+1}, \dots, a_j, \dots, a_J, a_{J+1(e)}, \dots, a_{J+s(e)}, \dots, a_{J+S(e)}\}; (5)$$

$$R' = \begin{pmatrix} 0 & \dots & 0 & R_{1(b), S(b)+1} & \dots & R_{1(b), J} & R_{1(b), J+1(e)} & \dots & R_{1(b), J+S(e)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & R_{S(b), S(b)+1} & \dots & R_{S(b), J} & R_{S(b), J+1(e)} & \dots & R_{S(b), J+S(e)} \\ 0 & \dots & 0 & R_{S(b)+1, S(b)+1} & \dots & R_{S(b)+1, J} & R_{S(b)+1, J+1(e)} & \dots & R_{S(b)+1, J+S(e)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & R_{J, S(b)+1} & \dots & R_{J, J} & R_{J, J+1(e)} & \dots & R_{J, J+S(e)} \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}; (6)$$

$$h'(t) = \begin{pmatrix} 0 & \dots & 0 & h_{1(b), S(b)+1} & \dots & h_{1(b), J} & h_{1(b), J+1(e)} & \dots & h_{1(b), J+S(e)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & h_{S(b), S(b)+1} & \dots & h_{S(b), J} & h_{S(b), J+1(e)} & \dots & h_{S(b), J+S(e)} \\ 0 & \dots & 0 & h_{S(b)+1, S(b)+1} & \dots & h_{S(b)+1, J} & h_{S(b)+1, J+1(e)} & \dots & h_{S(b)+1, J+S(e)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & h_{J, S(b)+1} & \dots & h_{J, J} & h_{J, J+1(e)} & \dots & h_{J, J+S(e)} \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}; (7)$$

состояния $A_b = \{a_{1(b)}, \dots, a_{s(b)}, \dots, a_{S(b)}\}$ являются начальными, а состояния $A_e = \{a_{J+1(e)}, \dots, a_{J+s(e)}, \dots, a_{J+S(e)}\}$ являются поглощающими.

Развитие полумарковского процесса (4) представляет собой блуждания по графу состояний (5) с переключениями (6), вероятностями и плотностями распределений (7). Каждая случайная последовательность переключений начинается в одном из начальных состояний подмножества A_b и оканчивается в одном из поглощающих состояний подмножества A_e (рис. 2).

Для матрицы $h'(t)$ может быть найдена характеристическая матрица по зависимости

$$h'(i\omega) = F[h'(t)],$$

где F – преобразование Фурье; ω – аргумент характеристической функции соответствующей плотности распределения (круговая частота); $i = \sqrt{-1}$.

По характеристической матрице могут быть найдены:

- взвешенная плотность распределения времени переключения, включая блуждания по промежуточным состояниям, полумарковского процесса из $a_{s(b)}$ в $a_{J+r(e)}$

$$h''_{s(b), r(e)}(t) = \sum_{n=1}^{\infty} F^{-1} [I_{s(b)} [h'(i\omega)]^n I_{J+r(e)}^{\theta}], (8)$$

где F^{-1} – обратное преобразование Фурье; θ – знак транспонирования;

$$I_{s(b)} = \begin{bmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 1(b) & s(b)-1 & s(b) & s(b)+1 & J & J+1(e) & J+S(e) & \end{bmatrix}$$

$$I_{s(e)} = \begin{bmatrix} 0 & \dots & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 1(b) & J & J+1(e) & J+s(e)-1 & J+s(e) & J+S(e)+1 & J+S(e) & \end{bmatrix}$$

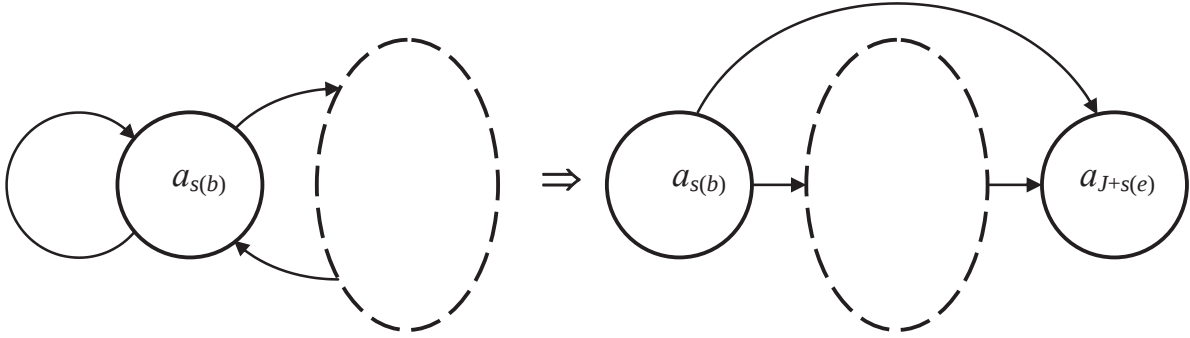


Рис. 2. Формирование начального и конечного состояний в эргодическом полумарковском процессе

- вероятность переключения из состояния $a_{s(b)}$ в $a_{r(e)}$ при блуждании по полумарковскому процессу

$$P''_{s(b),r(e)} = \int_0^{\infty} h''_{s(b),r(e)}(t) dt;$$

- плотность распределения времени переключения из состояния $a_{j(h\beta)}$ в состояние $a_{l(h\beta)}$

$$f''_{s(b),r(e)}(t) = \frac{h''_{s(b),r(e)}(t)}{P''_{s(b),r(e)}}.$$

В результате проведенных выкладок определены взвешенные плотности, вероятности и плотности распределения полумарковской матрицы процесса

$$M'' = \{A'', R'', h''(t)\}, \quad (9)$$

где $A'' = \{\alpha_1, \dots, \alpha_s, \dots, \alpha_S\}$ – множество состояний, соответствующих операторам алгоритма, генерирующим запрос на выдачу/получение данных от периферийного устройства; $R'' = [R''_{sr}]$, $R''_{sr} = 1$ – матрица смежности размером $S \times S$, отражающая структуру полного графа с петлями; $h''(t) = [h''_{sr}(t)]$ – полумарковская матрица размером $S \times S$; $h''_{sr}(t) = h''_{s(b),r(e)}(t)$; $1 \leq s(b), r(e)$, $s, r \leq S$.

Структура состояний полумарковского процесса (9) моделируется полным графом с петлями (рис. 3, а). Полумарковский процесс (9), так же как и полумарковский процесс (1), является эргодическим, причем каждое его переключение из состояния α_s в сопряженное состояние α_r , или по петле в то же самое состояние α_s , моделирует одно обращение алгоритма к периферийному устройству.

Не нарушая общности рассуждения, выберем из множества A'' состояния α_s и представим полумарковский процесс (9) в виде

$$\left\{ \alpha_s, M_s, \begin{bmatrix} R_{ss} & R_{ss'} \\ R_{ss} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} h_{ss}(t) & h_{ss'}(t) \\ h_{ss}(t) & 0 \end{bmatrix} \right\}, \quad (10)$$

где M_s – полумарковский подпроцесс, в который включены все состояния с их взаимными сопряжениями, кроме α_s ; R_{ss} моделирует все подмножество дуг, ведущих из состояния α_s в состояния полумарковского подпроцесса M_s ; M_{ss} моделирует все подмножество дуг, ведущих из состояний полумар-

ковского подпроцесса M_s в α_s ; $h_{ss}(t)$ – взвешенная плотность распределения прямого переключения из α_s в α_s по петле (9); $h_{ss'}(t)$ – взвешенная плотность распределения переключения из α_s в состояние M_s ; $h_{ss'}(t)$ – взвешенная плотность распределения переключения из состояний M_s в α_s .

Уничтожим петлю (α_s, α_s) . При ее уничтожении взвешенные плотности распределения $h''_{s(b),r(e)}(t)$ и вероятности пересчитываются по следующим зависимостям:

$$h'''_{s(b),r(e)}(t) = h_{sr}(t) = F^{-1} \left[F[h''_{s(b),r(e)}(t)] \cdot \prod_{n=0}^{\infty} F[h''_{s(b),s(b)}(t)] \right];$$

$$P'''_{s(b),r(e)}(t) = P_{sr} = \frac{P_{s(b),r(e)}}{1 - P_{s(b),s(b)}}.$$

Полумарковский процесс (10) приобретает вид (рис. 3, б)

$$\left\{ \alpha_s, M_s, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & h'_{ss}(t) \\ h'_{ss}(t) & 0 \end{bmatrix} \right\}. \quad (11)$$

Значения $h'_{ss}(t)$ и $h'_{ss'}(t)$ соответственно равны

$$h'_{ss}(t) = \sum_{\substack{r=1 \\ r \neq s}}^S h_{sr}(t);$$

$$h'_{ss'}(t) = \sum_{n=1}^{\infty} \mathfrak{T}^{-1} [\mathbf{I}_p [F[h_s(t)]]^n \mathbf{I}_s^{\theta}],$$

где

$$\mathbf{I}_p = [p_{s1}, \dots, p_{s,s-1}, 0, p_{s,s+1}, \dots, p_{sS}];$$

$$h_s(t) = \begin{bmatrix} h_{11} & \dots & h_{1,s-1} & h_{1s} & h_{1,s+1} & \dots & h_{1S} \\ & & & \dots & & & \\ h_{s-1,1} & \dots & h_{s-1,s-1} & h_{s-1,s} & h_{s-1,s+1} & \dots & h_{s-1,S} \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ h_{s+1,1} & \dots & h_{s+1,s-1} & h_{s+1,s} & h_{s+1,s+1} & \dots & h_{s+1,S} \\ & & & \dots & & & \\ h_{S,1} & \dots & h_{S,s-1} & h_{S,s} & h_{S,s+1} & \dots & h_{S,S} \end{bmatrix}$$

$$\mathbf{I}_p = [0, \dots, 0, 1, 0, \dots, 0].$$

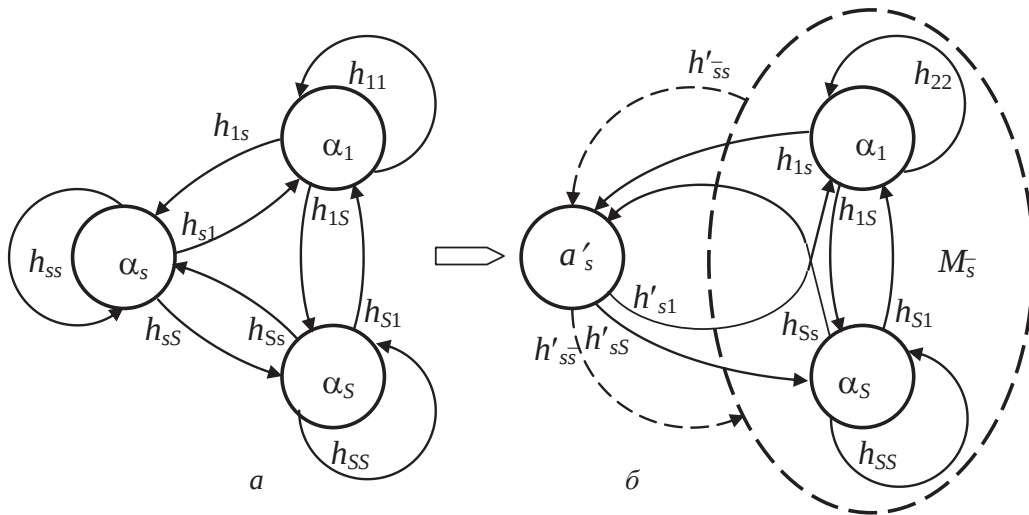


Рис. 3. Полумарковские процессы: а) из выделенных состояний; б) с разделенным состоянием α_s

Очевидно, что в эргодическом процессе (11) события переключения из состояния α'_s в состояния подпроцесса M_s образуют полную группу несовместных событий [19], поэтому $\int_0^\infty h'_{ss}(t)dt = 1$; и взвешенная плотность распределения равна плотности распределения. Точно так же события переключения из состояний подпроцесса M_s в состояние α'_s образуют полную группу несовместных событий, поэтому $\int_0^\infty h'_{ss}(t)dt = 1$ и взвешенная плотность распределения равна плотности распределения.

Применим к эргодическому процессу (11) закон больших чисел [20]. В соответствии с указанным законом, для внешнего по отношению к процессу (11) наблюдателя вероятность пребывания его в состоянии α'_s в произвольный момент времени определяется зависимостью

$$\pi_{s(b)} = \pi_s = \frac{T_s}{T_s + T_{\bar{s}}}, \quad (12)$$

где T_s – математическое ожидание времени пребывания процесса в состоянии s ; $T_{\bar{s}}$ – математическое ожидание времени пребывания процесса (11) в состояниях подпроцесса M_s ;

$$T_s = \int_0^\infty t h'_{ss}(t) dt; \quad T_{\bar{s}} = \int_0^\infty t h'_{\bar{s}\bar{s}}(t) dt.$$

Приведенные рассуждения позволяют сформулировать следующее *утверждение*: Для внешнего по отношению к эргодическому полумарковскому процессу наблюдателя вероятность пребывания процесса в произвольный момент наблюдения в наблюдаемой позиции равна отношению математических ожиданий времени пребывания к времени возврата в наблюдаемую позицию.

С учетом сформулированного и доказанного утверждения, плотность распределения времени

между двумя опросами периферийного устройства определяется по зависимости:

$$g_k(t) = \sum_{s(b)}^{S(b)} \sum_{r(e)=1}^{S(e)} \pi_{s(b)} h''_{s(b)r(e)}(t),$$

где $\pi_{s(b)}$ определяется выражением (12); $h''_{s(b)r(e)}(t)$ определяется выражением (8).

Интервал времени между опросами разных устройств

Пусть требуется определить интервал времени между опросом контроллеров разных периферийных устройств из числа изображенных на рис. 1, например, между опросом контроллера сигналов $u_k(t)$ и $v_l(t)$. Как и в предыдущем пункте, будем считать, что контроллер сигнала $u_k(t)$ опрашивается операторами, моделируемыми состояниями подмножества $A_s(3)$, а контроллер сигнала $v_l(t)$ опрашивается операторами, моделируемыми подмножеством $A \setminus A_q$, $A_q \cap A_s = \emptyset$. Без нарушения общности рассуждений будем считать, что состояния подмножества A_q стоят последними в списке состояний A , т. е. $A_q = \{a_{J-Q+1}, \dots, a_{J-q+1}, \dots, a_J\}$.

По методике, изложенной в предыдущем пункте, вероятности пребывания полумарковского процесса в состояниях подмножества A_s могут быть определены. Сформируем из матрицы $h(t)$ полумарковскую матрицу следующего вида:

$$\tilde{h}(t) = \begin{bmatrix} 0 & \dots & 0 & h_{1s+1} & \dots & h_{1J-Q} & h_{1J-Q+1} & \dots & h_{1J} \\ & & & & & & & & \\ 0 & \dots & 0 & h_{s+1} & \dots & h_{sJ-Q} & h_{sJ-Q+1} & \dots & h_{sJ} \\ 0 & \dots & 0 & h_{s+1J+1} & \dots & h_{s+1J-Q} & h_{s+1J-Q+1} & \dots & h_{s+1J} \\ & & & & & & & & \\ 0 & \dots & 0 & h_{J-Qs+1} & \dots & h_{J-QJ-Q} & h_{J-QJ-Q+1} & \dots & h_{J-QJ} \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ & & & & & & & & \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix}. \quad (13)$$

Тогда взвешенная плотность распределения времени блуждания полумарковского процесса из состояний подмножества A_s в состояния подмножества A_q определяется зависимостью

$$h_{kl}(t) = \sum_{n=1}^{\infty} F^{-1}[\mathbf{I}_{\pi}[F[\tilde{\mathbf{h}}(t)]]^n \mathbf{I}_Q^{\theta}],$$

где

$$\mathbf{I}_{\pi} = [\pi_1, \dots, \pi_s, \dots, \pi_S, 0, \dots, 0];$$

$$\mathbf{I}_Q = [0, \dots, 0, \underset{j-Q+1}{1}, \dots, 1].$$

Очевидно, что траектории блуждания, определенные матрицей (13), не составляют полную группу несовместных событий, поскольку из них исключаются траектории, проходящие более одного раза через состояния A_s , и траектории, проходящие через состояния подмножества A_q . Вероятность и плотность распределения времени достижения состояний подмножества A_q из A_s определяются по следующим зависимостям:

$$p_{kl} = \int_0^{\infty} h_{kl}(t) dt; \quad g_{kl}(t) = \frac{h_{kl}(t)}{p_{kl}}.$$

Для плотностей распределения $g_k(t)$ при решении практических задач могут быть определены начальный и центральные моменты различных порядков, в частности математическое ожидание и дисперсия:

$$\tau_{k,kl} = \int_0^{\infty} t g_{k,kl}(t) dt; \quad \sigma_{k,kl}^2 = \int_0^{\infty} t^2 g_{k,kl}(t) dt - \tau_{k,kl}^2.$$

СПИСОК ЛИТЕРАТУРЫ

1. Изерман Р. Цифровые системы управления / пер. с англ. – М.: Мир, 1984. – 541 с.
2. Дроздов Н.В., Мирошник И.В., Скорубский И.В. Системы автоматического управления с микроЭВМ. – Л.: Машиностроение, 1989. – 284 с.
3. Бесекерский В.А., Ефимов Н.Б., Зиядинов С.И. и др. Микропроцессорные системы автоматического управления / под общ. ред. В.А. Бесекерского. – Л.: Машиностроение, 1988. – 365 с.
4. Остром К., Виттенмарк Б. Системы управления с ЭВМ / пер. с англ. А.Н. Николаева, Т.С. Чеботаревой. – М.: Мир, 1987. – 480 с.
5. Олссон Г., Пиани Дж. Цифровые системы автоматизации и управления. – СПб.: Невский Диалект, 2001. – 558 с.
6. Байцер Б. Микроанализ производительности вычислительных систем. – М.: Радио и связь, 1983. – 360 с.
7. Барский А.Б. Планирование параллельных вычислительных процессов. – М.: Машиностроение, 1980. – 191 с.
8. Феррари Д. Оценка производительности вычислительных систем. – М.: Мир, 1981. – 576 с.
9. Корольюк В.С., Турбин А.Ф. Полумарковские процессы и их применения. – Киев: Наукова думка, 1976. – 138 с.
10. Привалов А.Н., Ларкин Е.В. Моделирование информационных процессов тренажерных систем: Концепция, методология, модели. – Saarbrücken, Deutschland: LAP LAMBERT Academic Publishing GmbH & Co., 2012. – 230 с.
11. Ярлыков М.С., Миронов М.А. Марковская теория оценивания случайных процессов. – М.: Радио и связь, 1993. – 460 с.

Заключение

В статье авторы представили формирование транзакций в цифровых контроллерах как результат блужданий по состояниям полумарковского процесса, каждое из которых моделирует интерпретацию соответствующего оператора управляющего алгоритма. Это позволило определить с точностью до плотностей распределения временные интервалы между опросами одного и того же периферийного устройства и опросами различных устройств. Для найденных плотностей распределения получены выражения для математических ожиданий и дисперсий.

Временные и вероятностные характеристики транзакций получены для управляющих алгоритмов и объектов управления самого общего вида. Они являются существенно важными с точки зрения достижения качественных параметров регулирования при управлении сложными многоконтурными объектами. Из общих выражений для плотностей распределения могут быть получены более простые выражения для временных интервалов при управлении одноконтурными объектами.

Дальнейшие исследования в этой области могут быть направлены на разработку простых инженерных методик оценки числовых характеристик временных интервалов, например методами имитационного моделирования, или разработку time-compiler – временного компилятора, который параллельно с трансляцией управляющей программы с языка высокого уровня будет давать оценку числовых характеристик плотностей распределения между выбранными операторами алгоритма.

12. Gotz N. et al. Multiprocessor and Distributed System Design: The Integration of Functional Specification and Performance Analysis Using Stochastic Process Algebras // Lecture Notes in Computer Science. – 1993. – V. 729. – P. 121–146.
13. Ajmone Marsan M., Conte G., Balbo G. A class of generalized stochastic Petri nets for the performance evaluation of multiprocessor systems // ACM Transactions on Computer Systems. – 1984. – V. 2. – № 2. – P. 93–122.
14. Wittenmark B., Nilsson J., Torngren M. Timing problems in real-time control systems // Proc. of 1995 American Control Conference – ACC'95. – Seattle, WA, American Automatic Control Council. – 1995. – P. 2000–2004.
15. Stability of linear feedback systems with random communication delays / R. Krtočila, U. Ozguner, H. Chan, H. Goktas, J. Winkelman, M. Liubakka // International Journal of Control. – 1994. – V. 59. – № 4. – P. 925–953.
16. Nyquist H. Certain topics in telegraph transmission theory // Trans. AIEE. – 1928. – V. 47. – P. 617–644.
17. Papoulis A. Probability, random variables, and stochastic processes. – New York: McGraw-Hill, 1991. – P. 427–442.
18. Limnios N., Oprisan G. Semi-Markov Processes and Reliability. – Boston, MA: Boston Imprint Birkhauser, 2001. – 222 p.
19. Loyer M. Equally Likely, Exhaustive and Independent Events // Mathematics Magazine. – 1980. – V. 53. – № 3. – P. 167–171.
20. Grimmett G.R., Stirzaker D.R. Probability and Random Processes. 3rd edition. – Oxford: Clarendon Press, 2001. – 608 p.

Поступила 05.03.2014 г.

DETERMINATION OF TIME GAPS IN CONTROL ALGORITHMS

Evgeny V. Larkin,

Dr. Sc., Tula State University, 92, Lenin avenue, Tula, 300012, Russia.

E-mail: elarkin@mail.ru

Aleksey N. Ivutin,

Cand. Sc., Tula State University, 92, Lenin avenue, Tula, 300012, Russia.

E-mail: alexey.ivutin@gmail.com

The urgency of the discussed issue is caused by the need to provide simple and effective methods for evaluating the characteristics of control algorithms that would achieve the quality control parameters in the management of complex multi-loop objects.

The main aim of the study: to determine up to density distribution the timing between polls of the same peripheral and surveys of various devices based on the model of formation of transactions as a result of digital controllers walks over states semi-Markov processes, each of which simulates the interpretation of the corresponding operator control algorithm.

The methods used in the study: to achieve the claimed scientific result the authors have used the fundamental scientific theory: Markov processes, probability theory, control theory, system analysis.

The results: The authors obtained the temporal and probabilistic characteristics of the transaction for control algorithms and facilities management of a general form, which allows evaluating quality control parameters in management of complex multi-loop objects. Time intervals between surveys of the same peripheral and surveys of different devices were defined up to a density distribution.

Key words:

Digital system, algorithm, transaction, semi-Markov process, ergodic process.

REFERENCES

1. Izerman R. *Tsifrovye sistemy upravleniya* [Digital Control Systems]. Moscow, Mir, 1984. 541 p.
2. Drozdov N.V., Miroshnik I.V., Skorubsky I.V. *Sistemy avtomaticheskogo upravleniya s mikroEVM* [Automatic control system with microcomputer]. Leningrad, Mashinostroyeniye, 1989. 284 p.
3. Besekersky V.A., Efimov N.B., Ziatdinov S.I. *Mikroprotsessornye sistemy avtomaticheskogo upravleniya* [Microprocessor based automatic control system]. Leningrad, Mashinostroyeniye, 1988. 365 p.
4. Ostrem K., Vittenmark B. *Sistemy upravleniya s EVM* [Computer Controlled Systems]. Moscow, Mir, 1987. 480 p.
5. Olsson G., Piani Dzh. *Tsifrovye sistemy avtomatizatsii i upravleniya* [Computer Systems for Automation and Control]. Saint Petersburg, Nevsky Dialekt, 2001. 558 p.
6. Baytser B. *Mikroanaliz proizvoditelnosti vychislitelnykh sistem* [Micro-Analysis of Computer System Performance]. Moscow, Radio i svyaz, 1983. 360 p.
7. Barsky A.B. *Planirovaniye parallelnykh vychislitelnykh protsessov* [Planning of parallel computational processes]. Moscow, Mashinostroyeniye, 1980. 191 p.
8. Ferrari D. *Otsenka proizvoditelnosti vychislitelnykh sistem* [Computer Systems Performance Evaluation]. Moscow, Mir, 1981. 576 p.
9. Korolyuk V.S., Turbin A.F. *Polumarkovskie protsessy i ikh primeneniya* [Semi-Markov processes and their application]. Kiev, Naukova dumka, 1976. 138 p.
10. Privalov A.N., Larkin E.V. *Modelirovaniye informatsionnykh protsessov trenazhernykh sistem: Kontseptsiya, metodologiya, modeli* [Modeling of information processes in training systems: Strategy, Methodology, Models]. Saarbrücken, Deutschland, LAP LAMBERT Academic Publishing GmbH & Co., 2012. 230 p.
11. Yarlykov M.S., Mironov M.A. *Markovskaya teoriya otsenivaniya sluchaynykh protsessov* [Markov's theory of stochastic processes estimation]. Moscow, Radio i svyaz, 1993. 460 p.
12. Gotz N. Multiprocessor and Distributed System Design: The Integration of Functional Specification and Performance Analysis Using Stochastic Process Algebras. *Lecture Notes in Computer Science*, 1993, vol. 729, pp. 121–146.
13. Ajmone Marsan M., Conte G., Balbo G. A class of generalized stochastic Petri nets for the performance evaluation of multiprocessor systems. *ACM Transactions on Computer Systems*, 1984, vol. 2, no. 2, pp. 93–122.
14. Wittenmark B., Nilsson J., Torngren M. Timing problems in real-time control systems. *Proc. of 1995 American Control Conference – ACC'95*. Seattle, WA, American Automatic Control Council, 1995, pp. 2000–2004.
15. Krtolica R., Ozguner U., Chan H., Goktas H., Winckelman J., Liubakka M. Stability of linear feedback systems with random communication delays. *International Journal of Control*, 1994, vol. 59, no. 4, pp. 925–953.
16. Nyquist H. Certain topics in telegraph transmission theory. *Trans. AIEE*, 1928, vol. 47, pp. 617–644.
17. Papoulis A. *Probability, random variables, and stochastic processes*. New York, McGraw-Hill, 1991. pp. 427–442.
18. Limnios N., Oprisan G. *Semi-Markov Processes and Reliability*. Boston, MA: Boston Imprint Birkhauser, 2001. 222 p.
19. Loyer M. Equally Likely, Exhaustive and Independent Events. *Mathematics Magazine*, 1980, vol. 53, no. 3, pp. 167–171.
20. Grimmett G.R., Stirzaker D.R. *Probability and Random Processes*. 3rd edition. Oxford: Clarendon Press, 2001. 608 p.

РОБАСТНАЯ СИСТЕМА УПРАВЛЕНИЯ С ПОВЫШЕННЫМ ПОТЕНЦИАЛОМ

Рустамов Газанфар Арастун оглы,

д-р техн. наук, проф. кафедры «Автоматика и управление»

Азербайджанского Технического Университета, Азербайджан,

Az1073, Баку, пр. Г. Джавида, 25. E-mail: gazanfar.rustamov@gmail.com

Предлагается подход к управлению априорно неопределенными динамическими объектами на основе метода функции Ляпунова, в частности, «метода квадратичных форм». Возможность беспределного увеличения коэффициента усиления PD-регулятора $n-1$ порядка позволяет подавлять генеральные составляющие неопределенной модели до сколь угодно малой величины. Этим обеспечивается высокая точность слежения эталонной траектории для широкого класса неопределенностей. В пределе система описывается линейным однородным уравнением гиперплоскости. Эта особенность позволяет определить параметры настройки регулятора на основе задаваемых показателей качества с использованием метода «модального управления». Предложенная методика позволяет обойти сложные математические трудности, возникающие при синтезе робастных систем, и управлять нелинейными нестационарными объектами в условиях существенной неопределенности с помощью простого физически понятного регулятора. К недостаткам предложенной методики следует отнести усиление высокочастотных помех, имеющих неопосредственный доступ в регулятор, а также использование производных выхода для формирования PD-регулятора. Решение модельных задач на Matlab/Simulink позволило сделать ряд положительных выводов, имеющих важное прикладное значение.

Ключевые слова:

Неопределенность, следящая робастная система, функция Ляпунова, партитурная модель, большой коэффициент усиления, параметрический маятник.

Введение

Проблема неопределенности. В отличие от понятия «грубость» по А.А. Андронову [1] термин «робастность» используется в случае, когда реквизиты объекта изменяются в широком диапазоне. Поэтому «робастность» можно характеризовать как «глобальную грубость», или «грубость в большом».

Все системы управления в той или иной степени функционируют в условиях неопределенности. Неопределенность не относится к реальному объекту, а является особенностью математической модели, построенной в условиях неполноты наших знаний о реальном объекте. Стремление улучшить модель путем добавления все новых факторов является нерациональным подходом, ибо каждый добавленный фактор вносит свою неопределенность. Кроме того, попытка формализовать непонятные закономерности может привести к появлению источников паразитной динамики.

Для практики автоматического управления важное значение имеет разработка методов синтеза на основе простых, но более содержательных моделей. Для таких моделей достаточно добиться «структурного соответствия», что нетрудно сделать, если знать основные закономерности поведения реального процесса.

Модели, предназначенные для целей управления, достаточно представить на уровне «партитур» (здесь имеется в виду «партитура, состоящая из небольшой модели для импровизации...» [2]). Генеральные составляющие таких моделей кроме известных компонент могут содержать и неопределенности по структуре, параметрам, внешним возмущениям и др.

Проблема управления. Обычно регулятор настраивается на номинальную модель объекта. Од-

нако при реализации регулятор подключается не к модели, а, естественно, к реальному объекту. При функционировании расхождение объекта от своей первоначальной модели (номинальная модель) может привести к непредсказуемым последствиям. Стратегия управления в условиях неопределенности заключается в компенсации этой разницы, не зная её природу.

При построении робастных систем полный отказ от использования средств адаптации, в частности алгоритмов идентификации и наблюдения неопределенностей, не позволяет получить ожидаемого качества переходных процессов и запасов устойчивости. Тем не менее, «очистка» номинальной динамики от всевозможных неопределенностей с использованием адаптивных методов является обременительной и часто невыполнимой задачей из-за невозможности разделения движения, особенно в нелинейном случае. Поэтому в основном рассматриваются аддитивные и медленно изменяющиеся неопределенности.

Особый интерес представляют методы управления в условиях неопределенности, непосредственно не использующие алгоритмы адаптации, однако обладающие высокими робастными показателями.

Неадаптивные методы управления в условиях неопределенности

В качестве претендентов на неадаптивные выделим следующие методы:

1. *Методы, основанные на H_∞ -теории* (Doyle, 1989) [3–5]. Основная задача этой теории основывается на минимизации H_∞ -нормы, характеризующей энергию выхода. Метод несколько эвристичен. Предполагается, что подобная минимизация кроме робастности должна обеспе-

- чивать и получение высоких динамических и статических показателей. Существуют упрощенные задачи до точности заданного уровня качества $\gamma > 0$.
2. *Метод, основанный на формировании фазовых ограничений* [6]. Идея этого метода заключается в «стеснении» системы в некоторую «трубку», где имеется свобода выбора решения в виде неравенств.
 3. *Методы, основанные на преднамеренной организации скользящих режимов* (С.В. Емельянов, В.И. Уткин (1967–1981), А.В. Левантовский, С.К. Коровин (1986–1997)) [7–9]. Соответствующая система обладает робастным свойством только при движении по линии скольжения и не является робастной до попадания на эту линию. Скользящий режим является уязвимым по отношению к структурным неопределенностям. Изменение топологии корней отдельных структур может привести к нарушению условия существования и устойчивости скользящего режима.
 4. *Методы на основе большого коэффициента усиления* (М.В. Мееров (1967–1986)) [10]. Эти методы обладают повышенной точностью и робастной устойчивостью по отношению к неопределенным параметрам и внешним низкочастотным возмущениям. Однако не обеспечивается робастность переходных процессов. Кроме того, не все объекты допускают реализацию чрезмерно большого коэффициента усиления.
 5. *Вибрационные методы управления*. Преднамеренная организация автоколебательных и вибрационных режимов позволяет поддерживать допустимую точность в установившемся режиме [11, 12].
 6. *Интервальные методы управления*. В этом направлении многие исследования делают акцент на известную работу В.Л. Харитонов [13, 14]. Указанные методы применимы только к линейным системам и гарантируют асимптотическую устойчивость, не гарантируя точность и показатели качества.

Вообще говоря, квалификация методов 3–5, как робастных, является нестрогой. Некоторые достоинства и недостатки указанных методов изложены в [15–17].

Средствами борьбы с широким классом неопределенностей, наиболее универсальными и технически просто реализуемыми, могут являться системы, допускающие беспредельное увеличение коэффициента усиления разомкнутой цепи без потери устойчивости. Тем не менее, классические методы в указанном направлении не нашли должного развития. По видимому, основными причинами этого являются:

- сложность структур, обеспечивающих беспредельное увеличение коэффициента усиления без нарушения устойчивости;
- применимость только к линейным системам, что гарантирует робастность статического ре-

жима (установившегося или статической ошибки), не гарантируя робастность динамических показателей качества;

- низкая репутация «большого коэффициента», в контексте классической элементной базы.

Во всех неадаптивных робастности системах в той или иной форме наблюдаются высокие коэффициенты усиления, обусловленные свойствами робастности. Можно показать множество примеров, где коэффициент усиления регулятора при настройке принимает неожиданно высокое значение. Так, в [18] оптимальный H_∞ -контроллер имеет коэффициент усиления $K_c = 505249$. В [19] установившееся значение коэффициента усиления периодической части регулятора $k = 380$, а в [20] оптимальный коэффициент усиления равен 100, и его дальнейшее увеличение приводит к нарушению устойчивости. Характерным для этих систем является незащищенность от беспредельного увеличения коэффициента усиления. Важно отметить, что, подобно системам с высоким коэффициентом усиления, также происходит усиление высокочастотных помех, налагаемых на координаты состояния.

Обобщая выявленную закономерность, для синтеза простых робастных регуляторов, востребованных в практических приложениях, можно предложить ряд рекомендаций:

- возможность беспредельного увеличения коэффициента усиления можно предусмотреть при постановке задачи, не дожидаясь результатов решения различных робастных задач, приводящих к появлению большого коэффициента усиления;
- для синтеза подобных систем следует использовать такие принципы, которые приводят к простым, реализуемым и работоспособным структурам. Основу таких принципов может составить метод функции Ляпунова;
- с целью обеспечения робастности переходных процессов (динамического режима) следует организовать движение системы по фиксированному многообразию, например по гиперплоскости. Системы, основанные на большом коэффициенте усиления, будем называть K_∞ -робастными системами.

В [15] рассмотрены K_∞ -системы, где в качестве регулятора использован PI-регулятор с большим коэффициентом усиления, представленный как гипернаблюдатель неопределенностей. В данной статье исследуется задача синтеза K_∞ -робастных систем на основе метода функции Ляпунова с использованием Р-регулятора.

Постановка задачи

Рассмотрим одномерный по входу и выходу управляемый объект n -го порядка, задаваемый дифференциальным уравнением в виде «партитурной» модели:

$$\dot{y}^{(n)} = f(y, t) + b(y, t)u, \quad t \in [0, \infty), \quad (1)$$

где $y = (y, \dot{y}, y^{(n-1)})^T = (x_1, x_2, \dots, x_n)^T \in R^n$ – доступный измерению или оценке вектор состояния; $y \in R$ –

управляемый выход; $u \in R$ – управляющее воздействие; $f(y, t)$, $b(y, t) > 0$ – нелинейные неизвестные ограниченные функции.

Генеральные составляющие $f(\cdot)$ и $b(\cdot)$ являются источниками неопределенностей, однако могут содержать и известные компоненты. В любом случае $f(\cdot)$ и $b(\cdot)$ принимаются как потенциальные источники паразитной динамики и подавляются (математически происходит деление на большое число). Взамен приведенная система наделяется желаемой динамикой.

Функции $f(\cdot)$ и $b(\cdot)$ могут содержать негладкие и разрывные нелинейности.

Рассматривается общая задача стабилизации, состоящая в таком выборе управления u , которое после завершения переходной составляющей обеспечивает движение выхода $y(t)$ объекта (1) по эталонной траектории $y_d(t)$. При этом состояние системы отслеживает эталонную траекторию $y_d = (y_d, \dot{y}_d, \dots, y_d^{(n-1)})^T$. Вектор ошибок слежения определяется выражением

$$e = y_d - y = (e, \dot{e}, \dots, e^{(n-1)})^T.$$

Решение задачи

В соответствии с принципом обратной связи стратегию управления будем строить на основе ошибки $e(t)$. В методе функции Ляпунова стратегия управления строится на основе фундаментального соотношения

$$dV(e)/dt < 0. \quad (2)$$

В качестве претендента на функцию Ляпунова примем распадающуюся на линейные множители квадратичную форму [21]:

$$V = (a_1 e + a_2 \dot{e} + \dots + a_n e^{(n-1)})(b_1 e + b_2 \dot{e} + \dots + b_n e^{(n-1)}).$$

Принимая $a_i = b_i = c_i$, $i = \overline{1, n}$, $c_n = 1$, после ввода нормирующего множителя, будем иметь:

$$V = \frac{1}{2} s^2, \quad (3)$$

где

$$s = c_1 e + c_2 \dot{e} + \dots + e^{(n-1)}. \quad (4)$$

Производная по времени функции (3):

$$dV/dt = \dot{V} = s\dot{s}, \quad (5)$$

где

$$\dot{s} = c_1 \dot{e} + c_2 \ddot{e} + \dots + e^{(n)}. \quad (6)$$

Из (5) следует, что для выполнения соотношения (2) достаточно поддерживать противоположность знаков s и $\dot{s}$. С этой целью запишем:

$$\dot{s} = u_s(s).$$

Функция u_s отвечает следующим требованиям:

$$\text{if } s > 0 \text{ then } u_s < 0; \text{ if } s < 0 \text{ then } u_s > 0. \quad (7)$$

В зависимости от вида u_s можно получить различные режимы движения. Например, при релейном $u_s = -k \cdot \text{sign}(s)$ и релейно-линейном $u_s = -k \cdot \text{sat}(s, \Delta)$ управлениях, соответственно, скользящее и квази-

скользящее движение, где Δ – ширина диапазона линейности при гладких Р-, РІ-управлениях асимптотическое.

Для определения управления u , обеспечивающего достаточное условие асимптотической устойчивости $\dot{V} < 0$, необходимо связать $\dot{V}$ с u . С этой целью в (6) выполним замену старшей производной $e^{(n)}$ выражением, получаемым в силу уравнения объекта (1):

$$e^{(n)} = y_d^{(n)} - y^{(n)} = y_d^{(n)} - f(y, t) - b(y, t)u.$$

Тогда можно записать

$$\dot{s} = x(t) - b(y, t)u, \quad b > 0. \quad (8)$$

Здесь

$$x(t) = \sum_{i=1}^{n-1} c_i e^{(i)} + y_d^{(n)} - f(y, t).$$

Из (9) нетрудно увидеть, что при $x(t)=0$ управление, обеспечивающее условие асимптотической устойчивости $dV(e)/dt = s\dot{s} < 0$, с учетом (7) можно записать в виде простого Р-закона управления:

$$u = u_s = k \cdot s(t). \quad (9)$$

По аналогии с [8] управление (10) будем называть «робастным эквивалентным управлением».

Влияние $x(t)$ на динамику системы можно ослабить путем увеличения коэффициента усиления регулятора k . Учитывая (9) в (8), представим последнее в виде:

$$s(t) = b^{-1}[x(t) - \dot{s}(t)]/k.$$

При устремлении k в бесконечность в пределе будем иметь:

$$\lim_{k \rightarrow \infty} s(t) = 0.$$

Отсюда следует, что система в пределе описывается уравнением гиперплоскости (4):

$$s = c_1 e + c_2 \dot{e} + \dots + e^{(n-1)} = 0, \quad s(0) = s_0.$$

При движении в малой окрестности вырожденной траектории $s=0$ порядок системы уменьшается на единицу. Поэтому для устойчивости движения угловые коэффициенты c_i гиперплоскости $s=0$ должны являться коэффициентами Гурвицевого полинома пониженного порядка:

$$H(p) = p^{n-1} + c_{n-1}p^{n-2} + \dots + c_1. \quad (10)$$

Вещественные части корней такого полинома $\text{Re}(p_i) < 0$.

Передаточные функции

Согласно (9) передаточная функция регулятора имеет вид:

$$W_c(p) = \frac{U_c(p)}{E(p)} = k(p^{n-1} + c_{n-1}p^{n-2} + \dots + c_3p^2 + c_2p + c_1). \quad (11)$$

Регулятор имеет пониженный $n-1$ порядок.

Согласно (10) характеристический полином системы:

$$D_{\text{sys}}(p) = H(p) = p^{n-1} + c_{n-1}p^{n-2} + \dots + c_3p^2 + c_2p + c_1. \quad (12)$$

Особенности реализации. Реализовать регулятор на основе выражения (10) невозможно, поскольку оно содержит операторы идеального дифференцирования. При реализации оператора дифференцирования p обычно используется реальное дифференцирующее звено с передаточной функцией:

$$W_d(p) = \frac{p}{T_d p + 1}. \quad (13)$$

Постоянная времени дифференциатора $T_d \approx 0,0001 - 0,02$ с.

В этом случае передаточную функцию регулятора (11) можно представить в виде:

$$W_C(p) = k(c_1 + c_2 W_d + c_3 W_d^2 + \dots + c_{n-1} W_d^{n-2} + c_n W_d^{n-1}). \quad (14)$$

Учитывая (13) в (14) и приравняв в числителе $T_d=0$, получим приближенную передаточную функцию, имеющую реализуемую форму в виде рациональной дроби:

$$W_C(p) \approx k \frac{p^{n-1} + c_{n-1}p^{n-2} + \dots + c_3p^2 + c_2p + c_1}{(T_d p + 1)^{n-1}}. \quad (15)$$

При $n=2$ имеем:

$$W_C(p) \approx k \frac{p + c_1}{T_d p + 1}.$$

Для $n=2$ при моделировании в качестве уравнения регулятора целесообразно использовать точное выражение

$$u = ks(t), \quad s = c_1 e + \dot{e}.$$

Определение параметров настройки

При движении из произвольной точки $s(0)=s_0$ система при достаточно большом коэффициенте усиления k почти мгновенно попадает на гиперплоскость $s(t)=0$ и в дальнейшем движется по этой плоскости в начало координат (рис. 3, б). Поэтому движение системы может быть описано уравнением гиперплоскости $s(t)=0$.

Настройка регулятора. Качественные показатели (время установления t_s , перерегулирование $\sigma\%$ и др.) переходной характеристики по ошибке $e(t)$ зависят от параметров настройки c_i , $i=1, \dots, n-1$ регулятора (11). Поскольку порядки регулятора (11) и характеристического уравнения замкнутой системы (12) одинаковы, то путем изменения параметров c_i , $i=1, \dots, n-1$, можно разместить все полюсы системы по желаемой схеме.

Монотонная переходная характеристика, $\sigma=0\%$. Потребуем, чтобы переходная характеристика по ошибке слежения $e(t)$ не имела перерегулирования, т. е. являлась монотонной. С этой целью для корней (полюсы системы) p_i характеристического полинома $D_{\text{sys}}(p)$ (12) выберем биномиальное распределение. При таком распределении $p_1=p_2=\dots=p_{n-1}$. Кратность корней позволяет опреде-

лить их на основе требуемого времени установления t_s переходной характеристики.

При кратных корнях выражение (12) можно представить в виде:

$$D_{\text{sys}}(p) = (p + p_1)^{n-1}.$$

Определим p_1 . При кратных корнях решение имеет вид:

$$e(t) = (C_1 + C_2 t + C_3 t^2 + \dots + C_n t^{n-2}) e^{-p_1 t}.$$

Представим решение в векторном виде:

$$e(t) = C^T \theta e^{-p_1 t}, \quad C = P e(0), \quad (16)$$

где P -квадратная матрица размера $(n-1) \times (n-1)$.

Для системы пятого порядка ($n=5$) включительно

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ p_1 & 1 & 0 & 0 \\ \frac{1}{2} p_1^2 & p_1 & \frac{1}{2} & 0 \\ \frac{1}{6} p_1^3 & \frac{1}{2} p_1^2 & \frac{1}{2} p_1 & \frac{1}{6} \end{bmatrix},$$

$$e(0) = \begin{bmatrix} e(0) \\ \dot{e}(0) \\ \ddot{e}(0) \\ \ddot{\ddot{e}}(0) \end{bmatrix}, \quad \theta = \begin{bmatrix} 1 \\ t \\ t^2 \\ t^3 \end{bmatrix}.$$

Пусть время установления для $\delta_s = \pm(2-5)\%$ -й допустимой ошибки задана $t=t_s$. Тогда, согласно (16), для определения p_1 получаем следующее выражение:

$$e(t_s) = \pm \delta_s / 100 = C^T \big|_{e=e(0)} \cdot \theta \big|_{t=t_s} e^{-p_1 t_s}.$$

Если $e(0)>0$, то и $\delta_s>0$ (приближение сверху), иначе $\delta_s<0$.

Соответствующее трансцендентное уравнение:

$$f(p_1) = C^T \big|_{e=e(0)} \cdot \theta \big|_{t=t_s} e^{-p_1 t_s} \mp \delta_s / 100 = 0. \quad (17)$$

Уравнение (18) можно решить с помощью функций Matlab: `solve('f(p1)=0')` или `fzero('f(p1)',[a,b],0.01)`.

После определения p_1 для описания уравнения регулятора во временной области (9) следуют определить параметры настройки c_i , $i=1, \dots, n-1$.

Эти параметры определяются как коэффициенты полинома

$$D_{\text{sys}}(p) = (p + p_1)^{n-1} = p^{n-1} + a_{n-1}p^{n-2} + \dots + a_2p + a_1. \quad (18)$$

Отсюда параметры настройки регулятора:

$$c_1 = a_1, c_2 = a_2, \dots, c_{n-1} = a_{n-1}.$$

При реализации уравнения регулятора в виде передаточной функции значения c_i , $i=1, \dots, n-1$, можно не вычислять, а передаточную функцию регулятора (15) реализовать на основе выражения:

$$W_C(p) \approx k \frac{(p + p_1)^{n-1}}{(T_d p + 1)^{n-1}}.$$

Определение коэффициента усиления k . Этот параметр связан с доминирующей неопределенностью и определяется экспериментально при компьютерном моделировании или на реальном объекте.

Если доминирующая неопределенность известна, то ее нужно реализовать, вызывая наихудший режим. Увеличивая k , следует контролировать ошибку слежения $e(t)$ в точке $t=t_s$ и добиться вхождения траектории $e(t)$ в коридор допустимой ошибки, т. е. выполнения условия $e(t) \leq |\delta_s|/100$ для всех $t \leq t_s$.

После некоторого значения k в системе устанавливаются заданные σ % и t_s . Однако не следует злоупотреблять значением k , поскольку это может привести к перерасходу энергии. Если энергия управления строго регламентирована, например, в подвижных или летательных аппаратах, то целесообразно использовать самонастройку k в зависимости от ошибки $e(t)$, поддерживая k на минимальном уровне.

Значения c_i также могут определяться с помощью *Optimization Toolbox/Signal Constraint* по девяти вводимым качественным показателям [22, 23].

Результаты моделирования

Рассматриваемая в статье робастная система относится к классу интервальных систем ввиду того, что источником неопределенности являются параметры модели, задаваемые своими предельными значениями.

Пример. Параметрический маятник. Решим задачу управления параметрическим маятником с периодически изменяющимися коэффициентами [24]:

$$\ddot{y} = -a \frac{\dot{R}}{R} \dot{y} - g \frac{1}{R} \sin(y) + \frac{1}{mR^2} u + n(t),$$

где $m=1\text{ kg}$, $g=9,81\text{ m/s}^2$, $a=2$; $0 < R_m \leq R \leq R_M$, $y(0)=(0,0)^T$.

Уравнение движения груза:

$$R = 0,8 + 0,1 \sin(8t) + 0,3 \cos(4t).$$

На объект в точке приложения управления действует Гауссовская помеха $n(t)$, представленная на рис. 1.

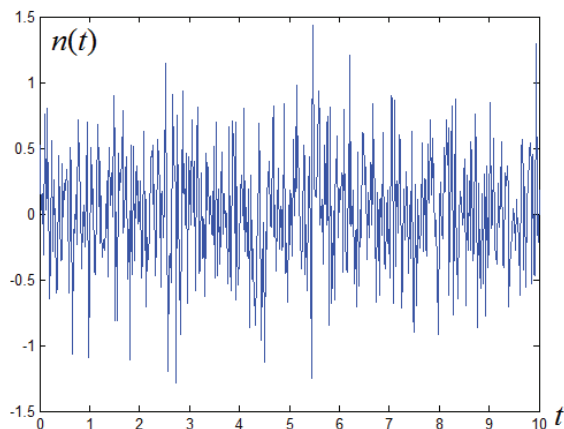


Рис. 1. Гауссовская помеха

Пусть эталонная траектория задана в виде:

$$y_d = 0,5 \sin(0,5t) + 0,5 \cos(t).$$

Начальное значение $y_d(0)=0,5$. Генеральные составляющие модели:

$$f(y, t) = -2 \frac{\dot{R}}{R} \dot{y} - g \frac{1}{R} \sin(y) + n(t), \quad b(y, t) = \frac{1}{mR^2}.$$

На рис. 2, а, б показаны кинематическая схема маятника и характер изменения коэффициентов модели.

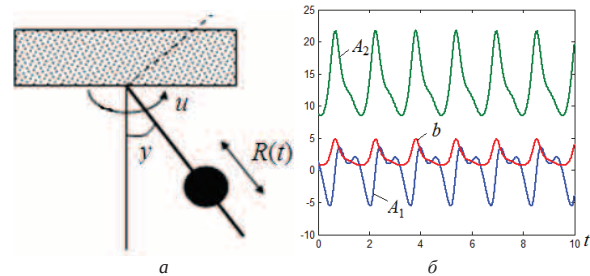


Рис. 2. а) кинематическая схема маятника; б) характер изменения коэффициентов модели

На рис. 2 $A_1=2\dot{R}/R$, $A_2=g/R$, $b=1/mR^2$. Требуемые показатели качества: перерегулирование $\sigma=0$ %; время установления $t_s \approx 1$ с для $\delta_s = \pm 2$ % допустимой ошибки регулирования.

Определим параметр настройки c_1 . При $n=2$ начальное условие $e(0)=y_d(0)-y(0)=(0,5;0,25)^T$. Уравнение (17) имеет вид:

$$f(p_1) = e(0)e^{-p_1 t_s} - 0,02 = 0,5e^{-p_1} - 0,02 = 0.$$

Используя функцию *solve()* находим $p_1=3,23$. Выражение (18) $D_{sys}=p+3,23 \Rightarrow c_1=3,23$. В результате, уравнение регулятора: $u=k(3,23e+\dot{e})$.

На рис. 3, а, б при $a=2$, $m=1$ показаны графики изменения ошибки регулирования $e(t)$ в зависимости от коэффициента усиления k и фазовый портрет системы.

При $k=120$ уже выполняются заданные показатели качества $\sigma=0$ %, $t_s=1$ с. Дальнейшее увеличение k не оказывает ощутимого влияния на показатели качества, однако приводит к увеличению энергии управления.

На рис. 3, б показан фазовый портрет системы при начальных условиях $y(0)=[2,1,-1,-2]$, $\dot{y}(0)=[2,2,-2,-2]$. Как видно, при $k=120$ все траектории стягиваются к линии $s=\dot{e}+3,23e=0$.

В начальный момент времени управление принимает довольно высокое значение $u(0)=-370$. Поэтому сигнал управления с помощью блока *Saturation* (насыщение) был ограничен на интервале $-60 \leq u(t) \leq 60$.

Рассмотрим случай параметрической неопределенности. Пусть коэффициент a при $\dot{y}(t)$ и масса груза m меняются на $\approx \pm 50$ %: $a=[1;1,5;2;2,5]$; $m=[0,5;1;1,5;2]$.

На рис. 4 показаны пучки переходных характеристик по выходу $\{y(t)\}$, ошибке $\{e(t)\}$ и сигналу управления $\{u(t)\}$ при $k=120$ и начальном условии $y(0)=(2;0)^T$.

Концентрация пучков $\{y(t)\}$ и $\{e(t)\}$ довольно высокая. Через $t_s \approx 1$ с выход объекта $\{y(t)\}$ начинает довольно точно отслеживать эталонную траек-

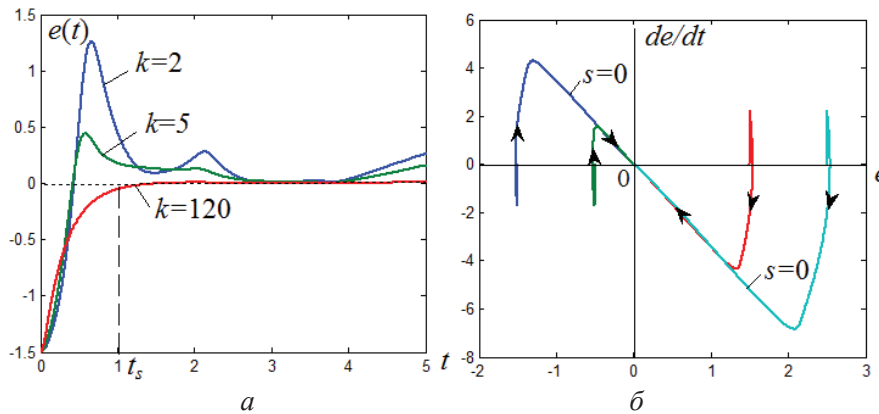


Рис. 3. а) графики изменения ошибки; б) фазовый портрет системы

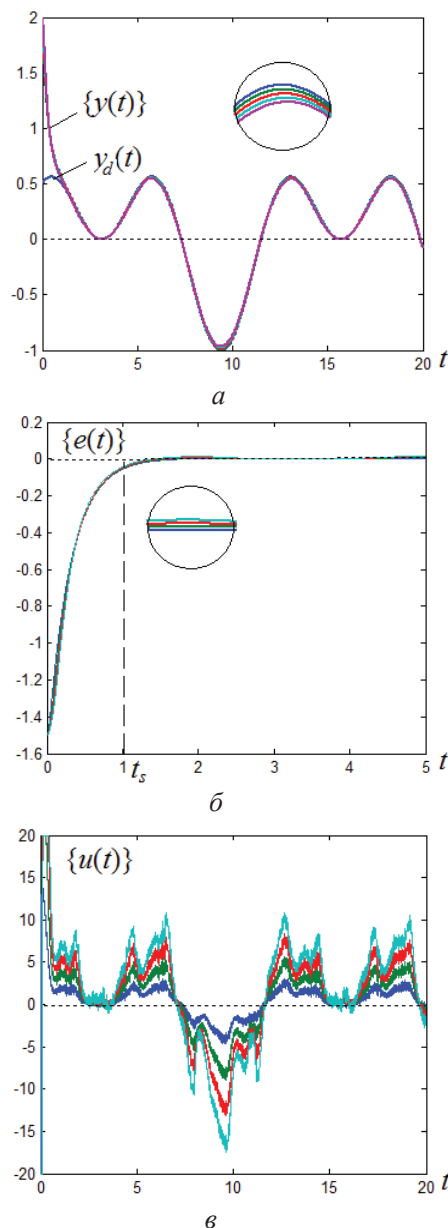


Рис. 4. Переходные характеристики (а), (б) и сигнал управления (в) при одновременном изменении параметров a и t

торию $y_d(t)$. В ослаблении действующих на вход объекта помех существенную роль играют фильтрующие свойства самого объекта.

Выводы

В работе рассмотрен один из возможных подходов к построению следящих робастных систем управления с высоким коэффициентом усиления. Основу синтеза составляет метод функции Ляпунова, в частности «метод квадратичных форм». Предложенная методика позволяет обойти сложные математические трудности, возникающие при синтезе робастных систем, и управлять нелинейными нестационарными объектами в условиях существенной неопределенности с помощью простого физически понятного регулятора.

К недостаткам предложенной методики следует отнести:

- отсутствие аналитической формулы для определения коэффициента усиления k ;
- робастный регулятор относительно ошибки регулирования получен в виде PD- (пропорционально-дифференциального) регулятора. При реализации таких регуляторов D-составляющие аппроксимируются реальным дифференцирующим звеном, что для систем высокого порядка может привести к ощутимым погрешностям. Если аппроксимацию не производить, то сигнал ошибки следует тщательно отфильтровать.
- происходит усиление высокочастотных помех, имеющих непосредственный доступ в регулятор. Этот недостаток является характерным для всех неадаптивных робастных систем;
- система настраивается на доминирующую неопределенность и должна «дежурить», ожидая появления наихудшего режима, оставаясь в напряжении. Этот недостаток можно исключить путем самонастройки коэффициента усиления регулятора, поддерживая его на минимальном уровне.
- не все объекты допускают реализацию чрезмерно высокого коэффициента усиления.

В результате литературного анализа можно прийти к следующему заключению общего характера:

- для обеспечения высоких робастных показателей коэффициент усиления регулятора должен принимать достаточно высокое значение. Это является диалектической закономерностью, поскольку связано с затрачиваемой энергией управления;
 - высокий коэффициент усиления в первую очередь вызывает чрезмерно высокое значение сигнала управления в начальный момент времени. Наложение же позиционного ограничения на управление снижает быстродействие и может привести к потере робастности;
- Преимущество неадаптивных систем заключается в следующем:
- используется незначительная информация об объекте – достаточно структурного представле-

ния модели. Исключение составляет H_∞ -теория, где используется номинальная модель объекта;

- не используются наблюдатели неопределенностей, которые уменьшают быстродействие системы и создают новые неудобства, связанные с известными проблемами сходимости в реальных условиях;
- имеется возможность построения простых робастных регуляторов, востребованных в промышленных приложениях.

Решение модельной задачи с использованием пакета *Matlab/Simulink* позволило сделать ряд положительных выводов, имеющих важное прикладное значение.

СПИСОК ЛИТЕРАТУРЫ

1. Андронов А.А., Поняткин Л.С. Грубые системы // Доклады АН СССР. – 1937. – Т. 14. – С. 247–249.
2. Наджаров А. Модель для сборки. URL: <http://alexnadzharov.ru> (дата обращения: 15.03.2014).
3. State-space solutions to standard H_2 and H_∞ control problems / J.C. Doyle, K. Glover, P.P. Khargonekar, B.A. Francis. IEEE Trans. Automat. Control. – 1989. – V. 34. – № 8. – P. 831–847.
4. Zhou K., Doyle J.C., Glover K. Robust and Optimal Control. – Upper Saddle River, NJ: Prentice Hall, 1996. – 586 p.
5. Позняк А.С. Основы робастного управления (H_∞ -теория). – М.: Изд-во МФТИ, 1991. – 128 с.
6. Пилишкин В.Н. Алгоритмы робастного управления в интеллектуальных системах // Вестник МГТУ. Серия приборостроение. – 1998. – № 1. – С. 23–28.
7. Емельянов С.В. Системы автоматического управления с переменной структурой. – М.: Наука, 1967. – 336 с.
8. Уткин В.И. Скользящие режимы в задачах оптимизации и управления. – М.: Наука, 1981. – 368 с.
9. Емельянов С.В., Коровин С.К. Новые типы обратных связей: Управление в условиях неопределенности. – М.: Наука, 1997. – 352 с.
10. Мееров М.В. Синтез структур систем автоматического управления высокой точности. – М.: Наука, 1967. – 423 с.
11. Меерков С.М. Принципы вибрационного управления: теория и применение // Автоматика и телемеханика. – 1973. – № 2. – С. 34–43.
12. Бейнарович В.А. Инвариантные системы автоматического управления с релейным усилителем // Доклады ТУСУРа. – 2010 (июнь). – № 1 (21). – Ч. 1. – С. 70–73.
13. Харитонов В.Л. Об асимптотической устойчивости положения равновесия семейства систем линейных дифференциальных уравнений // Дифференциальные уравнения. – 1978. – Т. 14. – № 11. – С. 2086–2088.
14. Кунцевич В.М. Управление в условиях неопределенности: гарантированные результаты в задачах управления и идентификации. – Киев: Наукова думка, 2006. – 264 с.
15. Rustamov G.A. Absolutely robust control systems // Automatic Control and Computer Sciences. – 2013. – V. 47. – № 5. – P. 227–241.
16. Потапенко Е.М., Казурова А.Е., Савранская А.В. Обзор работ по динамике многомассовых неопределенных электромеханических систем, выполненных на кафедре электропривода ЗНТУ // Электротехника та електроенергетика. – 2011. – № 1. – С. 7–8.
17. Грязина Е.Н., Поляк Б.Т., Тремба А.А. Синтез регуляторов низкого порядка по критерию H_∞ : Параметрический подход // Автоматика и телемеханика. – 2007. – № 3. – С. 94–105.
18. Xue D., Chen Y.Q., Atherton D.P. Linear Feedback Control. Analysis and Design with MATLAB. – Philadelphia: Society for Industrial and Applied Mathematics, 2007. – 354 p.
19. Шеленок Е.А. Адаптивно-робастная система управления нелинейными объектами периодического действия // Информатика и системы управления. – 2012. – № 4 (34). – С. 128–137.
20. Robust Control Toolbox™ Getting Started Guide / G. Balas, R. Chiang, A. Packard, M. Safonov. – Natick, MA: Math Works, Inc., 2005–2010. – 135 p.
21. Шевцов Г.С. Линейная алгебра. – М.: Гардарики, 1999. – 360 с.
22. Mamedov G.A., Rustamov G.A., Rustamov R.G. Construction of a Logical Control by Means of Optimization of the Function When an Object Model is Indeterminate // Automatic Control and Computer Sciences. – 2010. – V. 44. – № 3. – P. 119–123.
23. Рустамов Г.А., Абдуллаева А.Т., Рустамов Р.Г. Оптимизация функции Ляпунова при релейном управлении динамическими объектами // Автоматизация и Современные Технологии. – 2013. – № 3. – С. 21–25.
24. Levant A. Principles of 2-sliding mod design // Automatica (Journal of IFAC). – 2007. – V. 43. – Iss. 4. – P. 576–586.

Поступила 23.03.2014 г.

ROBUST CONTROL SYSTEM WITH HIGH POTENTIAL

Gazanfar A. Rustamov,

Dr. Sc., Azerbaijan Technical University, 25, G. Javid avenue,
Baku, Az1073, Azerbaijan. E-mail: gazanfar.rustamov@gmail.com

The paper introduces the approach to control of a priori uncertain dynamic objects based on Lyapunov function method, in particular «the method of quadratic forms». Possibility of infinite raising the gain coefficient of PD-controller of $n-1$ order allows suppressing general components of uncertain model to an arbitrarily small value. This ensures high accuracy of tracking of the reference trajectory for a wide class of uncertainties. The system is described by a homogeneous linear equation of the hyper-plane in the limit. This feature allows defining the controller settings parameters based on quality indicators using the «modal control» method. The proposed technique allows avoiding the complex mathematical difficulties occurring in the synthesis of robust control systems and controlling nonlinear non-stationary objects under substantial uncertainty by a simple physically intuitive controller. The disadvantages of the proposed technique are the increased frequency interference with direct access to the controller, the usage of the output derivatives for forming PD-controller as well. Solving model problems on Matlab/Simulink allowed the author to make a number of positive findings of the important practical value.

Key words:

Uncertainty, robust tracking system, Lyapunov function, partitura model, high gain coefficient, parametric pendulum.

REFERENCES

- Andronov A.A., Pontryagin L.S. Grubyie sistemy [Rude systems]. *Reports of AoS of USSR*, 1937, vol. 14, pp. 247–249.
- Nadzharov A. *Model dlya sborki* [Installation Model]. Available at: <http://alexnadzharov.ru> (accessed 15 March 2014).
- Doyle J.C., Glover K., Khargonekar P.P., Francis B.A. State-space solutions to standard and H_∞ control problems. *IEEE Trans. Automat. Control*, 1989, vol. 34, no. 8, pp. 831–847.
- Zhou K., Doyle J.C., Glover K. *Robust and Optimal Control*. Upper Saddle River, NJ, Prentice Hall, 1996. 586 p.
- Poznyak A.S. *Osnovy robastnogo upravleniya (H_∞ -teoriya)* [Fundamentals of robust control]. Moscow, Publishing house of Moscow Physical-Technical University, 1991. 128 p.
- Pilishkin V.N. Algoritmy robastnogo upravleniya v intellektualnykh sistemakh [Robust control algorithms in intelligent systems]. *Vestnik MGTU. Seriya priborostroenie – Bulletin of MSTU. Instrumentation series*, 1998, no. 1, pp. 23–28.
- Emelyanov S.V. *Sistemy avtomaticheskogo upravleniya s pere-mennoy strukturoy* [Automatic control system with variable structure]. Moscow, Nauka Publ., 1967. 336 p.
- Utkin V.I. *Skolzhashchie rezhimy v zadachakh optimizatsii i upravleniya* [Sliding modes in Optimization and Control Problems]. Moscow, Nauka Publ., 1981. 368 p.
- Emelyanov S.V., Korovin S.K. *Novye tipy obratnykh svyazey: Upravleniye v usloviyakh neopredelennosti* [New types of feedback: Control under uncertainty]. Moscow, Nauka Publ., 1997. 352 p.
- Meerov M.V. *Sintez struktur sistem avtomaticheskogo upravleniya vysokoy tochnosti* [Synthesis of structures of automatic control systems with high precision]. Moscow, Nauka Publ., 1967. 423 p.
- Meerkov S.M. Printsipy vibratsionnogo upravleniya: teoriya i primeneniye [Principles of vibration control: theory and application]. *Avtomatika i telemekhanika – Automation and telemekhanics*, 1973, no. 2, pp. 34–43.
- Beynarovich V.A. Invariantnye sistemy avtomaticheskogo upravleniya s reley-nym usilitелеm [Invariant automatic control system with switching amplifier]. *Doklady TUSURa – Reports of Tomsk State University of Control Systems and Radioelectronics*, 2010 (June), no. 1 (21), P. 1, pp. 70–73.
- Kharitonov V.L. Ob asimptoticheskoy ustoychivosti polozheniya ravновесiya semeystva sistem lineynykh differentsialnykh uravneniy [On asymptotic stability of equilibrium of the family of linear differential equations systems]. *Differentsialnye uravneniya – Differential equations*, 1978, vol. 14, no. 11, pp. 2086–2088.
- Kuntsevich V.M. *Upravlenie v usloviyakh neopredelennosti: garantirovannye rezultaty v zadachakh upravleniya i identifikatsii* [Control under uncertainty: guaranteed results in problems of control and identification]. Kiev, Naukova Dumka Publ., 2006. 264 p.
- Rustamov G.A. Absolutely robust control systems. *Automatic Control and Computer Sciences*, 2013, vol. 47, no. 5, pp. 227–241.
- Potapenko E.M., Kazurova A.E., Savranskaya A.B. Obzor rabot po dinamike mnogomassovykh neopredelennykh elektromekhanicheskikh sistem, vypolnennykh na kafedre elektroprivoda ZNTU [Studies on dynamics of multibody uncertain electromechanical systems performed at the Department of electric drive of ZNTU]. *Elektrotehnika i elektroenergetika – Electrical engineering and electricity*, 2011, no. 1, pp. 7–8.
- Gryazina E.N., Polyak B.T., Tremba A.A. Sintez regulatorov niz-kogo poradka po kriteriyu H_∞ : Parametricheskii podkhod [Synthesis of low-order controller by H_∞ criteria: parametric approach]. *Avtomatika i telemekhanika – Automation and Telemechanics*, 2007, no. 3, pp. 94–105.
- Xue D., Chen Y.Q., Atherton D.P. *Linear Feedback Control. Analysis and Design with MATLAB*. Philadelphia, Society for Industrial and Applied Mathematics, 2007. 354 p.
- Shelenok E.A. Adaptivno-robastnaya sistema upravleniya nelineynymi obektami periodicheskogo deystviya [Adaptive robust control system of nonlinear batch objects]. *Informatika i sistemy upravleniya – Computer Science and Control Systems*, 2012, no. 4 (34), pp. 128–137.
- Balas G., Chiang R., Packard A., Safonov M. *Robust Control Toolbox™ Getting Started Guide*. Natick, MA, Math Works, Inc., 2005–2010. 135 p.
- Shevtsov G.S. *Lineynaya algebra* [Linear algebra]. Moscow, Gardariki Publ., 1999. 360 p.
- Mamedov G.A., Rustamov G.A., Rustamov R.G. Construction of a Logical Control by Means of Optimization of the Function When an Object Model is Indeterminate. *Automatic Control and Computer Sciences*, 2010, vol. 44, no. 3, pp. 119–123.
- Rustamov G.A., Abdullaeva A.T., Rustamov R.G. Optimizatsiya funktsii Lyapunova pri reley-nom upravlenii dinamichskimi obektami [Optimization of Lyapunov function by the relay control of dynamic objects]. *Asovremennyye tekhnologii – Automation and Modern Technology*, 2013, no. 3, pp. 21–25.
- Levant A. Principles of 2-sliding mode design. *Automatica (Journal of IFAC)*, 2007, vol. 43, Iss. 4, pp. 576–586.

УДК 004.5;620.162.4;620.193.21;681.513.2

АВТОМАТИЗИРОВАННАЯ СИСТЕМА КОНТРОЛЯ ПАРАМЕТРОВ КЛИМАТИЧЕСКОГО И ТЕРМОВАКУУМНОГО ОБОРУДОВАНИЯ ДЛЯ ИСПЫТАНИЯ ИЗДЕЛИЙ КОСМИЧЕСКОЙ ТЕХНИКИ

Андросов Владимир Яковлевич,

начальник отдела климатических, механических и электрических испытаний открытого акционерного общества «Научно-производственный центр "Полус"», Россия, 634050, г. Томск, пр. Кирова, 56 «в». E-mail: polus@online.tomsk.net

Плотников Андрей Владимирович,

ведущ. инженер-программист отдела климатических, механических и электрических испытаний открытого акционерного общества «Научно-производственный центр "Полус"», Россия, 634050, г. Томск, пр. Кирова, 56 «в». E-mail: polus@online.tomsk.net

Вагайцев Семен Александрович,

инженер-электроник отдела климатических, механических и электрических испытаний открытого акционерного общества «Научно-производственный центр "Полус"», Россия, 634050, г. Томск, пр. Кирова, 56 «в». E-mail: polus@online.tomsk.net

Будников Евгений Анатольевич,

генеральный директор общества с ограниченной ответственностью «ПРОМХОЛ», Россия, г. Москва, 105062, Лялин пер., д. 20. E-mail: promhol@mail.ru

Актуальность работы обусловлена необходимостью автоматизации процесса термовакuumных и климатических испытаний изделий космической техники на предприятии космической отрасли.

Цель работы: создание специализированной системы контроля режимов и параметров процессов термовакuumных и климатических испытаний, обоснование выбранных технологий сбора, обработки первичных данных о заданном и текущем испытательном профиле, а также данных о техническом состоянии применяемых средств измерений и испытательного оборудования, имеющих различное исполнение.

Методы и технологии: анализ технологических процессов испытаний изделий, проектирование и макетирование автоматизированных систем контроля и управления, проведение испытаний в программном режиме, анализ заданного испытательного профиля, параллельное программирование, объектно-ориентированное программирование, связывание и внедрение объектов (OLE), хранение и обработка XML-данных, человеко-машинное взаимодействие, клиент-серверное взаимодействие.

Результаты: создана автоматизированная система контроля параметров испытаний, разработаны ее основные программно-технические решения и компоненты, проведена оценка эффективности их применения по первым результатам эксплуатации.

Выводы: применение автоматизированной системы, включающей до нескольких десятков единиц испытательного оборудования, объединенных в сеть и управляемых из единого центра процессом термовакuumных и климатических испытаний изделий, обеспечивает экономическую эффективность, высокий организационно-технический и метрологический уровень наземной обработки изделий космической техники.

Ключевые слова:

Климатические и термовакuumные испытания, внешние воздействующие факторы, система охлаждения и нагрева жидкости, косвенный нагрев и охлаждение жидкости, каскадная холодильная установка, диспетчерское управление и сбор данных, человеко-машинный интерфейс, контроллер-регулятор, программируемый логический контроллер, автоматизация проведения испытаний, автоматизированная система контроля параметров испытаний.

В настоящее время к изделиям космической техники нового поколения предъявляются повышенные требования по точностным, ресурсным и эксплуатационным характеристикам, в частности по стойкости изделий к факторам открытого космического пространства, напрямую влияющим на надежность и ресурс космического аппарата и характеризующимся температурами до минус 70 °С и атмосферным давлением до $1 \cdot 10^{-6}$ мм рт. ст.

Имитация космического пространства при наземной отработке изделий осуществляется путем

проведения испытаний на прочность и стойкость к термовакuumным и климатическим факторам [1].

Испытания проводятся с применением специализированного термовакuumного испытательного оборудования (ИО), технологической оснастки, электронных и электромеханических нагрузочных устройств, а также программно-технических средств измерений [2].

Изделие во включенном состоянии испытывается в различных режимах работы, при этом изменяются как параметры нагрузок, так и климатиче-

ский и вакуумный профили [3, 4]. Процесс выполняется непрерывно с различным количеством и длительностью циклов при общей продолжительности от нескольких суток до месяца.

Таким образом, испытания являются технически сложными и дорогостоящими, их реализация в полном объеме предъявляет повышенные требования как к применяемым техническим системам, так и к персоналу испытательной лаборатории.

С целью комплексной обработки разрабатываемых изделий на стойкость к факторам открытого космического пространства, эффективного использования применяемых технических средств и минимизации влияния инженерно-технического персонала на процесс в ОАО «НПЦ «Полус» создана автоматизированная система контроля параметров испытаний (АС КПИ).

АС КПИ (рис. 1) представляет собой ИО, объединенное промышленным интерфейсом в единую сеть и подключенное к управляющему компьютеру (ПК) с разработанным специализированным программным обеспечением (ПО).

Рассмотрим одну из типовых функциональных схем отдельно взятой единицы ИО (рис. 2), входящей в состав АС КПИ. Блоки испытуемого изделия устанавливаются в вакуумную камеру на «этажерку» соединенных определенным образом между собой термостабилизированных плит, входные и выходные фланцы которых подключены посредством термостойких шлангов к программно-управляемой системе охлаждения и нагрева жидкости.

В системе охлаждения и нагрева жидкости реализован принцип косвенного нагрева и охлаждения термостабилизированных плит, установленных в вакуумной камере, позволяющий достигать указанных температур поверхности с использованием специального теплоносителя.

Необходимая теплоотводящая мощность плит рассчитывается исходя из удельной тепловой мощности, выделяемой блоками изделия на «этажерку» плит, и соответствует производительности системы охлаждения и нагрева жидкости, выраженной в единицах тепловой мощности нагрева и охлаждения в единицу времени согласно неравенству

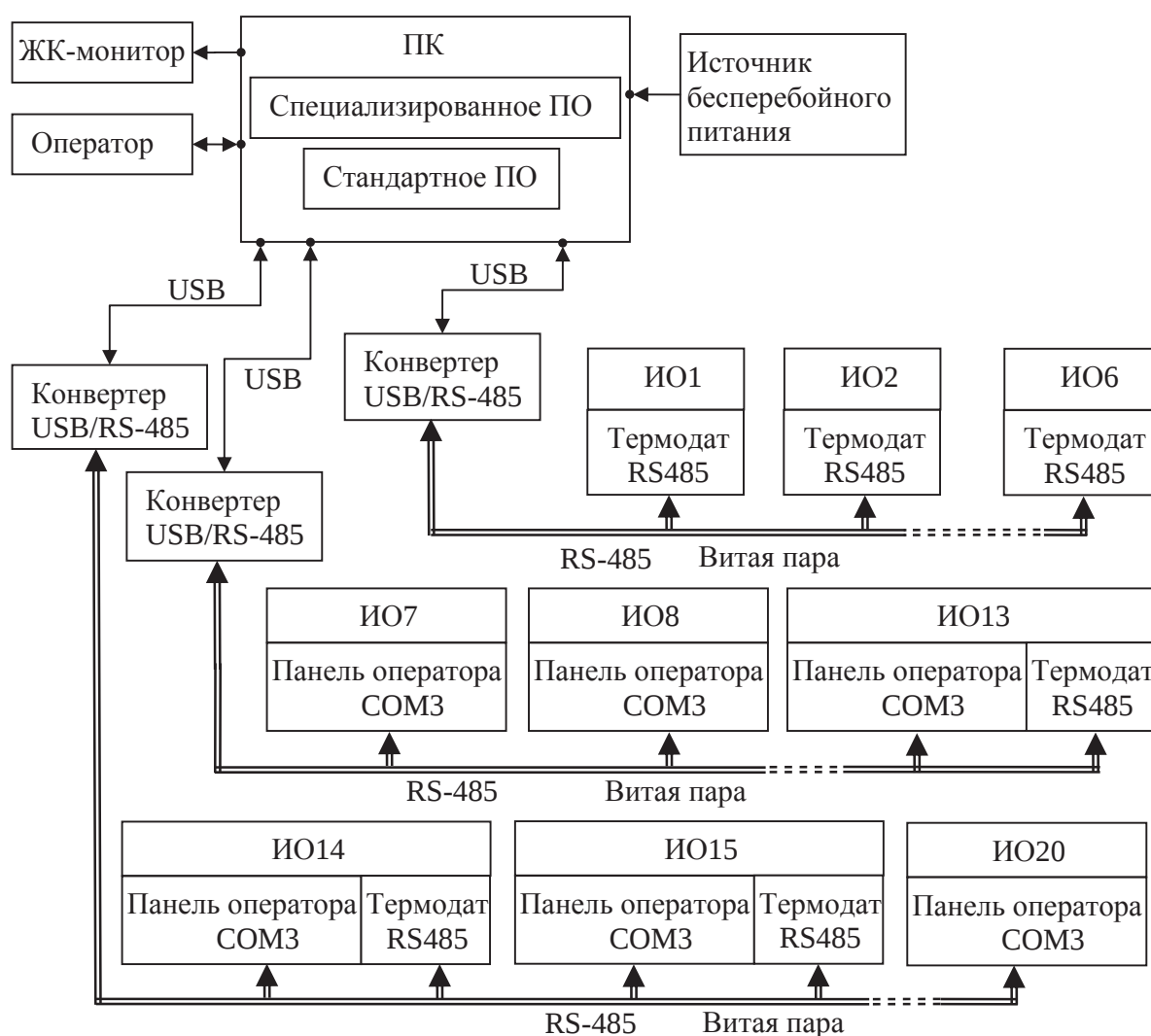


Рис. 1. Функциональная схема АС КПИ

$$P_{\text{изд}} < P_{\text{пл}} < P_{\text{сонж}},$$

где $P_{\text{изд}}$ – тепловая мощность, выделяемая изделием; $P_{\text{пл}}$ – теплоотводящая мощность плиты; $P_{\text{сонж}}$ – производительность системы охлаждения и нагрева жидкости.

Тепловые расчеты указанного процесса не приводятся, поскольку выходят за рамки данной статьи.

Изготовление термостабилизированных плит – сложный технологический процесс, заключающийся в заливке алюминия в специальные формы с установленными внутри трубками сложной геометрической конфигурации.

«Термоконтрукция», состоящая из блоков изделия и плит, масса которой достигает 500 кг, посредством грузоподъемного механизма опускается в вертикальную вакуумную камеру большого объема, и изделие циклически испытывается в течение длительного периода времени.

Испытания проводятся по заданному в соответствии с требованиями испытательному профилю, имеющему граничные значения:

- по температуре: от -70 до $+150$ °С;
- по относительной влажности: от 10 до 98 % при температуре от 10 до 95 °С со скоростью изменения не менее 1 °С/мин;
- по давлению: от 760 до $1 \cdot 10^{-6}$ мм рт. ст. при температуре от -60 до $+120$ °С, с точностью поддержания температуры $1-2$ °С.

Для реализации заданных тепловых режимов испытаний в ИО вместе с системами управления используются каскадные холодильные машины, электрические нагреватели, системы осушения и увлажнения, позволяющие с высокой точностью (по температуре $\pm(0,1-0,3)$ °С и влажности 1–3 %) поддерживать заданные параметры.

Как правило, на предприятии одновременно проходят испытания изделий различных типов, что требует значительного количества испытатель-

ного, технологического оборудования, программно-технических средств и высококвалифицированного персонала.

В целях оптимизации процесса испытаний, с учетом вышеизложенного, была спроектирована автоматизированная система, включающая до 20 единиц ИО (потенциально до нескольких десятков), структурированную кабельную сеть, конвертеры интерфейса и разработанное специализированное ПО.

АС КПИ в реальном масштабе времени контролирует процесс климатических и термовакуумных испытаний во всех единицах ИО, включенных в сеть, с отслеживанием текущих параметров каждой единицы ИО в графическом и табличном виде и возможностью изменения оператором профиля, останова и продолжения испытаний, а также с автоматическим архивированием результатов на жесткий диск ПК.

При разработке АС КПИ с целью оценки производительности и масштабируемости данной системы проведено ее макетирование, а также рассмотрены различные подходы к созданию автоматизированных систем управления технологическими процессами [5–9]. Для управления такими сложными процессами средства сбора и передачи технологической информации должны иметь высокую производительность, при этом обладать надежностью и безопасностью.

Авторами рассмотрены различные спецификации обмена данными с технологическим оборудованием и пакеты программного обеспечения, реализующие эти спецификации [10]. Несмотря на насыщенность рынка программным обеспечением для обмена данными с промышленным оборудованием [11], нет программного продукта, согласованного с системами управления применяемого ИО.

Для внедрения существующих на рынке систем диспетчерского управления и сбора данных [12, 13] необходима как минимум поддержка на

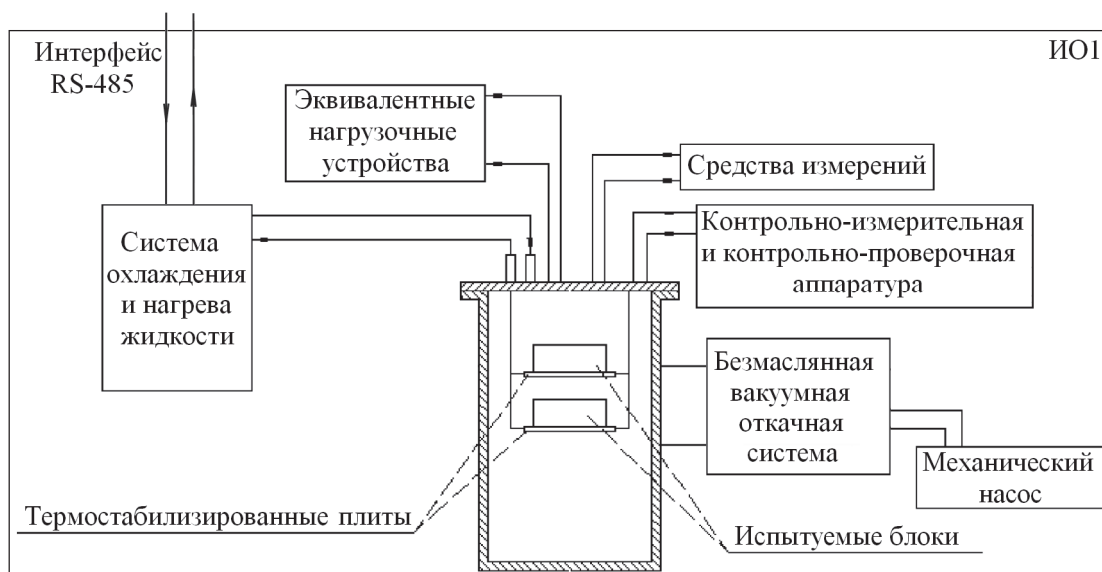


Рис. 2. Функциональная схема типовой единицы ИО

уровне драйверов со стороны производителя систем управления ИО.

С учетом уникальности систем управления эксплуатируемых единиц ИО, их различного исполнения и, как следствие, отсутствия драйверов для передачи данных в общеизвестные автоматизированные системы, а также особенностей процесса термовакуумных и климатических испытаний, авторами было решено разработать автоматизированную систему на базе стандартных промышленных интерфейсов, при этом принимая во внимание опыт разработки подобных систем другими компаниями [14–16].

Для создания АС КПИ климатическое и термовакуумное оборудование объединено в единую сеть интерфейсом RS-485 (стандарт TIA/EIA-485) с подключением по нескольким линиям к ПК посредством конвертеров интерфейса USB-RS485 (тип IFD6500 фирмы DELTA Electronics) [17]. Таким образом, сеть организована в соответствии с двумя базовыми сетевыми топологиями – «звезда» и «шина».

Интерфейс RS-485 использует балансную (дифференциальную) схему передачи сигнала [18]. Максимальная скорость передачи данных 115 200 бит/с по линии длиной до 1000 м.

Конвертер интерфейса USB-RS-485 [19] предназначен для обеспечения связи ПК, имеющего интерфейс USB, с удаленными объектами по интерфейсу RS-485. Устройство преобразует сигналы USB-порта компьютера в сигналы интерфейса RS-485 и направляет их к удаленному объекту по каналу связи (витой паре).

Подключение ИО, распределенного на значительной площади, по нескольким линиям к ПК позволило минимизировать длину кабельной сети и обеспечить передачу данных без искажений.

Обмен данными в сети между ПК и ИО выполняется по протоколу MODBUS-ASCII [20, 21].

На базе программируемых логических контроллеров (базовый блок типа DVP-14SS с модулями расширения типа DVP-14SP), контроллеров температуры (базовый блок типа DTC-1000 с модулями расширения типа DTC-2000), человеко-машинных интерфейсов (панелей оператора) типа DOP-B и источников питания (24 В, 60 Вт) типа DRP024V060W1AZ фирмы DELTA Electronics, регуляторов температуры типа Термодат-16ЕЗ фирмы «Системы контроля», входных устройств, многоканальных микропроцессорных регистраторов типа PMT 39DM/3 фирмы «Элемер» и исполнительных устройств (реле), являющихся частями системы управления единицы ИО (рис. 3), и ПК разработано специализированное ПО автоматизированной системы (рис. 4).

Для отладки и тестирования автоматизированной системы разработано ПО, эмулирующее всю сеть с задаваемым количеством единиц ИО [22].

Подключение каждой единицы ИО выполнялось в следующем порядке:

- настройка параметров и модификация ПО, встроенного в систему управления ИО, с целью реализации функций удаленного управления им и передачи параметров в ПК (значения климатических и термовакуумных параметров, информация о состоянии ИО (включено или выключено), информация об авариях и т. д.);
- внесение ИО в таблицу опроса, заключающееся в задании в специализированном ПО идентификаторов измерительных каналов, типа каналов (контрольный или измерительный), типа измеряемых параметров (температура, влажность или давление), цвета линий на графике испытательного профиля;

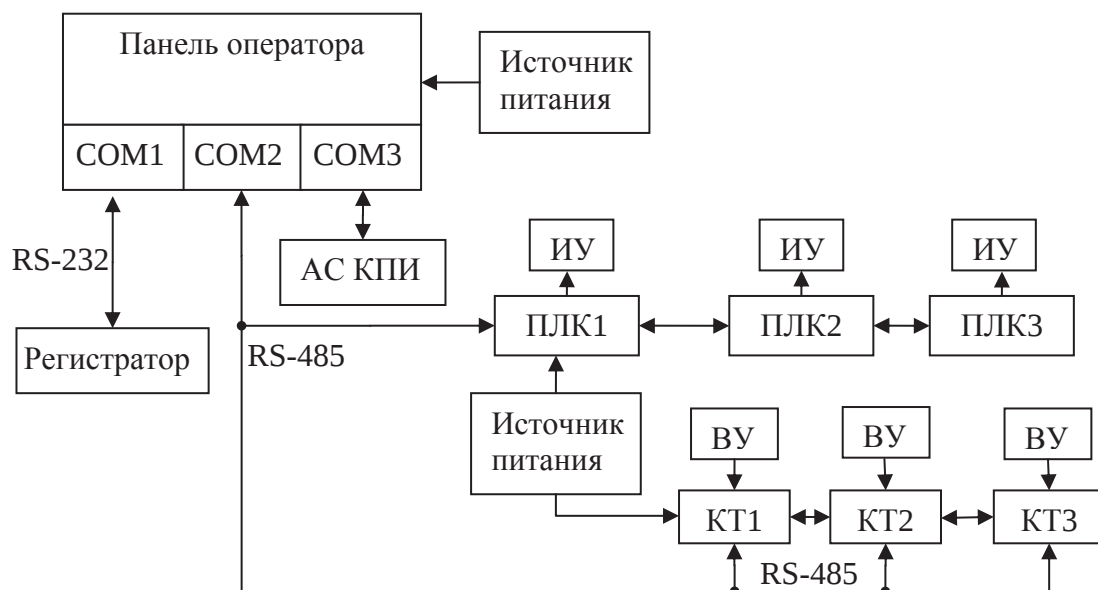


Рис. 3. Функциональная схема системы управления климатическим и термовакуумным ИО на основе микропроцессорной техники фирмы DELTA Electronics: ИУ – исполнительное устройство; ПЛК – программируемый логический контроллер; ВУ – входное устройство; КТ – контроллер температуры

- нанесение изображения ИО на план помещения.

ПО АС КПИ разработано в среде программирования с открытым кодом Lazarus, работает в операционных системах Microsoft Windows 2000/XP/Vista/7, имеет удобный интерфейс взаимодействия с пользователем, модульную структуру, открытую для реализации дополнительных функций, поддерживает взаимодействие с различным климатическим и термовакуумным ИО, имеющим интерфейс RS-485.

Для хранения данных широко применяется технология XML [23]. ПО автоматизированной си-

стемы непрерывно с задаваемой периодичностью опрашивает всю сеть ИО и сохраняет данные на жесткий диск ПК. Работа пользовательского интерфейса и модулей ПО, отвечающих за такие ресурсоемкие задачи, как сбор данных и расчет профиля, выполняется параллельно, что позволяет задействовать преимущества многопроцессорных (многоядерных) ПК [24].

ПО АС КПИ в реальном масштабе времени опрашивает всю сеть ИО, при этом во вложенном окне (рис. 5) на графике отображается процесс изменения во времени значений одного из выбирае-

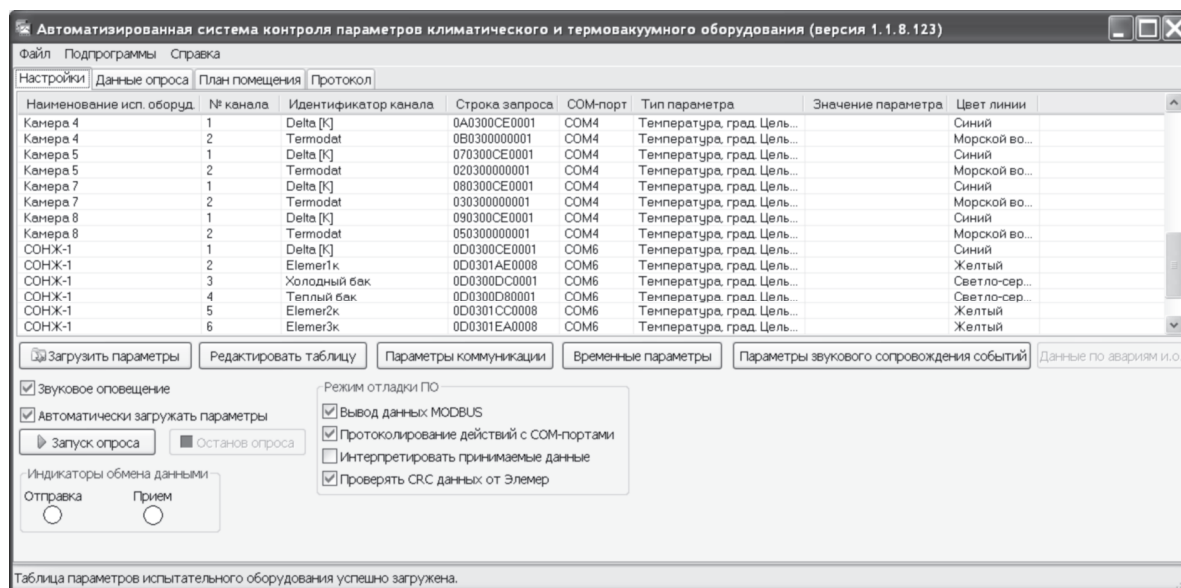


Рис. 4. Основное окно программного обеспечения автоматизированной системы с заполненной таблицей параметров подключенного испытательного оборудования

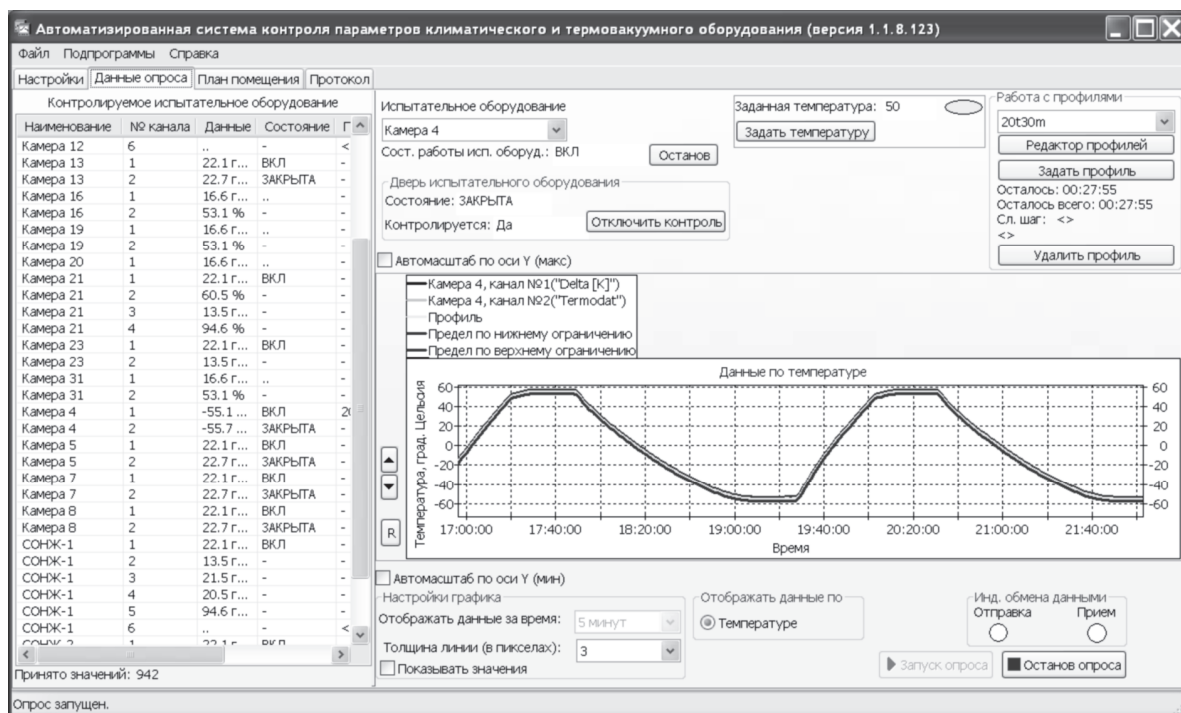


Рис. 5. Окно с данными контроля параметров функционирующей сети испытательного оборудования

мых оператором параметров ИО (температура, влажность, давление), а также линии профиля и заданных ограничений.

При выборе оператором наименования ИО из списка в этом окне дополнительно отображается информация о его работе:

- значения заданных параметров ИО (температура, влажность, давление) и индикация их достижения;
- состояние ИО (включено или отключено);
- состояние исполнительных переключателей ИО, а также информация об их контроле («да» или «нет»).

В случае возникновения аварии в ИО, открытии двери ИО при работе, а также отклонении контролируемых параметров от допустимых значений ПО АС КПИ осуществляет соответствующее аудиовизуальное оповещение.

Также необходимо отметить встроенные функции безопасности при работе с ПО АС КПИ: протоколируются все действия оператора, а при включении/выключении, изменении значений параметров ИО у оператора запрашивается код подтверждения.

Большое внимание уделено надежности функционирования АС КПИ [25], в частности дополнительно проверяется информация при обмене данными с ИО, применяется кодирование данных с вычислением их контрольных сумм.

Непрерывная работа АС КПИ в круглосуточном режиме в течение длительного времени подтверждает высокий уровень безотказности АС КПИ и ее устойчивость к некорректным действиям обслуживающего персонала, различным состояниям ИО, в том числе к аварийным ситуациям и т. д.

Посредством реализованной подпрограммы обработки измеренных данных по испытаниям различных изделий можно просмотреть значения климатических и термовакуумных параметров за долгий период времени (месяц) и сформировать отчет в формате приложений Microsoft Office.

Разработанная АС КПИ позволяет:

- выполнять непрерывный сбор данных от многоканальных систем управления с графической визуализацией в реальном масштабе времени;
- работать с различными входными устройствами: датчиками температуры, влажности, давления, газового состояния среды и др.;
- задавать испытательный профиль в табличном и графическом режиме с помощью редактора режимов;
- осуществлять непрерывный контроль значений климатических и термовакуумных параметров на соответствие заданному испытательному профилю, что оказывает непосредственное влияние на метрологические характеристики текущих испытаний;

- осуществлять аудиовизуальное оповещение о выходе контролируемых параметров за допустимые пределы, отображаемые графически красными линиями;
- управлять ИО с рабочего места оператора (задавать значения температуры, влажности и давления, производить удаленный запуск/останов оборудования);
- отображать на плане помещения значения контролируемых параметров и информацию о состоянии работы (включено, отключено, авария) для каждой единицы ИО;
- отслеживать действия оператора в автоматизированной системе и сохранять информацию о них на жестком диске с формированием протокола.

В настоящее время АС КПИ проходит наработку и метрологическую аттестацию, необходимые для внедрения в производственный процесс в ОАО «НПЦ "Полус"».

Выводы

В ОАО «НПЦ "Полус"» исследована задача автоматизации процесса термовакуумных и климатических испытаний изделий космической техники.

По результатам проведенных исследований предложены решения, базирующиеся на применении микропроцессорных систем управления, встроенных в испытательное оборудование, и промышленных стандартах обмена информации.

Практика предварительной эксплуатации созданной на предприятии автоматизированной системы контроля процесса испытаний показывает, что обеспечивается высокопроизводительный и надежный обмен данными технологических параметров и управляющими командами в реальном масштабе времени, а также данными о текущем состоянии климатического и термовакуумного оборудования.

Эффективность применения автоматизированной системы контроля параметров испытаний в целом можно оценить по следующим показателям:

- значительно возрастает эффективность и надежность применяемого испытательного оборудования;
- уменьшается (минимизируется) вероятность ошибок и несанкционированного воздействия на изделие в процессе испытаний;
- улучшаются метрологические характеристики проводимых испытаний, и обеспечивается возможность их постоянного контроля;
- сокращается количество занятого инженерно-технического персонала.

Внедрение в НПЦ «Полус» АС КПИ позволит значительно повысить качество климатических и термовакуумных испытаний изделий космической техники при уменьшении затрат на их проведение.

СПИСОК ЛИТЕРАТУРЫ

1. Разработка стенда для исследования стойкости материалов к воздействию факторов космического пространства / С.В. Янович, Ю.Н. Литвак, П.А. Михалев, В.М. Башков // Инженерный вестник. – 2012. – № 11. URL: http://engbul.bmstu.ru/file/511839.html?_s=1 (дата обращения: 20.01.2014).
2. Комплексные системы для отработки и испытания космических аппаратов / С.В. Кравченко, С.Б. Нестеров, В.А. Романько, Н.А. Тестоедов, В.И. Халиманович, В.В. Христин // Вакуумная техника и технология. – 2012. – № 4. – С. 213–230.
3. Термовакuumные испытания системы электроснабжения на космическом аппарате / Б.М. Рассмакин, С.М. Хайранов, В.И. Хоминич, А.В. Буденный, Е.Ю. Коваленко, Е.Н. Елисеева // Технология и конструирование в электронной аппаратуре. – 2011. – № 5. – С. 6–10.
4. Термовакuumные испытания полых композитных стержней для космических конструкций / С.В. Резник, О.В. Денисов, П.В. Просунцов, В.П. Тимошенко, А.В. Шуляковский // Все материалы: Энциклопедический справочник. – 2012. – № 7. – С. 8–12.
5. Бобырь М., Титов В., Емельянов С. Теоретические основы построения автоматизированных систем управления технологическими процессами на основе нечеткой логики. – Старый Оскол: ТНТ, 2011. – 232 с.
6. Иванов А. Проектирование автоматизированных систем манипулирования объектами обработки и сборки. – М.: Форум, 2012. – 352 с.
7. Копцюх Е.С. Основные задачи этапа обследования в рамках проектов автоматизации // Прикладная информатика. – 2010. – № 6. – С. 10–17.
8. Никонов А.В. Внедрение средств автоматизации в процессы контроля и управления на производстве // Технические науки – от теории к практике. – 2012. – № 7-1. – С. 127–131.
9. Реализация человеко-машинного интерфейса в интегрированной технологии проектирования автоматизированных систем контроля и управления / Г.А. Егоров, А.Д. Белоногов, М.А. Островский, Я.А. Рейзман // Мехатроника, автоматизация, управление. – 2011. – № 7. – С. 56–62.
10. Козлецов А.П., Решетников И.С. Современные способы организации обмена данными с системами управления // Информационные технологии в проектировании и производстве. – 2010. – № 2. – С. 17–23.
11. Вейбер В.В., Кудинов А.В., Марков Н.Г. Задача сбора и передачи технологической информации распределенного промышленного предприятия // Известия Томского политехнического университета. – 2011. – Т. 319. – № 5. – С. 69–74.
12. Реймген Ю.Э. Автоматизированные системы управления технологическими процессами. SCADA система // Научный вестник Московского государственного горного университета. – 2012. – № 12. – С. 114–132.
13. SCADA // Википедия. Свободная энциклопедия. 2013. URL: <http://ru.wikipedia.org/wiki/SCADA> (дата обращения: 20.01.2014).
14. Черпаков Б.А., Стрижевский М.Н. Многоканальная автоматизированная система для измерения температуры при проведении теплопрочностных испытаний // Вестник Московского авиационного института. – 2010. – Т. 17. – № 4. – С. 27–31.
15. Синтез интервальной системы управления с адаптивным ПИД-регулятором / В.Ф. Лубенцов, Е.В. Лубенцова, Г.В. Масюткина, Т.А. Рудакова // Автоматизация в промышленности. – 2013. – № 9. – С. 62–64.
16. Баскур О., Тюнякин А.В., Хертлер К. Диагностика и мониторинг оборудования по состоянию. Мониторинг состояния оборудования в реальном времени // Автоматизация в промышленности. – 2012. – № 9. – С. 26–29.
17. Руководство по построению систем управления на базе продукции Delta Electronics (сети Modbus, CANOpen, Ethernet) // ООО «НПО "Стоик Лтд"». 2013. URL: http://www.stoikltd.ru/images/manual/Training_kit_OP_Manual.pdf (дата обращения: 20.01.2014).
18. Термодат-16Е5. Руководство пользователя // НПП «Системы контроля». 2013. URL: <http://termodat.ru/pics/16E5.pdf> (дата обращения: 20.01.2014).
19. IFD-6500. Инструкция по применению // ООО «НПО "Стоик Лтд"». 2011. URL: http://www.stoikltd.ru/images/manual/module/IFD6500_manual_rus.pdf (дата обращения: 20.01.2014).
20. Инструкция по сопряжению контроллеров Delta DVP со SCADA-системами // ООО «НПО "Стоик Лтд"». 2013. URL: http://www.stoikltd.ru/images/manual/DVP_and_SCADA.pdf (дата обращения: 20.01.2014).
21. Коммуникационный модуль RTU-485 для удаленного сбора данных от модулей расширения контроллеров Delta DVP серии S по протоколу Modbus // ООО «НПО "Стоик Лтд"». 2013. URL: http://www.stoikltd.ru/images/manual/modules/RTU_485_manual_rus.pdf (дата обращения: 20.01.2014).
22. Использование режима Virtual Com (виртуальный COM-порт) в коммуникационном модуле IFD9506 // ООО «НПО "Стоик Лтд"». 2013. URL: http://www.stoikltd.ru/images/manual/module/IFD9506_Virtual_com_rus.pdf (дата обращения: 20.01.2014).
23. Гринев М., Щеклеин И. Ориентированные на приложения методы хранения XML-данных // Труды ИСП РАН. – 2009. – Т. 16. – С. 143–154.
24. Немнюгин С.А., Стесик О.Л. Параллельное программирование для многопроцессорных вычислительных систем. – СПб.: БХВ-Петербург, 2002. – 400 с.
25. Шишмарев В. Диагностика и надежность автоматизированных систем. – М.: Академия, 2013. – 352 с.

Поступила 10.02.2014 г.

UDC 004.5;620.162.4;620.193.21;681.513.2

AUTOMATED CONTROL SYSTEM FOR CLIMATIC AND THERMOVACUUM EQUIPMENT PARAMETERS TO TEST SPACE INDUSTRY PRODUCTS

Vladimir Ya. Androsov,

«Scientific and Production Center "Polus"», 56 «v», Kirov avenue,
Tomsk, 634050, Russia. E-mail: polus@online.tomsk.net.

Andrey V. Plotnikov,

«Scientific and Production Center "Polus"», 56 «v», Kirov avenue,
Tomsk, 634050, Russia. E-mail: polus@online.tomsk.net.

Semen A. Vagaytsev,

«Scientific and Production Center "Polus"», 56 «v», Kirov avenue,
Tomsk, 634050, Russia. E-mail: polus@online.tomsk.net

Evgeny A. Budnikov,

Limited Liability Company «PROMHOL», 20, Lyalin pereulok,
Moscow, 105062, Russia. E-mail: promhol@mail.ru

The work is caused by necessity to automate thermovacuum and climatic tests of space industry products at space branch enterprise.

The aim of the work is the development of specialized system for monitoring modes and parameters of thermovacuum and climatic tests processes; substantiation of the chosen technologies of collecting, processing of primary data on pre-defined and current profile test as well as the data on technical condition of the applied test equipment and measuring instruments with various modifications.

Methods and technologies: analysis of technological processes of product tests, design and layout of the automated control and monitoring systems, test running in a program mode, analysis of pre-defined test profile, parallel programming, object-oriented programming, object linking and embedding (OLE), storage and processing of XML-data, human-machine interaction, client-server interaction.

Results: The authors have developed the automated control system for test parameters and its basic software and engineering solutions and components; have assessed the effectiveness of their application by the first results of operation.

Conclusions: The use of the automated system consisting of up to several tens of test equipment units, networked and controlled by a single center by thermovacuum and climatic test processes, provides economic effectiveness, high technical-organizational and metrology level for on-ground test of space industry products.

Key words:

Climatic and thermovacuum tests, external influencing factors, liquid heating and cooling system, indirect liquid heating and cooling, cascade refrigeration system, supervisory control and data acquisition, human-machine interface, controller-regulator, programmable logic controller, automation of test operation, automated system of test parameters control.

REFERENCES

1. Yanovich S.V., Litvak Yu.N., Mikhalev P.A., Bashkov V.M. Razrabotka stenda dlya issledovaniya stoykosti materialov k vozdeystviyu faktorov kosmicheskogo prostranstva [Developing stand for researching material resistance to the effects of space factors]. *Inzhenerny vestnik – Engineering bulletin*, 2012, no. 11. Available at: http://engbul.bmstu.ru/file/511839.html?_s=1 (accessed 20 January 2014).
2. Kravchenko S.V., Nesterov S.B., Romanko V.A., Testodov N.A., Khalimanovich V.I., Khristich V.V. Kompleksnye sistemy dlya otrabotki i ispytaniya kosmicheskikh apparatov [Complex systems for processing and testing space vehicles]. *Vakuumnaya tekhnika i tekhnologiya – Vacuum equipment and technology*, 2012, no. 4, pp. 213–230.
3. Rassamakin B.M., Khayrnasov S.M., Khominich V.I., Budenny A.V., Kovalenko E.Yu., Eliseeva E.N. Termovakuumnye ispytaniya sistemy elektrosnabzheniya nanosputnika NTUU «KPI» [Thermovacuum testing of power supply system of nanosatellite NTUU «KPI»]. *Tekhnologiya i konstruirovaniye v elektronnoy apparature – Technology and design in electronic equipment*, 2011, no. 5, pp. 6–10.
4. Reznik S.V., Denisov O.V., Prosuntsov P.V., Timoshenko V.P., Shulyakovskiy A.V. Termovakuumnye ispytaniya polykh kompozitnykh sterzhney dlya kosmicheskikh konstruktsey [Thermovacuum tests of hollow composite rods for space structures]. *Vse materialy. Entsiklopedicheskiy spravochnik – All the materials. Encyclopedic Reference*, 2012, no. 7, pp. 8–12.
5. Bobyr M., Titov V., Emelyanov S. *Teoreticheskie osnovy postroeniya avtomatizirovannykh sistem upravleniya tekhnologicheskimi protsessami na osnove nechetkoy logiki* [Theoretical foundations of constructing automated control systems based on fuzzy logic]. Staryy Oskol, TNT Publ., 2011. 232 p.
6. Ivanov A. *Proektirovaniye avtomatizirovannykh sistem manipulirovaniya obektami obrabotki i sborki* [Design of automated systems of handling processing and assembly objects]. Moscow, Forum Publ., 2012. 352 p.
7. Koptsyukh E.S. Osnovnye zadachi etapa obsledovaniya v rambakh proektov avtomatizatsii [Main objectives of pre-implementation screening phase]. *Prikladnaya informatika – Applied informatics*, 2010, no. 6, pp. 10–17.
8. Nikonov A.V. Vnedreniye sredstv avtomatizatsii v protsessy kontrolya i upravleniya na proizvodstve [Application of automation in monitoring and controlling in industries]. *Tekhnicheskie nauki – ot teorii k praktike – Technical sciences – from theory to practice*, 2012, no. 7–1, pp. 127–131.
9. Egorov G.A., Belonogov A.D., Ostrovskiy M.A., Reyzman Ya.A. Realizatsiya cheloveko-mashinnogo interfeysa v integrirovannoy tekhnologii proektirovaniya avtomatizirovannykh sistem kontro-

- lya i upravleniya [Implementation of human-machine interface in integrated design technology of automated monitoring and control systems]. *Mekhatronika, avtomatizatsiya, upravlenie – Mechatronics, automation, control*, 2011, no. 7, pp. 56–62.
10. Kozletsov A.P., Reshetnikov I.S. Sovremennye sposoby organizatsii obmena dannymi s sistemami upravleniya [Modern methods of data exchange with the control systems]. *Informatsionnye tekhnologii v proektirovanii i proizvodstve – Information technologies in designing and manufacturing*, 2010, no. 2, pp. 17–23.
 11. Veyber V.V., Kudinov A.V., Markov N.G. Zadacha sbora i pere-dachi tekhnologicheskoy informatsii raspredelenno go promyshlennogo predpriyatiya [The task of collecting and transferring technological information of the distributed industrial enterprise]. *Bulletin of the Tomsk Polytechnic University*, 2011, vol. 319, no. 5, pp. 69–74.
 12. Reyngen Yu.E. Avtomatizirovannyye sistemy upravleniya tekhnologicheskimi protsessami. SCADA sistema [Automated control systems of technological processes. SCADA system]. *Nauchnyy vestnik Moskovskogo gosudarstvennogo gornogo universiteta – Scientific bulletin of the Moscow State Mountain University*, 2012, no. 12, pp. 114–132.
 13. SCADA. *Wikipedia. The free encyclopedia*. 2013. Available at: <http://ru.wikipedia.org/wiki/SCADA> (accessed 20 January 2014).
 14. Cherpakov B.A., Strizhevskiy M.N. Mnogokanalnaya avtomatizirovannaya sistema dlya izmereniya temperatury pri provedenii teploprochnostnykh ispytaniy [Multichannel automated system for measuring temperature when testing heat resistance]. *Vestnik Moskovskogo Aviatcionnogo Instituta – Bulletin of the Moscow Aviation University*, 2010, vol. 17, no. 4, pp. 27–31.
 15. Lubentsov V.F., Lubentsova E.V., Masyutina G.V., Rudakova T.A. Sintez intervalnoy sistemy upravleniya s adaptivnym PID-regulyatorom [Synthesis of interval control system with adaptive PID controller]. *Industrial Automation*, 2013, no. 9, pp. 62–64.
 16. Baskur O., Tyunyatkin A.V., Khertler K. Diagnostika i monitoring oborudovaniya po sostoyaniyu. Monitoring sostoyaniya oborudovaniya v realnom vremeni [Diagnosis and equipment monitoring on a condition. Monitoring of equipment condition in real time]. *Industrial Automation*, 2012, no. 9, pp. 26–29.
 17. Rukovodstvo po postroeniyu sistem upravleniya na baze produkt-sii Delta Electronics (seti Modbus, CANOpen, Ethernet) [Guide to building management systems based on products of Delta Electronics (network Modbus, CANOpen, Ethernet)]. *NPO "Stoik Ltd"*, 2013. Available at: http://www.stoikltd.ru/images/manual/Training_kit_OP_Manual.pdf (accessed 20 January 2014).
 18. Termodat-16E5. Rukovodstvo polzovatela [User manual]. *NPP «Sistemy kontrolya»*, 2013. Available at: <http://termo-dat.ru/pics/16E5.pdf> (accessed 20 January 2014).
 19. IFD-6500. Instruktsiya po primeneniyu [Application instruction]. *NPO "Stoik Ltd"*, 2011. Available at: http://www.stoikltd.ru/images/manual/commodule/IFD6500_manual_rus.pdf (accessed 20 January 2014).
 20. Instruktsiya po sopryazheniyu kontrollerov Delta DVP so SCADA sistemami [The instruction on mating control units Delta DVP with SCADA-systems]. *NPO "Stoik Ltd"*, 2013. Available at: http://www.stoikltd.ru/images/manual/DVP_and_SCADA.pdf (accessed 20 January 2014).
 21. Kommunikatsionny modul RTU-485 dlya udalennogo sbora danykh ot moduley rasshireniya kontrollerov Delta DVP serii S po protokolu Modbus [Communication Module RTU-485 for remote data collection from expansion modules of controllers Delta DVP series S by Modbus protocol]. *NPO "Stoik Ltd"*, 2013. Available at: http://www.stoikltd.ru/images/manual/modules/RTU_485_manual_rus.pdf (accessed 20 January 2014).
 22. Ispolzovanie rezhima Virtual Com (virtualny COM-port) v kommunikatsionnom module IFD9506 [Using the Virtual Com mode (virtual COM-port) on the communication module IFD9506]. *NPO "Stoik Ltd"*, 2013. Available at: http://www.stoikltd.ru/images/manual/commodule/IFD9506_Virtual_com_rus.pdf (accessed 20 January 2014).
 23. Grinev M., Shcheklein I. Orientirovannyye na prilozheniya metody khraneniya XML-dannykh [The methods of storage the XML-data oriented on applications]. *Trudy ISP RAN – Proc. ISP RAN*, 2009, vol. 16, pp. 143–154.
 24. Nemnyugin S.A., Stesik O.L. *Parallelnoe programmirovaniye dlya mnogoprotssessornykh vychislitelnykh sistem* [Parallel programming for multiprocessor computing systems]. Saint Petersburg, BHV-Petersburg Publ., 2002. 400 p.
 25. Shishmarev V. *Diagnostika i nadezhnost avtomatizirovannykh sistem* [Diagnosis and reliability of automated systems]. Moscow, Academy Publ., 2013. 352 p.

УДК 004.492.2

ОЦЕНКА РИСКОВ ФИЗИЧЕСКОЙ БЕЗОПАСНОСТИ ЛИНЕЙНОЙ ЧАСТИ МАГИСТРАЛЬНОГО НЕФТЕПРОВОДА

Кукало Иван Анатольевич,

аспирант кафедры радиоэлектроники и защиты информации
ФГБОУ ВПО «Томский государственный университет систем управления и
радиоэлектроники», Россия, 634050, г. Томск, пр. Ленина, 40.
E-mail: i@kukalo.ru

Гривцов Сергей Николаевич,

ведущий специалист отдела охраны магистральных нефтепроводов
Службы безопасности ОАО «Центрсибнефтепровод»,
Россия, 634050, г. Томск, ул. Набережная р. Ушайки, д. 24.
E-mail: GrivtsovSN@tom.transneft.ru

Актуальность работы обусловлена необходимостью защиты объектов топливно-энергетического комплекса от совершения актов незаконного вмешательства.

Цель работы: обобщение и развитие методики оценки рисков в иерархических структурах критически важных объектов для пространственно-распределенных линейных объектов топливно-энергетического комплекса типа линейная часть магистрального нефтепровода.

Методы исследования: Математическое моделирование распределения вероятности рисков события, экспертная оценка методом Дельфи, методом ранжирования, метод непосредственной оценки, численное интегрирование двумерной функции риска во времени и пространстве, интерполяция функции траектории проведенного мероприятия.

Результаты: Предложена методика оценки риска физической безопасности для линейной части магистрального нефтепровода, реализующая концепцию ГОСТ Р ИСО 31000–2010. Проведена идентификация рисков физической безопасности линейной части магистрального нефтепровода, разработана математическая модель оценки рисков для пространственно-распределенных линейных объектов топливно-энергетического комплекса типа линейная часть магистрального нефтепровода для проведения анализа и сравнительной оценки риска. Модель позволяет рассчитать пространственно-временное распределение величины риска физической безопасности, которое основывается на предположении о двумерном дискретном случайном распределении вероятности рисков события. Результирующее значение величины риска физической безопасности определяется на основе данных о проведенных мероприятиях по контролю защищенности линейного объекта. Произведен расчет выходных параметров модели для участков магистрального нефтепровода «Александровское–Анжеро-Судженск» с учетом входных параметров, предоставленных ОАО «Центрсибнефтепровод».

Ключевые слова:

Оценка рисков, магистральный нефтепровод, физическая безопасность, матрица распределения риска, пространственно-временное распределение, накопленный риск, уровень риска.

Основные меры по обеспечению безопасности линейных объектов топливно-энергетического комплекса (ТЭК) в Российской Федерации устанавливаются Федеральным законом [1]. Известные подходы к управлению безопасностью систем масштаба ТЭК в значительной мере базируются на процедуре категорирования опасных объектов, в результате которой объектам одной категории предъявляются единые требования по обеспечению безопасности [2]. Альтернативой процедуре категорирования, как известно, является подход по управлению рисками систем физической безопасности (УР-СФБ). Создание методик, реализующих данный подход, сопряжено с необходимостью решения нескольких взаимосвязанных задач, включающих установление области применения, оценку риска и обработку риска [3]. Пример формальной реализации такого подхода для иерархических структур критически важных пространственно-локализованных объектов дан в работе [4]. В работах [5–7] предложены аналогичные концепции управления рисками различных систем –

информационных, кибернетических, транспортных, таможенных, банковских и др.

К сожалению, указанные методики управления рисками не могут охватить всего многообразия действующих систем физической безопасности (ФБ). Одним из факторов, ограничивающих метод оценки рисков [4], является допущение о точном характере объекта УР-СФБ. Это допущение исключает из рассмотрения множество пространственно-распределенных систем. К таковым, в частности, относится линейная часть магистрального нефтепровода (ЛЧ МН), характеризующаяся пространственно-протяженной и многосвязной инфраструктурой [8]. Особенности ЛЧ МН осложняют задачу обеспечения безопасности ТЭК, не позволяя использовать для этой цели стандартные методики УР-СФБ, ориентированные на точечные объекты.

Целью настоящей работы является обобщение и развитие методики оценки рисков в иерархических структурах критически важных объектов [4] для пространственно-распределенных линейных

объектов ТЭК типа ЛЧ МН, учитывая зарубежные [9–11] и отечественные [12, 13] разработки по управлению эксплуатационными рисками МН.

Процедура оценки риска, как известно, является ключевым этапом управления рисками и представляет собой процесс, объединяющий идентификацию, анализ и сравнительную оценку риска [14]. Данная процедура основывается на моделях злоумышленника и угроз СФБ, которые для объектов ЛЧ МН были описаны в работе [15]. Таким образом, разрабатываемая методика является основой для создания специализированной модели управления рисками линейных объектов ТЭК, которая позволит реализовать заключительный этап процесса управления рисками – обработку риска ФБ [3].

На первом этапе оценки риска системы ЛЧ МН рассмотрим перечень источников риска и событий. Считаем, что множество источников рисков событий для ЛЧ МН совпадает с множеством угроз [15].

Возможный перечень источников рисков событий для типового линейного объекта – ЛЧ МН, представлен в табл. 1.

Таблица 1. Перечень источников рисков событий

$n_{i,j}$	Источник рисков событий	Описание источника
n_1	Группа преступных элементов	Лицо или группа лиц, действующие согласованно с целью совершения акта незаконного вмешательства для получения выгоды
n_2	Внутренние нарушители	Лицо или группа лиц, обладающих правом доступа на объект в силу выполнения служебных или иных обязанностей
n_3	Диверсионно-террористическая группа	Подразделение, используемое для совершения диверсионно-террористических актов уничтожения или временного выведения из строя важнейших объектов топливно-энергетического комплекса

Анализ известных угроз [15] и рисков событий позволил выявить множество возможных рисков событий для линейного объекта S как $r_{i,j}$. Отметим, что для ЛЧ МН каждому риску событию соответствует одна угроза безопасности.

Рассмотрим возможный вариант задания элементов множества рисков событий, представленный в табл. 2.

Перейдем далее к анализу риска, включающего оценку вероятности и последствий идентифицированных опасных событий с учетом наличия и эффективности применяемых способов управления [14]. Для учета особенностей пространственно-протяженных линейных объектов, руководствуясь общими принципами моделирования [16], опишем концептуальную модель оценки риска. С этой целью выделим основные понятия:

- акт незаконного вмешательства – противоправное действие (бездействие), угрожающее безопасному функционированию линейного объек-

та, повлекшее за собой повреждение или уничтожение имущества либо создавшее угрозу наступления таких последствий [1];

- риск ФБ – количественная оценка сочетания вероятности и последствий совершения акта незаконного вмешательства [3];
- линейный объект – линия электропередачи, линия связи (в том числе линейно-кабельные сооружения), трубопровод, автомобильная дорога, железнодорожная линия и другие подобные сооружения [17];
- контрольные мероприятия – мероприятия по контролю защищенности линейного объекта, направленные на выявление совершенных актов незаконного вмешательства.

Таблица 2. Перечень рисков событий системы S

$r_{i,j}$	Наименование рисков событий	Источники рисков событий	Описание угрозы
r_1	Хищение нефти с использованием действующей врезки в ЛЧ МН либо регулярный отбор нефти через вдоль-трассовое оборудование	n_1, n_2	Установка патрубка с задвижкой с последующим нарушением целостности трубопровода для несанкционированной кражи нефти, взлом вантузов, камер приема-пуска средств очистки и диагностики, запорной арматуры и др. технологических устройств с целью несанкционированной кражи нефти
r_2	Кража вдоль-трассового оборудования	n_1, n_2	Кража вдоль-трассового оборудования и его компонентов (высоковольтные линии, компоненты из цветных металлов и др.)
r_3	Установка взрывчатого устройства на ЛЧ МН	n_3	Реализация террористического акта с целью нанесения материального, экологического и другого ущерба объекту топливно-энергетического комплекса

Далее, следуя [16], в качестве входных параметров математической модели принимаем параметры линейного объекта S , интенсивность инцидентов нарушения ФБ линейного объекта, характеристики проводимых контрольных мероприятий.

Важной характеристикой рассматриваемой модели является ее динамический характер, учитывающий суточные и сезонные колебания риска вдоль различных участков ЛЧ МН.

Расчет этой динамики основывается на априорном показателе интенсивности инцидентов нарушения ФБ и степени опасности для того или иного участка линейного объекта. Модель должна также учитывать информацию о мероприятиях по контролю защищенности линейного объекта, снижающих величину риска ФБ в местах их проведения на величину их рископонижающего потенциала [4]. Выходным параметром модели управления рисками S является динамика пространственно-временного распределения величины риска ФБ ЛЧ МН. Аналогичные подходы по использованию про-

странственно-временного распределения вероятности событий применяются в информационной безопасности [18], экологии [19], криминалистике [20] и др.

Каждое рисковое событие имеет три основные характеристики: c_{i^R} – цену риска (оценку ущерба, который может быть нанесен событием риска), p_{i^R} – вероятность рискового события и величину риска w_{i^R} , рассчитываемую по формуле [4]

$$w_{i^R} = c_{i^R} \cdot p_{i^R}.$$

На основе входного параметра модели λ – периода повторяемости инцидентов нарушения физической безопасности линейного объекта – зададим границы области определения функции p_{i^R} во времени $t \in \Delta t \in (t_{\text{нач}}, t_{\text{кон}})$. Оценку величины риска для каждого события определим как функцию

$$w_{i^R}(t) = c_{i^R} \cdot p_{i^R}(t), \quad (1)$$

где $w_{i^R}(t)$ – значение величины риска ФБ в момент времени t ; c_{i^R} – цена риска, т. е. оценка ущерба, который может быть нанесен событием риска; $p_{i^R}(t)$ – значение вероятности реализации рискового события в момент времени t .

Определим линейный объект как отрезок числовой прямой $S(x_{\text{нач}}, x_{\text{кон}})$, где $x_{\text{нач}}$ – начальная граничная точка отрезка, $x_{\text{кон}}$ – конечная граничная точка отрезка. Зададим границы области определения функции p_{i^R} вдоль оси линейного объекта как $x \in \Delta x \in (x_{\text{нач}}, x_{\text{кон}})$. Следуя подходу авторов [12], добавим второй аргумент в формулу (1) и определим функцию $w_{i^R}(x, t)$

$$w_{i^R}(x, t) = c_{i^R} \cdot p_{i^R}(x, t).$$

Рассмотрим рисковое событие как двумерную дискретную случайную величину, определенную на равных интервалах длиной h_t и h_x . Тогда интегральное значение риска на интервале $\Delta t \times \Delta x$ будет определяться уравнением

$$\begin{aligned} w_{i^R}(\Delta x, \Delta t) &= c_{i^R} p_{i^R}(x, t) = \\ &= \sum_{j=1}^n \sum_{i=1}^m w_{i^R}(\bar{x}_i, \bar{t}_j) = c_{i^R} \sum_{j=1}^n \sum_{i=1}^m p_{i^R}(\bar{x}_i, \bar{t}_j), \end{aligned}$$

где $h_x = \frac{x_{\text{кон}} - x_{\text{нач}}}{m}$, $h_t = \frac{t_{\text{кон}} - t_{\text{нач}}}{n}$ – шаги сетки; n –

количество шагов сетки для Δt ; m – количество шагов сетки для Δx ; $j \in J^M$ – множество индексов, определенных на интервале Δt ; $i \in I^M$ – множество индексов, определенных на интервале Δx ; $\bar{t}_j \in ((j-1)h_t, jh_t)$, $\bar{x}_i \in ((i-1)h_x, ih_x)$ – интервалы; $w_{i^R}(\bar{x}_i, \bar{t}_j)$ – значение величины риска на интервале $\bar{t}_j, \bar{x}_i$.

Значения $w_{i^R}(\bar{x}_i, \bar{t}_j)$ могут быть записаны в виде матрицы $\|w_{i^R}(\bar{x}_i, \bar{t}_j)\|$ (табл. 3).

В качестве допущения модели, значение цены риска линейного объекта считаем постоянной величиной, не зависящей от времени и ее конкретного расположения на линейном объекте. Для линейных объектов транспортной инфраструктуры цену риска c_{i^R} (величину ущерба) можно определить как обобщающий показатель некоторых частных зна-

чений ущерба физической безопасности ЛЧ МН для рискового события [2] по следующей формуле

$$c_{i^R} = c_{i^R_{\text{пл}}} + c_{i^R_{\text{гр}}} + c_{i^R_{\text{бал}}} + c_{i^R_{\text{экол}}},$$

где $c_{i^R_{\text{пл}}}$ – финансовый ущерб, определяемый численностью погибших и пострадавших; $c_{i^R_{\text{гр}}}$ – финансовый ущерб от кражи транспортируемого сырья и уменьшения грузопотока; $c_{i^R_{\text{бал}}}$ – балансовая стоимость сооружения или стоимость восстановления; $c_{i^R_{\text{экол}}}$ – стоимостное выражение ожидаемого экологического ущерба.

Таблица 3. Матрица распределения рисков $\|w_{i^R}(\bar{x}_i, \bar{t}_j)\|$

Δt	Δx				
	$\bar{x}_1$	...	$\bar{x}_i$	...	$\bar{x}_m$
$\bar{t}_1$	$w_{i^R}(\bar{x}_1, \bar{t}_1)$	...	$w_{i^R}(\bar{x}_i, \bar{t}_1)$	...	$w_{i^R}(\bar{x}_m, \bar{t}_1)$
...	...	...	...	...	...
$\bar{t}_j$	$w_{i^R}(\bar{x}_1, \bar{t}_j)$	...	$w_{i^R}(\bar{x}_i, \bar{t}_j)$	...	$w_{i^R}(\bar{x}_m, \bar{t}_j)$
...	...	...	...	...	...
$\bar{t}_n$	$w_{i^R}(\bar{x}_1, \bar{t}_n)$	...	$w_{i^R}(\bar{x}_i, \bar{t}_n)$	...	$w_{i^R}(\bar{x}_m, \bar{t}_n)$

Определим матрицу распределения величины $\|p_{i^R}(\bar{x}_i, \bar{t}_j)\|$ (табл. 4), на основе которой можно вычислить $\|w_{i^R}(\bar{x}_i, \bar{t}_j)\|$

$$\|w_{i^R}(\bar{x}_i, \bar{t}_j)\| = c_{i^R} \|p_{i^R}(\bar{x}_i, \bar{t}_j)\|.$$

Таблица 4. Матрица распределения $\|p_{i^R}(\bar{x}_i, \bar{t}_j)\|$

Δt	Δx				
	$\bar{x}_1$	...	$\bar{x}_i$	...	$\bar{x}_m$
$\bar{t}_1$	$p_{i^R}(\bar{x}_1, \bar{t}_1)$	...	$p_{i^R}(\bar{x}_i, \bar{t}_1)$	...	$p_{i^R}(\bar{x}_m, \bar{t}_1)$
...	...	...	...	...	...
$\bar{t}_j$	$p_{i^R}(\bar{x}_1, \bar{t}_j)$	...	$p_{i^R}(\bar{x}_i, \bar{t}_j)$	...	$p_{i^R}(\bar{x}_m, \bar{t}_j)$
...	...	...	...	...	...
$\bar{t}_n$	$p_{i^R}(\bar{x}_1, \bar{t}_n)$	...	$p_{i^R}(\bar{x}_i, \bar{t}_n)$	...	$p_{i^R}(\bar{x}_m, \bar{t}_n)$

Примем, что события $p_{i^R}(\bar{t}_j)$ и $p_{i^R}(\bar{x}_i)$ независимы, тогда

$$p_{i^R}(\bar{x}_i, \bar{t}_j) = p_{i^R}(\bar{t}_j) \cdot p_{i^R}(\bar{x}_i).$$

Указанное допущение свидетельствует о том, что рисковое событие может произойти в любой момент времени $\bar{t}_j$ на произвольном участке линейного объекта $\bar{x}_i$ и корреляция между функциями вероятности $p_{i^R}(\bar{t}_j)$ и $p_{i^R}(\bar{x}_i)$ отсутствует. Соответственно, для вероятности справедливы формулы

$$p_{i^R}(\bar{t}_j) = \sum_{i=1}^m p_{i^R}(\bar{x}_i, \bar{t}_j), \quad p_{i^R}(\bar{x}_i) = \sum_{j=1}^n p_{i^R}(\bar{x}_i, \bar{t}_j).$$

По условию нормировки

$$\sum_{j=1}^n \sum_{i=1}^m p_{i^R}(\bar{x}_i, \bar{t}_j) = \sum_{j=1}^n p_{i^R}(\bar{t}_j) = \sum_{i=1}^m p_{i^R}(\bar{x}_i) = 1. \quad (2)$$

Пусть в качестве допущения модели величина t имеет равномерное дискретное распределение. Тогда значение вероятности составляющей $p_i(\bar{x}_i)$ системы случайных величин $p_i(\bar{x}_i, \bar{t}_j)$ будет определяться уравнением

$$p_{i^R}(\bar{t}_j) = \frac{1}{n}.$$

Вероятность составляющей $p_i(\bar{x}_i)$ системы случайных величин $p_i(\bar{x}_i, \bar{t}_j)$ определим экспертным методом [15]. Для всех участков $\bar{x}_i \in (\bar{x}_1, \dots, \bar{x}_m)$ определим среднее значение вероятности реализации рискованного события $p_{i^R}(\bar{x}_{cp})$

$$p_{i^R}(\bar{x}_{cp}) = \frac{1}{m}.$$

Для того чтобы учесть факторы, влияющие на вероятность реализации угрозы на том или ином участке линейного объекта, введем нормирующие коэффициенты ξ_i , определяющие степень опасности конкретного участка ЛЧ МН

$$p_{i^R}(\bar{x}_i) = \xi_i \cdot p_{i^R}(\bar{x}_{cp}). \quad (3)$$

Для определения коэффициентов ξ_i в формуле (3) для каждого участка $\bar{x}_i \in (\bar{x}_1, \dots, \bar{x}_m)$ линейного объекта воспользуемся методом балльной оценки факторов [13, 21]. На первоначальном этапе определим группы факторов влияния $Гр_a$ с весовыми коэффициентами ρ_a , определяющими вероятность их влияния на реализацию угрозы ФБ. В пределах каждой группы $Гр_a$ определим факторы влияния в количестве B_a . Каждый фактор имеет буквенно-цифровое обозначение F_{ab} , где a – номер группы, b – номер фактора в группе. Относительный вклад фактора F_{ab} внутри своей группы учитывается с помощью весового коэффициента (доли) q_{ab} .

Расчет коэффициента ξ_i производился нами с использованием балльной оценочной системы, при которой каждому значению фактора F_{ab} ставится в соответствие определенное, назначаемое на основании расчета или экспертной оценки, количество баллов B_{ab} (по 10-балльной шкале), отражающее интенсивность его влияния. При рассмотрении конкретного i -го участка линейного объекта последовательно оценивается степень влияния каждого из факторов. Полученные для всех факторов влияния балльные оценки $\{B_{ab}, a \in (1 \dots A), b \in (1 \dots B)\}$ подставляются в формулу

$$\xi_i = \frac{B_i}{B_{cp}},$$

где B_i – балльная оценка i -го участка линейного объекта, определяемая как:

$$B_i = \sum_{a=1}^A \sum_{b=1}^{B(a)} \rho_a \cdot q_{ab} \cdot B_{ab},$$

где B_{cp} – средняя балльная оценка линейного объекта (среднее арифметическое), полученная на основе балльной оценки каждого участка линейного объекта $i \in (1 \dots m)$, а m – общее количество рассматриваемых участков линейного объекта.

В работе [15] приведены группы факторов, факторы и их веса для определения значения ξ_i типовой модели угроз ЛЧ МН. Факторы и веса для рассмотренных рискованых событий идентичны факторам модели угроз ЛЧ МН, поскольку каждому рискованному событию ставится в соответствие только одна угроза безопасности. Группы факторов и веса рискованых событий представлены в табл. 5.

Таблица 5. Группы факторов влияния

Обозначение и наименование группы факторов		Доля группы, ρ_a		
		r_1^s	r_2^s	r_3^s
$Гр_1$	Предпосылки акта незаконного вмешательства	0,20	0,20	0,80
$Гр_2$	Реализация акта незаконного вмешательства	0,40	0,40	0,10
$Гр_3$	Извлечения выгоды в результате акта незаконного вмешательства	0,40	0,40	0,10

Факторы влияния и их весовые коэффициенты представлены в табл. 6.

Таблица 6. Факторы влияния

Обозначение и наименование фактора влияния		Доля в группе q_{ab}		
		r_1^s	r_2^s	r_3^s
Группа 1: Предпосылки незаконного вмешательства				
F_{11}	Криминогенная обстановка	0,20	0,30	0,70
F_{12}	Населенность прилегающих территорий	0,60	0,40	0,20
F_{13}	Уровень занятости населения прилегающих территорий	0,20	0,30	0,10
Группа 2: Реализация незаконного вмешательства				
F_{21}	Защищенность объекта	0,10	0,10	0,60
F_{22}	Наличие подъездных путей	0,60	0,60	0,10
F_{23}	Наличие специального оборудования, персонала и транспорта для реализации акта незаконного вмешательства	0,30	0,30	0,30
Группа 3: Извлечения выгоды в результате незаконного вмешательства				
F_{31}	Наличие в прилегающих территориях пунктов сбыта похищенной продукции	0,90	0,80	0
F_{32}	Репутационные, экологические и политические последствия	0,10	0,20	1

В качестве примера в табл. 7 приведены значения отдельных составляющих фактора F_{31} и соответствующие им балльные оценки $B_{31}^{(m)}$, где m – номер составляющей. Итоговая балльная оценка для данного фактора рассчитывается как сумма балльных оценок нижеприведенных пяти составляющих. Если сумма баллов превышает 10, то $B_{31}=10$.

В работе [22] предложен способ хранения информации о GPS траекториях движущихся объектов в виде двумерного пространственно-временного распределения в системах обработки неструктурированных данных. Следуя данному подходу и переходя от абсолютных GPS координат к относительному позиционированию на ЛЧ МН, представим данные о мероприятиях по обеспечению безопасности объекта S в виде следующего системного множества

$$M = \{m_{i^M}(t), k_{i^M}\},$$

где $i^M \in I^M$ – множество индексов мероприятий; $k_{i^M} \in K^M$ – множество типов мероприятий; $m_{i^M}(t)$ – от-носительная координата мероприятия на ЛЧ МН.

Таблица 7. Балльная оценка фактора B_{31}

m	Наименование составляющей m	$B_{31}^{(m)}$
1	В районе пролегания участка нефтепровода находят-ся потребители, использующие нефть в качестве то-плива (малые котельные, теплицы, хлебозаводы, ас-фальтобитумные заводы и др.)	5
2	В районе пролегания участка нефтепровода находят-ся нелегальные нефтеперерабатывающие заводы	10
3	В районе пролегания участка нефтепровода находят-ся частные малые нефтеперерабатывающие заводы	7
4	В районе пролегания участка нефтепровода находят-ся крупные нефтеперерабатывающие заводы	2
5	В районе пролегания участка нефтепровода осу-ществляют предпринимательскую деятельность юри-дические лица, занимающиеся торговлей нефте-продуктами	3

Функция $m_{i^M}: t \rightarrow t_{i^M} \rightarrow x \rightarrow x_{i^M}$ является таблично заданной (табл. 8).

Таблица 8. Значения функции $m_{i^M}(t)$

t	$t_1^{i^M}$	$t_2^{i^M}$	...	...	$t_z^{i^M}$	...	...	...	$t_{i^M}^{i^M}$
$m_{i^M}(t)$	$x_1^{i^M}$	$x_2^{i^M}$	...	...	$x_z^{i^M}$	...	...	...	$x_{i^M}^{i^M}$

Функция $m_{i^M}(t)$ представляет собой траекторию мероприятия во времени и пространстве линейно-го объекта, а множество $z \in (1 \dots l_{i^M})$ представляет со-бой индексы координат мероприятия. В качестве допущения определим функцию $m_{i^M}(t)$ как непре-рывную.

Таким образом, каждому мероприятию будут соответствовать дискретные значения $p_{i^M}(\bar{x}_i, \bar{t}_j)$ в области определения $\Delta x \times \Delta t$ (табл. 2). Для вычи-сления промежуточных значений $(\bar{x}_i, \bar{t}_j)$ для каждо-го мероприятия $m_{i^M}(t)$ значение функции $m_{i^M}(t)$ ап-проксимируем в виде непрерывной кусочно-ли-нейной функции, построенной по значениям функ-ции в двух соседних точках

$$m_{i^M}(t) = \begin{cases} \frac{(t - t_1^{i^M})(x_2^{i^M} - x_1^{i^M})}{(t_2^{i^M} - t_1^{i^M})} + x_1^{i^M}, & t_1^{i^M} < t < t_2^{i^M} \\ \dots \\ \frac{(t - t_z^{i^M})(x_{z+1}^{i^M} - x_z^{i^M})}{(t_{z+1}^{i^M} - t_z^{i^M})} + x_z^{i^M}, & t_z^{i^M} < t < t_{z+1}^{i^M} \\ \dots \\ \frac{(t - t_{n_{i^M}-1}^{i^M})(x_{n_{i^M}}^{i^M} - x_{n_{i^M}-1}^{i^M})}{(t_{n_{i^M}}^{i^M} - t_{n_{i^M}-1}^{i^M})} + x_{n_{i^M}-1}^{i^M}, & t_{n_{i^M}-1}^{i^M} < t < t_{n_{i^M}}^{i^M} \end{cases}.$$

Тогда каждое контрольное мероприятие в каж-дой точке его проведения снижает вероятность ри-скового события до момента предыдущего меро-приятия на величину

$$p_{i^M}^{i^M}(\bar{x}_i, \bar{t}_j) = p_{i^M}(\bar{x}_i, \bar{t}_j) \cdot (1 - \sigma_{k_{i^M}}),$$

$$\bar{x}_i = x_z^{i^M}, \bar{t}_j \in (t_{z_{пред}}^{i^M}, t_z^{i^M}),$$

где $t_{z_{пред}}^{i^M}$ – момент времени проведения предыдуще-го мероприятия; $(t_{z_{пред}}^{i^M}, t_z^{i^M})$ – интервал времени с момента проведения предыдущего мероприятия; $\sigma_{k_{i^M}} \in (0, 1)$ – коэффициент эффективности сниже-ния рискового события для типа мероприятия k_{i^M} .

Коэффициент $\sigma_{k_{i^M}}$ определим с применением ме-тодов экспертного оценивания путем непосредствен-ного оценивания эффективности снижения рисково-го события для мероприятия типа k_{i^M} . Чем выше ко-эффициент $\sigma_{k_{i^M}}$, тем эффективнее мероприятие сни-жает величину риска на временном интервале до мо-мента проведения предыдущего мероприятия, по-скольку в результате контрольного мероприятия мы получаем информацию об отсутствии реализован-ных рисковых событий на обследуемом участке.

Общий рискоснижающий потенциал меро-приятия n_{i^M} можно вычислить по формуле

$$u_{i^M}^{i^M} = \sum p_{i^M}^{i^M}(\bar{x}_i, \bar{t}_j), \bar{x}_i = x_z^{i^M}, \bar{t}_j \in (t_{z_{пред}}^{i^M}, t_z^{i^M}).$$

С учетом основного свойства закона распреде-ления (2) значение вероятности рискового события перераспределится в «будущие» участки времени и на «соседние» участки линейного объекта

$$p_{i^M}^{i^M}(\bar{x}_i, \bar{t}_j) = \frac{u_{i^M}^{i^M}}{\left| \bar{t}_z^{i^M} \dots \bar{t}_j^{i^M} \right|}, \bar{x}_i = x_z^{i^M}, \bar{t}_j \in (t_z^{i^M}, t_n^{i^M}),$$

$$p_{i^M}^{i^M}(\bar{x}_i, \bar{t}_j) = \frac{p_{i^M}^{i^M}(\bar{x}_z^{i^M}, \bar{t}_z^{i^M})}{\left| x_1 \dots x_{z-1}^{i^M}, x_{z+1}^{i^M} \dots x_m^{i^M} \right|},$$

$$\bar{x}_i \in (x_1 \dots x_{z-1}^{i^M}, x_{z+1}^{i^M} \dots x_m^{i^M}), \bar{t}_j = \bar{t}_z^{i^M}.$$

Каждое мероприятие снижает отдельные эле-менты общей матрицы распределения $\|p_{i^M}(\bar{x}_i, \bar{t}_j)\|$ на величину $\|p_{i^M}^{i^M}(\bar{x}_i, \bar{t}_j)\|$ и повышает другие элементы матрицы $\|p_{i^M}(\bar{x}_i, \bar{t}_j)\|$ на величину $\|p_{i^M}^{i^M}(\bar{x}_i, \bar{t}_j)\|$. Для за-данной координаты $\bar{x}_z^{i^M}$, $\bar{t}_z^{i^M}$ мероприятия $m_{i^M}(t)$ может быть матрица корректировки значения ве-роятности $\|p_{i^M}^{i^M}(\bar{x}_i, \bar{t}_j)\|$ на всей области определения функции $p_{i^M}(\bar{x}_i, \bar{t}_j)$ (табл. 9).

Для оценки рискоснижающего потенциала множества мероприятий $M = \{m_{i^M}(t), k_{i^M}\}$ необходимо осуществить последовательный расчет значений $\|p_{i^M}^{i^M}(\bar{x}_i, \bar{t}_j)\|$ для каждого $\bar{x}_i = m_{i^M}(t) = x_z^{i^M}$, упорядо-

чивая множество значений $\bar{t}_z^{i^M}$ всех мероприятий по возрастанию, для того чтобы достоверно опреде-лить дату проведения предыдущего мероприятия.

Таблица 9. Матрица корректировки значения вероятности $\|p_{iR}^{iM}(\bar{x}_i, \bar{t}_j)\|$

Δt	Δx					
	$\bar{x}_1^{iM}$	$\bar{x}_2^{iM}$	...	$\bar{x}_i^{iM}$	...	$\bar{x}_{n_{iM}}^{iM}$
$\bar{t}_1$	0	0	...	0	...	0
...	...	...	...	...	...	...
$\bar{t}_j^{\text{пред}}$	0	0	...	$p_{iR}(\bar{x}_i^{iM}, \bar{t}_j^{\text{пред}})$	...	0
...	...	...	...	...	...	...
$\bar{t}_{n_{iM}}^{iM}$	$p_{iR}''(\bar{x}_1^{iM}, \bar{t}_{n_{iM}}^{iM})$	$p_{iR}''(\bar{x}_2^{iM}, \bar{t}_{n_{iM}}^{iM})$	...	$p_{iR}'(\bar{x}_i^{iM}, \bar{t}_{n_{iM}}^{iM})$	...	$p_{iR}''(\bar{x}_{n_{iM}}^{iM}, \bar{t}_{n_{iM}}^{iM})$
$\bar{t}_{n_{iM}+1}^{iM}$	0	0	...	$p_{iR}''(\bar{x}_i^{iM}, \bar{t}_{n_{iM}+1}^{iM})$	...	0
...	...	...	...	...	...	...
$\bar{t}_n$	0	0	0	$p_{iR}''(\bar{x}_i, \bar{t}_j)$	0	0

Если траектории нескольких мероприятий проходят через точку $(\bar{x}_z^{iM}, \bar{t}_z^{iM})$, то для мероприятия с наибольшим $\sigma_{k_{iM}}(\bar{t}_{z_{\text{пред}}}^{iM}, \bar{t}_z^{iM}) > 0$, для остальных мероприятий $\bar{t}_{z_{\text{пред}}}^{iM} = \bar{t}_z^{iM}$. Таким образом, при расчете рискообразующего потенциала учитывается наиболее эффективное мероприятие из всех проведенных.

Множество мероприятий M образуют совокупную матрицу корректирующих мероприятий $\|p_{iR}^M(\bar{x}_i, \bar{t}_j)\|$, на основе которой возможно определить результирующее значение риска матрицы распределения рисков $\|w_{iR}^{\text{рез}}(\bar{x}_i, \bar{t}_j)\|$ как набора арифметических операций с матрицами:

$$\begin{aligned} \|w_{iR}^{\text{рез}}(\bar{x}_i, \bar{t}_j)\| &= c_{iR} \cdot \|p_{iR}^{\text{рез}}(\bar{x}_i, \bar{t}_j)\| = \\ &= c_{iR} (\|p_{iR}^M(\bar{x}_i, \bar{t}_j)\| - \|p_{iR}^M(\bar{x}_i, \bar{t}_j)\|) = \\ &= c_{iR} \left(\|p_{iR}^M(\bar{x}_i, \bar{t}_j)\| - \sum_{i^M=1}^{i^M} \sum_{z=1}^{i^M} \|p_{iR}^M(\bar{x}_i, \bar{t}_j)\| \right). \end{aligned}$$

Результирующая матрица $\|w_{iR}^{\text{рез}}(\bar{x}_i, \bar{t}_j)\|$ определяет динамическое, пространственно-временное распределение риска ФБ. Матрица является источником формирования дополнительных интегральных количественных показателей, участвующих в процессе поддержки принятия решений по распределению имеющихся ресурсов для снижения риска.

В качестве примера такого показателя введем понятие накопленного риска $W_{iR}(\bar{x}_i, \bar{t}_j)$, описывающего уровень риска в монетарном выражении для участка линейного объекта $\bar{x}_i$ на момент времени $\bar{t}_j$ как

$$\begin{aligned} W_{iR}(\bar{x}_i, \bar{t}_j) &= w_{iR}^{\text{рез}}(\bar{x}_i, \Delta t') = \\ &= \sum_{j'=1}^j w_{iR}^{\text{рез}}(\bar{x}_i, \bar{t}_{j'}) = c_{iR} \sum_{j'=1}^j p_{iR}(\bar{x}_i, \bar{t}_{j'}), \end{aligned}$$

где $\Delta t' \in (t_{\text{нач}}, t_j)$ – интервал времени для расчета накопленного риска.

Введенный выше показатель $W_{iR}(\bar{x}_i, \bar{t}_j)$ основан на накопленной вероятности события риска, отражает финансовую составляющую риска и может применяться при обосновании затрат по проведению контрольных мероприятий, а также для оценки необходимого количества денежных средств для обеспечения безопасности линейных объектов [13].

Также введем понятие уровня риска $E_{iR}(\bar{x}_i, \bar{t}_j)$, выраженного в долях для участка линейного объекта $\bar{x}_i$ на момент времени $\bar{t}_j$ как:

$$\begin{aligned} E_{iR}(\bar{x}_i, \bar{t}_j) &= \frac{w_{iR}^{\text{рез}}(\bar{x}_i, \Delta t')}{w_{iR}(\bar{x}_i, \Delta t')} = \\ &= \frac{\sum_{j'=1}^j w_{iR}^{\text{рез}}(\bar{x}_i, \bar{t}_{j'})}{\sum_{j'=1}^j w_{iR}(\bar{x}_i, \bar{t}_{j'})} = \frac{\sum_{j'=1}^j p_{iR}^{\text{рез}}(\bar{x}_i, \bar{t}_{j'})}{\sum_{j'=1}^j p_{iR}(\bar{x}_i, \bar{t}_{j'})}, \end{aligned}$$

$E_{iR}(\bar{x}_i, \bar{t}_j)$ позволяет оценить динамику риска произвольного участка линейного объекта с учетом проведенных мероприятий. Показатель можно выразить в процентном отношении.

Значения $W_{iR}(\bar{x}_i, \bar{t}_j)$ и $E_{iR}(\bar{x}_i, \bar{t}_j)$, определенные на интервале $\Delta t \times \Delta x$, могут быть также определены в виде матриц $\|W_{iR}(\bar{x}_i, \bar{t}_j)\|$ и $\|E_{iR}(\bar{x}_i, \bar{t}_j)\|$.

Показатели $W_{iR}(\bar{x}_i, \bar{t}_j)$ и $E_{iR}(\bar{x}_i, \bar{t}_j)$, на наш взгляд, являются альтернативами действующим показателям эффективности процесса анализа и планирования мероприятий для поддержки заданного уровня защищенности линейного объекта, а именно:

- количество мероприятий на участке $\bar{x}_i$ на момент времени $\bar{t}_j$;
- количество дней с последнего мероприятия на участке $\bar{x}_i$ на момент времени $\bar{t}_j$.

Показатели $W_{iR}(\bar{x}_i, \bar{t}_j)$ и $E_{iR}(\bar{x}_i, \bar{t}_j)$ позволяют более эффективно решать заключительную задачу

процесса оценки рисков – сравнительную оценку риска, за счет следующих преимуществ:

- учитываются различные степени опасности для участков линейного объекта $\bar{x}_i$ со стороны злоумышленников;
- учитывается различная степень эффективности множества типов контрольных мероприятий $\{k_{pi}\}$.

В качестве примера использования модели рассмотрим результаты расчета показателей $W_{pi}(\bar{x}_i, \bar{t}_j)$ и $E_{pi}(\bar{x}_i, \bar{t}_j)$ для одного из МН Томской области. При этом будем использовать следующие входные параметры:

- λ – 1 инцидент/год, $t_{нач}$ – 01.01.2013, $t_{кон}$ – 31.12.2013;
 - S – МН «Александровское–Анжеро-Судженск» ($x_{нач}=0,0$; $x_{кон}=818,0$ км);
 - M – данные о мероприятиях, проведенных экипажем службы безопасности (СБ) ОАО «Центр-сибнефтепровод» за смену с 01.01.2013 20:07 по 02.01.2013 2:42 (табл. 10).
- В качестве внутренних параметров определим:
- $C_{pi}=30$ млн р.;
 - обследование на автомобиле $k=0,3$; пешее обследование $k=0,4$; обследование подъездных путей к МН $k=0,1$;
 - $h_i=1$ мин и $h_x=100$ м;
 - $\xi_i=1$.

Аппроксимированный график выходных параметров модели $W_{pi}(\bar{x}_i, \bar{t}_j)$ и $E_{pi}(\bar{x}_i, \bar{t}_j)$ для $\bar{x}_i \in (394, \dots, 440)$ км и $\bar{t}_j = 02.01.2013$ 2:42 представлен на рис. 1. На рисунке представлена зависимость влияния информации о проведенных контрольных мероприятиях на выходной параметр модели – величину риска ФБ.

На рис. 2, 3 представлены графики интегральных значений $W_{pi}(\Delta x', \bar{t}_j)$ и $E_{pi}(\Delta x', \bar{t}_j)$ для участка

$\Delta x'$ МН «Александровское–Анжеро-Судженск» длиной 400 км при $\Delta t=1$ год. Начальная и конечная координаты участка не указываются, поскольку реальные значения величин с привязкой к конкретному участку составляют коммерческую тайну ОАО «Центрсибнефтепровод».

На рис. 4, 5 представлены трехмерные графики распределения величин $W_{pi}(\bar{x}_i, \bar{t}_j)$ и $E_{pi}(\bar{x}_i, \bar{t}_j)$ для аналогичного участка $\Delta x'$ МН «Александровское–Анжеро-Судженск» длиной 400 км при $\bar{t}_j \in \Delta t$, $\Delta t=1$ год. Графики указывают на слабые и незащищенные места ЛЧ МН и обеспечивают лиц, осуществляющих принятие решений [10], наглядным представлением о защищенности ЛЧ МН во времени и пространстве.

Для построения указанных графиков использовался трехмерный графопостроитель «contourf» пакета Matlab, в котором топология поверхности массивов $W_{pi}(\bar{x}_i, \bar{t}_j)$ и $E_{pi}(\bar{x}_i, \bar{t}_j)$ в трехмерном пространстве изображалась в виде цветового набора двумерных линий постоянного уровня. При этом наглядную информацию об изменении функции дает заливка прямоугольника на плоскости ХУ (участок ЛЧ МН/время) цветом, зависящим от значения функции в точках плоскости (значение риска/уровня риска). Более светлые области графика плоскости соответствуют областям с наибольшим значением риска (уровня риска) и выделяют наиболее опасные места (во времени для участка ЛЧ МН), для которых вероятность появления рискованного события наиболее велика. Темные участки выделяют наименее опасные места, где за счет информации, полученной в ходе проведения контрольных мероприятий, вероятность появления рискованного события уменьшается.

Таблица 10. Мероприятия, проведенные экипажем СБ за смену

Время начала мероприятия, $t_1^{ин}$	Время окончания мероприятия, $t_2^{ин}$	Начальный участок осмотра, $x_1^{ин}$	Конечный участок осмотра, $x_2^{ин}$	Тип мероприятия, k_{pi}
01.01.2013 20:07	01.01.2013 20:37	394	398,4	Обследование на автомобиле
01.01.2013 20:37	01.01.2013 20:52	398,5	398,4	Пешее обследование
01.01.2013 20:52	01.01.2013 21:19	398,5	400,4	Обследование на автомобиле
01.01.2013 21:19	01.01.2013 21:53	400,2	400,4	Пешее обследование
01.01.2013 21:53	01.01.2013 22:11	408,9	400,4	Обследование на автомобиле
01.01.2013 22:11	01.01.2013 22:43	408,9	409	Пешее обследование
01.01.2013 22:43	01.01.2013 22:54	413,1	409	Обследование на автомобиле
01.01.2013 22:54	01.01.2013 23:19	413,1	413,2	Пешее обследование
01.01.2013 23:19	01.01.2013 23:42	416,4	413,2	Обследование на автомобиле
01.01.2013 23:42	01.01.2013 23:58	414,4	414,6	Пешее обследование
01.01.2013 23:58	02.01.2013 0:12	416,6	418,2	Обследование на автомобиле
02.01.2013 0:12	02.01.2013 0:27	418,3	418,2	Пешее обследование
02.01.2013 0:27	02.01.2013 0:58	418,3	427,9	Обследование на автомобиле
02.01.2013 0:58	02.01.2013 1:18	428	427,9	Пешее обследование
02.01.2013 1:18	02.01.2013 1:34	428	431	Обследование на автомобиле
02.01.2013 1:34	02.01.2013 1:43	431	431	Обследование подъездных путей к МН
02.01.2013 2:24	02.01.2013 2:35	435	435,2	Пешее обследование
02.01.2013 1:52	02.01.2013 2:03	436,2	436,2	Обследование подъездных путей к МН
02.01.2013 2:03	02.01.2013 2:24	435	437,4	Обследование на автомобиле
02.01.2013 2:35	02.01.2013 2:42	437,4	437,5	Пешее обследование

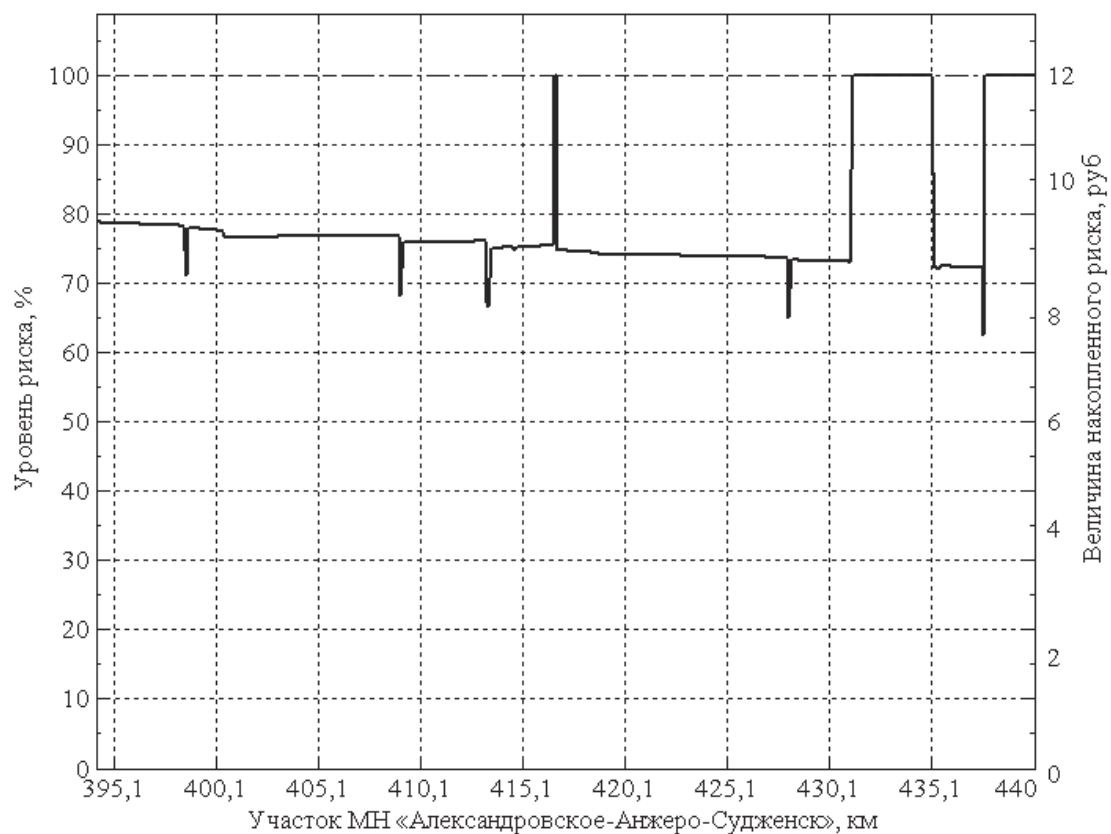


Рис. 1. Значение уровня риска и накопленного риска для участка МН 394–440 км

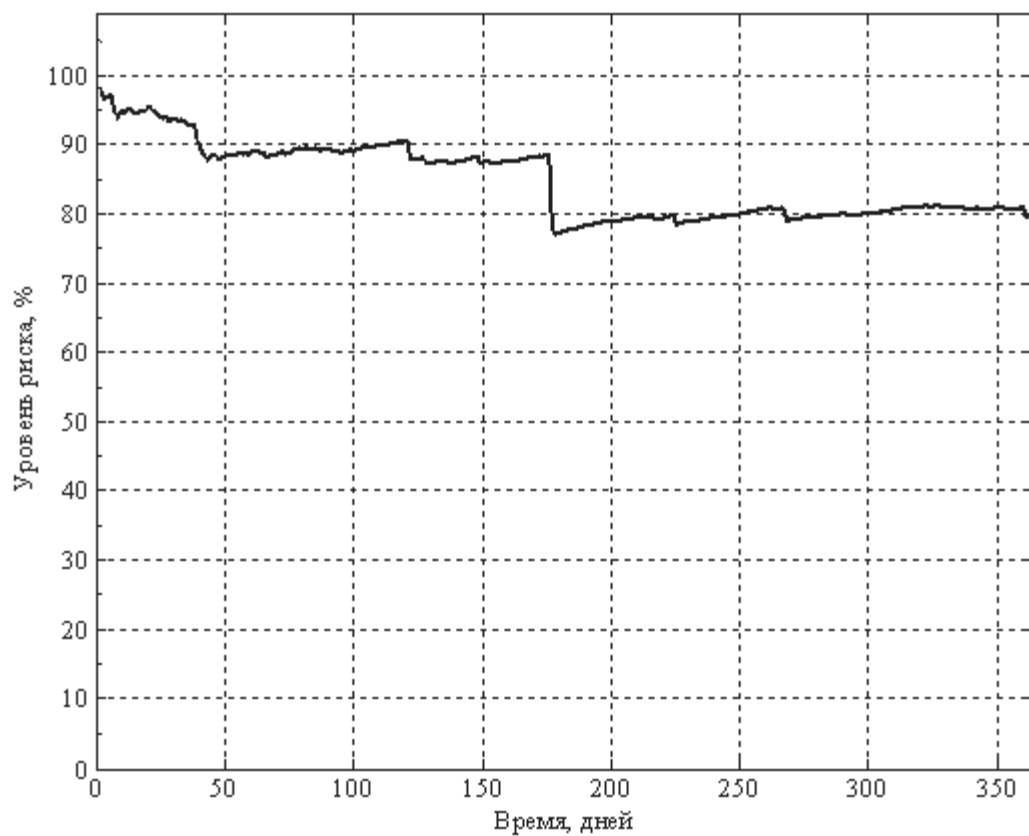


Рис. 2. Значение уровня риска на участке МН длиной 400 км в течение года

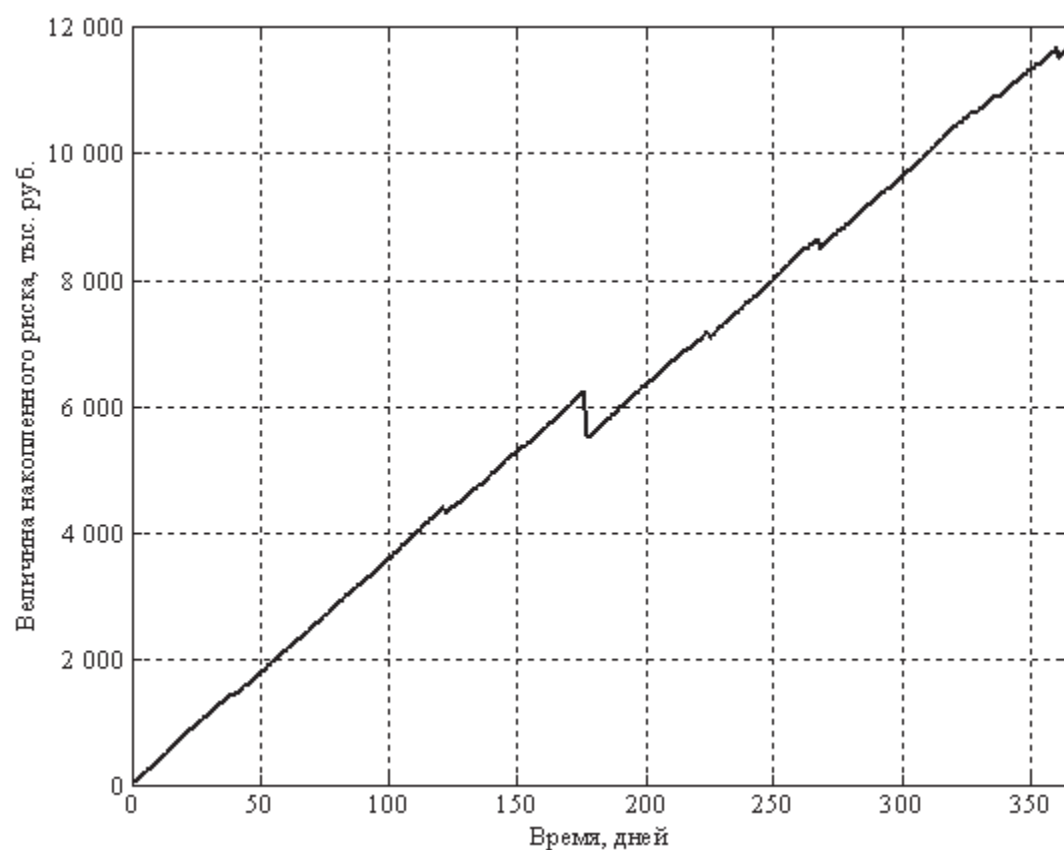


Рис. 3. Значение накопленного риска на участке МН длиной 400 км в течение года

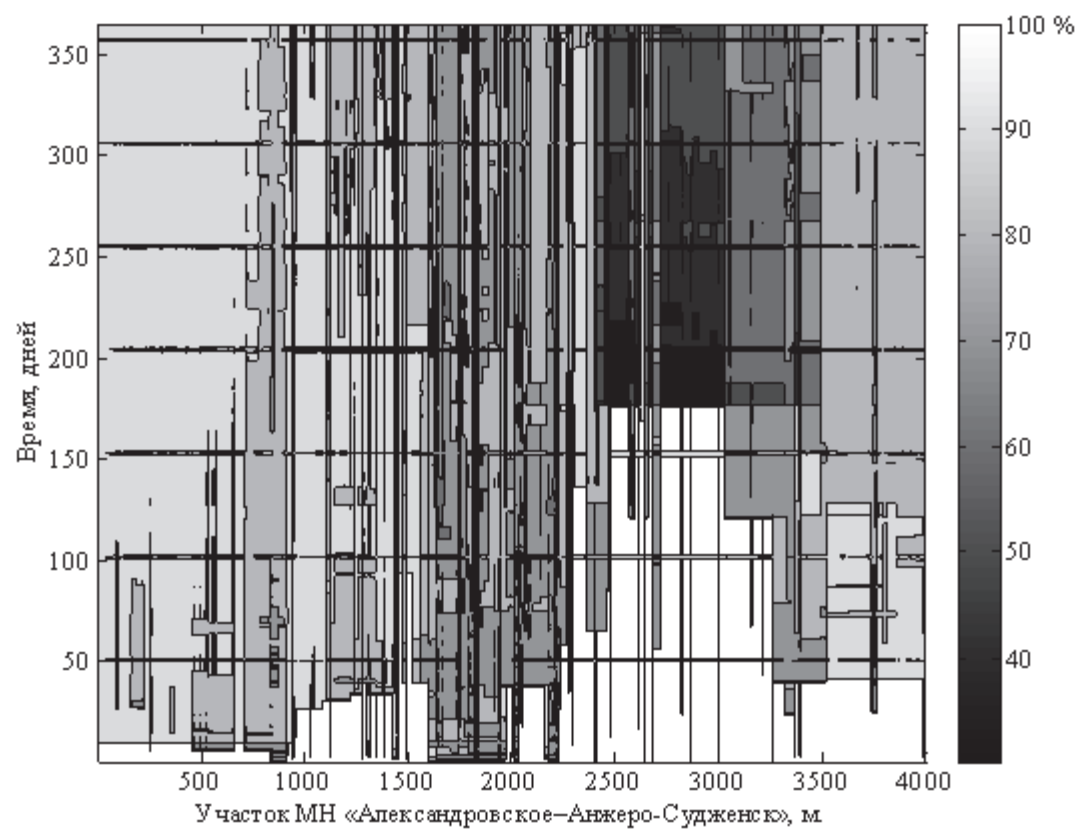


Рис. 4. Распределение уровня риска на участке МН длиной 400 км

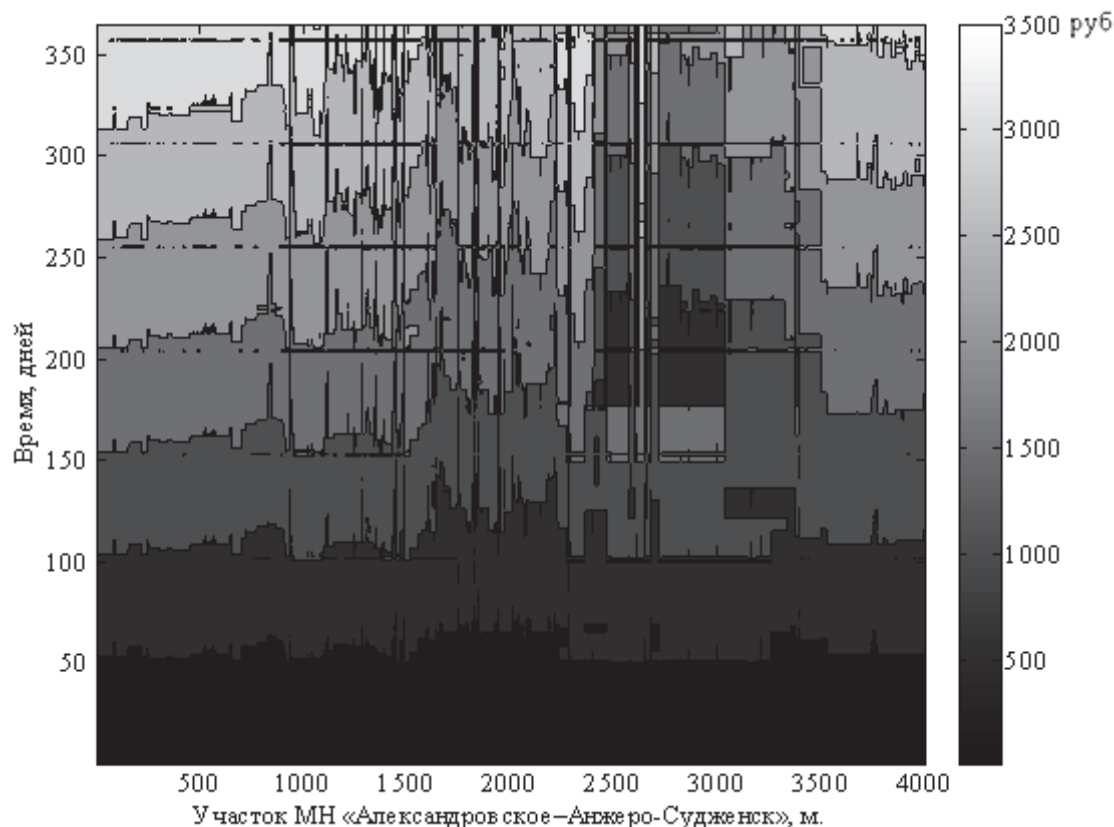


Рис. 5. Распределение накопленного риска на участке МН длиной 400 км

В ОАО «Центрсибнефтепровод» происходит внедрение автоматизированной системы (АС), реализующей указанную математическую модель по оценке величины риска физической безопасности. Среди специализированных АС схожей направленности можно выделить разработки зарубежных компаний по управлению рисками эксплуатации МН [23–25].

Первая очередь автоматизированной системы введена в промышленную эксплуатацию и используется во всех подразделениях СБ ОАО «Центрсибнефтепровод». АС позволяет осуществлять централизованный сбор и хранение данных для автоматизированного формирования входных параметров математической модели. На сегодняшний момент в АС зарегистрирована информация о 19 тыс. обследованных участках МН «Самотлор–Александровское», «Александровское–Анжеро-Судженск», «Игольско-Таловое–Парабель». Эти данные в совокупности образуют основной параметр модели – множество контрольных мероприятий M . Проведено категорирование МН, на основе которого рассчитаны коэффициенты ξ_i , произведена оценка коэффициентов k_{ij} .

Работы по внедрению второй очереди АС находятся на стадии разработки технического проекта. Завершение работ по внедрению АС позволит интегрировать разработанную математическую модель в производственные процессы по охране ЛЧ МН ОАО «Центрсибнефтепровод». Использование ко-

личественных показателей $W_{ip}(\bar{x}_i, \bar{t}_j)$ и $E_{ip}(\bar{x}_i, \bar{t}_j)$ позволит обеспечить руководящий персонал дополнительной информацией для принятия более эффективных управленческих решений при планировании мероприятий по охране ЛЧ МН.

Выводы

На основе методики оценки рисков в иерархических структурах критически важных объектов [4], с учетом отдельных элементов методик по управлению эксплуатационными рисками МН [9–13], авторами предложена методика оценки риска ФБ для пространственно-протяженных объектов типа ЛЧ МН, реализующая принципы ГОСТ Р ИСО/МЭК 31010–2011 «Менеджмент риска. Методы оценки риска». В рамках первого этапа оценки проведена идентификация рисков ФБ ЛЧ МН, предложен возможный вариант перечня источников риска и соответствующих им рисков событий. Для проведения дальнейших этапов по анализу и сравнительной оценке риска разработана математическая модель оценки рисков для пространственно-распределенных линейных объектов ТЭК типа ЛЧ МН.

Модель позволяет рассчитать пространственно-временное распределение величины риска ФБ, которое основывается на предположении о двумерном дискретном случайном распределении вероятности рисков события. Учитывается разная степень защищенности участков МН путем использо-

вания метода балльной оценки факторов, влияющих на степень опасности участков МН [15]. Результирующее значение величины риска ФБ определяется на основе данных о проведенных мероприятиях по контролю защищенности линейного объекта. Модель учитывает различную степень эффективности множества типов мероприятий.

В качестве модельной иллюстрации произведен расчет выходных параметров модели $W_i(x_i, t_i)$ и

$E_i(\bar{x}_i, \bar{t}_i)$ для различных участков МН «Александровское–Анжеро-Судженск» при входных параметрах модели, основанных на реальных данных и произвольных значениях внутренних параметров (реальные значения составляют коммерческую тайну ОАО «Центрсибнефтепровод»).

Выражаем благодарность Геннадию Наумовичу Глазову, ведущему научному сотруднику ЗАО «НПФ «Микран», за полезные консультации при подготовке статьи.

СПИСОК ЛИТЕРАТУРЫ

1. Федеральный закон Российской Федерации от 21 июля 2011 г. № 256-ФЗ «О безопасности объектов топливно-энергетического комплекса» // Российская газета – издание Правительства Российской Федерации, официальный публикатор документов. 2011. URL: <http://www.rg.ru/2011/07/26/tek-dok.html> (дата обращения: 28.01.2014).
2. Кононов А.А., Стилиславский А.Б., Цыгичко В.Н. Управление рисками нарушения транспортной безопасности. – М.: АС-Траст, 2008. – 210 с.
3. ГОСТ Р ИСО/МЭК 31000–2010. Менеджмент риска. Принципы и руководство. – М.: Стандартиформ, 2011. – 28 с.
4. Оценка рисков в иерархических структурах критически важных объектов / А.А. Кононов, К.В. Черныш, Д.С. Гуревич, А.К. Поликарпов // Труды ИСА РАН / под ред. Д.С. Черешкина. – М.: ЛЕНАНД, 2010. – Т. 52. – С. 5–15.
5. Управление кибербезопасностью автоматизированных информационных систем / В.Н. Бекетов, А.А. Кононов, С.Н. Осипов, А.А. Фобьянчук // Труды ИСА РАН. – 2011. – Т. 61. – С. 8–13.
6. Цыгичко В.Н., Кононов А.А., Поликарпов А.К. Концепция системы управления рисками нарушения таможенного законодательства // Труды ИСА РАН. – 2011. – Т. 31. – С. 21–41.
7. Стилиславский А.Б., Кононов А.А. Автоматизированная система управления безопасностью объектов транспортной инфраструктуры // Управление рисками и безопасностью: Труды ИСА РАН. – М.: ЛЕНАНД, 2009. – Т. 31. – С. 52–73.
8. Безопасность России. Правовые, социально-экономические и научно-технические аспекты. Анализ риска и проблем безопасности. В 4-х частях. Ч. 4. Научно-методическая база анализа риска и безопасности / науч. руковод. К.В. Фролов. – М.: МГОФ «Знание», 2007. – 864 с.
9. D'Atri M.F., Rodriguez D., Garcia-Martinez R. Improving pipeline risk models by using data mining techniques // 24th World Gas Conference. – Buenos Aires, 2009. URL: <http://www.igu.org/html/wgc2009/papers/docs/wgcFinal00663.pdf> (дата обращения: 28.01.2014).
10. Selva R., Eng C., Mech F.I. Risk Based Inspection (RBI) Best Practice: The Technical Specification for Ensuring Successful Implementation // 13th International Conference on Pressure Vessel & Piping Technology. – London, 2012. URL: http://www.ppsimtech.com/downloads/risk_based_inspection_best_practice-the_technical_specification_ron_selva_pp_simtech_keynote_paper-1_13th_icpvt_2012_london.pdf (дата обращения: 28.01.2014).
11. Selvik J.T., Scarf P., Aven T. An extended methodology for risk based inspection planning // Reliability: Theory & Applications. – 2011. – № 20. URL: http://gnedenko-forum.org/Journal/2011/012011/RTA_1_2011-12.pdf (дата обращения: 28.01.2014).
12. Оценка риска аварий на магистральных нефтепроводах КТК-Р и БТС / Ю.А. Дадонов и др. // Управление промышленной безопасностью. – 2002. URL: http://programs.safety.ru/BTP/2002_6/02_06_2-6.pdf (дата обращения: 28.01.2014).
13. Анализ риска магистральных нефтепроводов при обосновании проектных решений, компенсирующих отступления от действующих требований безопасности / Лисанов М.В. и др. // Безопасность Труда в Промышленности. Анализ рисков. – 2010. – № 3. URL: http://www.safety.ru/downloads/publications/analiz_riska/4.6.%20%D0%9B%D0%B8%D1%81%D0%B0%D0%BD%D0%BE%D0%B2.PDF (дата обращения: 28.01.2014).
14. ГОСТ Р ИСО/МЭК 31010–2011. Менеджмент риска. Методы оценки риска. – М.: Стандартиформ, 2011. – 74 с.
15. Кукало И.А. Модель угроз системы физической защиты линейной части магистрального нефтепровода // Известия Томского политехнического университета. – 2013. – Т. 322. – № 5. – С. 37–41.
16. Тарасенко Ф.П. Прикладной системный анализ. – М.: КНО-РУС, 2010. – 224 с.
17. Постановление Правительства Российской Федерации от 16 февраля 2008 г. № 87 г. Москва // Российская газета – издание Правительства Российской Федерации, официальный публикатор документов. 2014. URL: <http://www.rg.ru/2011/07/26/tek-dok.html> (дата обращения: 28.01.2014).
18. Jiang G., Cybenko G. Fellow Temporal and Spatial Distributed Event Correlation for Network Security // Proceeding of the 2004 American Control Conference. – Boston, 2004. URL: http://www.ntnu.no/users/skoge/prost/proceedings/acc04/Papers/0173_WeM11.4.pdf (дата обращения: 28.01.2014).
19. Fink D., Hochachka W.M. Spatiotemporal exploratory models for broad-scale survey data // The Ecological Society of America. – 2010. – № 20. URL: <http://faculty.bscb.cornell.edu/~hooker/STEM.pdf> (дата обращения: 28.01.2014).
20. Wang X., Brown D.E. The spatio-temporal modeling for criminal incidents // Security Informatics. 2012. URL: <http://www.security-informatics.com/content/1/1/2> (дата обращения: 28.01.2014).
21. Безопасность России. Правовые, социально-экономические и научно-технические аспекты. Анализ рисков и управление безопасностью: методические рекомендации. / науч. руковод. К.В. Фролов. – М.: МГОФ «Знание», 2008. – 672 с.
22. Querying Uncertain Spatio-Temporal Data / T. Emrich, H.-P. Kriegel, N. Mamoulis, M. Renz, A. Zuffe. URL: http://i.cs.hku.hk/~nikos/ICDE12_UncertainST.pdf (дата обращения: 28.01.2014).
23. The Synergi Pipeline software. URL: http://www.dnv.com/services/software/products/synergi/synergi_pipeline/index.asp (дата обращения: 28.01.2014).
24. Pipeline Risk Controller Version 2.0 for Windows. URL: <http://www.gulfpub.com/product.asp?ProductID=2428> (дата обращения: 28.01.2014).
25. The Multi-Analytic Risk Visualization. URL: http://www.dnv.com/moreondnv/research_innovation/programmes/materials/sensors/marv.asp (дата обращения: 28.01.2014).

Поступила 03.02.2014 г.

RISK ASSESSMENT OF LINEAR PART PHYSICAL SAFETY IN THE MAIN OIL PIPELINE

Ivan A. Kukalo,

Tomsk State University of Control Systems and Radioelectronics,
40, Lenin Avenue, Tomsk, 634050, Russia. E-mail: i@kukalo.ru

Sergey N. Grivtsov,

JSC «TsentrSibnefteprovod», 24, Naberezhnaya Reki Ushayki Street,
Tomsk, 634050, Russia. E-mail: GrivtsovSN@tom.transneft.ru

Relevance of the work is caused by need to protect the objects of the fuel and energy complex from illegal intervention.

The main aim of the study is to generalize and to develop a risk assessment technique in hierarchical structures of crucial objects for spatial distributed linear objects of energy industry of the type linear part of the main oil pipeline.

The methods used in the study: Mathematical modeling of risk event probability distribution, expert assessment by the Delfi method, ranging method, direct assessment method, numerical integration of risk two-dimensional function in time and space, interpolation of function of the held event trajectory.

The results: The authors have proposed the technique for assessing physical safety risk for linear part of the main oil pipeline, implementing GOST P ISO 31000–2010 concept. Physical safety risks of linear part of the main oil pipeline were identified; the mathematical model for assessing risks for spatial spread linear objects of energy industry like linear part of the main oil pipeline was developed to carry out the analysis and comparative assessment of risk. The model allows calculating existential distribution of physical safety risk dimension which is based on the assumption of two-dimensional discrete casual distribution of risk event probability. Resultant value of size of physical safety risk dimension is defined on the basis of data on the held events on control of linear object security. The authors calculated the output parameters of the model for the main pipeline «Aleksandrovsky–Anzhero-Sudzhensk» considering input parameters provided by JSC «TsentrSibnefteprovod».

Key words:

Risk assessment, oil pipeline, physical safety, matrix of risk distribution, spatiotemporal distribution, cumulative risk, level of risk.

REFERENCES

1. *Federalny zakon Rossiyskoy Federatsii ot 21 iyulya 2011 g. № 256-FZ «O bezopasnosti obektov toplivno-energeticheskogo kompleksa»* [Federal Law of 21 July 2011 № 256-FZ «On the safety of the fuel and energy complex»]. Available at: <http://www.rg.ru/2011/07/26/tek-dok.html> (accessed 28 January 2014).
2. Kononov A.A. *Upravlenie riskami narusheniya transportnoy bezopasnosti* [Risk management for transport safety violations]. Moscow, AS-Trast Publ., 2008. 210 p.
3. *GOST R ISO/MEK 31000–2010. Menedzhment riska. Printsipy i rukovodstvo* [State Standard 31000–2010. Risk management. Principles and Guidelines on Implementation]. Moscow, Standartinform, 2010. 28 p.
4. Kononov A.A., Chernysh K.V., Gurevich D.S., Polikarpov A.K. Otsenka riskov v ierarkhicheskikh strukturakh kriticheskikh vaznykh obektov [Risk assessment in hierarchical structures of critical facilities]. *Trudy ISA RAN*, 2010, no. 52, pp. 5–15.
5. Beketov V.N., Kononov A.A., Osipov S.N., Fobyanchuk A.A. Upravlenie kiberbezopasnostyu avtomatizirovannykh informatsionnykh sistem [Cybersecurity management of automated information systems]. *Trudy ISA RAN*, 2011, no. 61, pp. 8–13.
6. Tsygichko V.N., Kononov A.A., Polikarpov A.K. Kontseptsiya sistemy upravleniya riskami narusheniya tamozhennogo zakonodatelstva [The concept of risk management violations of customs legislation]. *Trudy ISA RAN*, 2011, no. 31, pp. 21–41.
7. Stislavskiy A.B., Kononov A.A. Avtomatizirovannaya sistema upravleniya bezopasnostyu obektov transportnoy infrastruktury [Automated safety management system of transport infrastructure]. *Trudy ISA RAN*, 2009, no. 31, pp. 52–73.
8. Frolov K.V. *Bezopasnost Rossii. Pravovye, sotsialno-ekonomicheskie i nauchno-tehnicheskie aspekty. Analiz riska i problem bezopasnosti* [Russia's security. Legal, socio-economic and scientific-technical aspects. Risk analysis and security issues. Scientific and methodological basis of risk analysis and security]. Moscow, Znanie Publ., 2007. P. 4, 864 p.
9. D'Atri M.F., Rodriguez D., Garcia-Martinez R. Improving pipeline risk models by using data mining techniques. *Proc. 24th World Gas Conference*. Buenos Aires, 2009. Available at: <http://www.igu.org/html/wgc2009/papers/docs/wgcFinal00663.pdf> (accessed 28 January 2014).
10. Selva R., Eng C., Mech F.I. Risk Based Inspection (RBI) Best Practice: The Technical Specification for Ensuring Successful Implementation. *Proc. 13th International Conference on Pressure Vessel & Piping Technology*. London, 2012. Available at: http://www.ppsimtech.com/downloads/risk_based_inspection_best_practice-the_technical_specification_ron_selva_pp_simtech_keynote_paper-1_13th_icpvt_2012_london.pdf (accessed 28 January 2014).
11. Selvik J.T., Scarf P., Aven T. An extended methodology for risk based inspection planning. *Reliability: Theory & Applications*, 2011, no. 20. Available at: http://gnedenko-forum.org/Journal/2011/012011/RTA_1_2011-12.pdf (accessed 28 January 2014).
12. Dadonov Yu.A. Otsenka riska avariyn na magistralnykh nefteprovodakh KTK-R i BTS [Assessment of risk of accidents on the main oil KTK-R and BTS]. *Upravlenie promyshlennoy bezopasnostyu*, 2002. Available at: http://programs.safety.ru/BTP/2002_6/02_06_2-6.pdf (accessed 28 January 2014).
13. Lisanov M.V. Analiz riska magistralnykh nefteprovodov pri obosnovanii proektnykh resheniy, kompensiruyushchikh otstupleniya ot deystviyushchikh trebovaniy bezopasnosti [Risk Analysis of main oil pipelines in justification of design decisions to compensate deviations from safety requirements]. *Bezopasnost Truda v Promyshlennosti. Analiz riskov*, 2010, no. 3. Available at: http://www.safety.ru/downloads/publik_s/analiz_riska/4.6.%20%D0%9B%D0%B8%D1%81%D0%B0%D0%BD%D0%BE%D0%B2.PDF (accessed 28 January 2014).

14. *GOST R ISO/MEK 31010–2011. Menedzhment riska. Metody otsenki riska* [State Standard 31010–2011. Risk management. Risk assessment techniques]. Moscow, Standartinform, 2011. 74 p.
15. Kukalo I.A. Model ugroz sistemy fizicheskoy zashchity lineynoy chasti magistralnogo nefteprovoda [The threat model of the physical protection of the linear part of the main pipeline]. *Bulletin of the Tomsk Polytechnic University*, 2013, vol. 322, no. 5, pp. 37–41.
16. Tarasenko F.P. *Prikladnoy sistemny analiz* [Applied System Analysis]. Moscow, KNORUS Publ., 2010. 224 p.
17. *Postanovlenie Pravitelstva Rossiyskoy Federatsii ot 16 fevralya 2008 g. № 87 g. Moskva* [Government Decree of 16 February 2008 № 87 Moscow]. Available at: <http://www.rg.ru/2011/07/26/tek-dok.html> (accessed 28 January 2014).
18. Jiang G., Cybenko G. Fellow Temporal and Spatial Distributed Event Correlation for Network Security. *Proc. of the 2004 American Control Conference*. Boston, 2004. Available at: http://www.nt.ntnu.no/users/skoge/prost/proceedings/acc04/Papers/0173_WeM11.4.pdf (accessed 28 January 2014).
19. Fink D., Hochachka W.M. Spatiotemporal exploratory models for broad-scale survey data. *The Ecological Society of America*, 2010, no. 20. Available at: <http://faculty.bscb.cornell.edu/~hooiker/STEM.pdf> (accessed 28 January 2014).
20. Wang X., Brown D.E. The spatio-temporal modeling for criminal incidents. *Security Informatics*. 2012. Available at: <http://www.security-informatics.com/content/1/1/2> (accessed 28 January 2014).
21. Frolov K.V. *Bezopasnost Rossii. Pravovye, sotsialno-ekonomicheskie i nauchno-tekhnicheskie aspekty. Analiz riskov i upravlenie bezopasnostyu: metodicheskie rekomendatsii* [Russia's security. Legal, socio-economic and scientific-technical aspects. Risk analysis and safety management: guidelines]. Moscow, Znanie Publ., 2008. 672 p.
22. Emrich T., Kriegel H.-P., Mamoulis N., Renz M., Zulfle A. *Querying Uncertain Spatio-Temporal Data*. Available at: http://i.cs.hku.hk/~nikos/ICDE12_UncertainST.pdf (accessed 28 January 2014).
23. *The Synergi Pipeline software*. Available at: http://www.dnv.com/services/software/products/synergi/synergi_pipeline/index.asp (accessed 28 January 2014).
24. *Pipeline Risk Controller Version 2.0 for Windows*. Available at: <http://www.gulfpub.com/product.asp?ProductID=2428> (accessed 28 January 2014).
25. *The Multi-Analytic Risk Visualization*. Available at: http://www.dnv.com/moreondnv/research_innovation/programmes/materials/sensors/marv.asp (accessed 28 January 2014).

УДК 622.276.5.001.42:519.688

МЕТОД АДАПТИВНОЙ ИДЕНТИФИКАЦИИ И ДИАГНОСТИКИ НЕФТЯНЫХ ПЛАСТОВ В ПРОЦЕССЕ НОРМАЛЬНОЙ ЭКСПЛУАТАЦИИ СКВАЖИН И КРАТКОВРЕМЕННЫХ ОСТАНОВОК

Романова Екатерина Валерьевна,аспирант кафедры геологии и разработки нефтяных месторождений
Института природных ресурсов Томского политехнического университета,
Россия, 634050, г. Томск, пр. Ленина, 30. E-mail: Romanova88EV@mail.ru**Сергеев Виктор Леонидович,**д-р техн. наук, профессор кафедры геологии и разработки нефтяных
месторождений Института природных ресурсов Томского политехнического
университета, Россия, 634050, г. Томск, пр. Ленина, 30.
E-mail: SergeevVL@ignd.tpu.ru

Актуальность работы обусловлена необходимостью определения параметров нефтяных пластов в процессе нормальной эксплуатации скважин, оснащенных стационарными измерительными системами для гидродинамического мониторинга и управления разработкой нефтяных месторождений.

Цель исследования: разработка метода адаптивной идентификации и диагностики, позволяющего определять параметры нефтяных пластов в процессе нормальной эксплуатации скважин и кратковременных остановок, учитывать дополнительную априорную информацию.

Методы исследования: использованы теоретические и практические разработки в области гидродинамических исследований скважин, системного анализа, идентификации систем с учетом дополнительной априорной информации, оптимизации функций и линейной алгебры. Решение задач идентификации и диагностики проводилось на основе промысловых данных забойного давления и дебита скважин, оснащенных стационарными измерительными системами, с учетом экспертных оценок фильтрационных параметров пласта и с использованием компьютерной программы Saphir.

Результаты: разработан метод адаптивной идентификации и диагностики, позволяющий повысить точность и устойчивость оценок фильтрационных параметров нефтяных пластов в процессе получения данных о забойном давлении и дебите скважин в условиях их нормальной эксплуатации и кратковременных остановок. На примере обработки промысловых данных гидродинамических исследований скважин по кривой восстановления давления нефтяного месторождения проведен сравнительный анализ методов определения фильтрационных параметров однородно-пористого нефтяного пласта.

Ключевые слова:

Идентификация, диагностика, адаптация, гидродинамические исследования скважин, кривая восстановления давления, априорная информация, нефтяные пласты.

Введение

В настоящее время в связи с возможностью получения информации в режиме реального времени стационарными информационно измерительными системами (СИИС) актуальным является создание новых технологий идентификации и интерпретации гидродинамических исследований скважин (ГДИС), позволяющих оперативно, в процессе нормальной эксплуатации скважин (без их длительных остановок), определять параметры нефтяных пластов и скважин, необходимые для решения задач мониторинга и управления разработкой месторождений нефти и газа [1, 2].

Следует отметить, что традиционные, наиболее информативные технологии идентификации и интерпретации ГДИС на неустановившихся режимах фильтрации, требуют проведения достаточно длительных остановок скважин с детальным анализом кривой восстановления давления (КВД) и ее производной, выделением на КВД участков радиального притока, использованием графо-аналитических методов интерпретации и т. п. [2–7]. Данные технологии носят достаточно трудоемкий и затратный характер, поскольку интерпретация

КВД производится после завершения заранее спланированных исследований, что часто приводит к продолжительным простоям скважин и значительной потере добычи нефти. Для решения отмеченных выше проблем, сокращения времени простоя скважин в [8, 9] предложен метод адаптивной интерпретации ГДИС по КВД, позволяющий получать параметры нефтяных пластов в процессе проведения исследований, не планируя заранее время их завершения.

На сегодняшний день в связи с тем, что на нефтегазовых месторождениях России большинство скважин механизированного фонда оборудованы телеметрическими системами, позволяющими осуществлять длительный мониторинг гидродинамических параметров (забойное давление на приеме насоса, фазовый дебит скважины и т. д.), актуальным является определение параметров пласта в процессе нормальной эксплуатации скважин. Однако в процессе эксплуатации месторождения по разным техническим причинам (срыв подачи, остановка насоса, ожидание геолого-технических мероприятий на скважине и т. п.) осуществляются остановки скважин различной длительностью – от

часов до нескольких суток с регистрацией случайных КВД измерительными системами. В этой связи актуальной является задача идентификации и диагностики параметров нефтяных пластов в условиях нормальной эксплуатации скважин и их кратковременных остановок.

В работе приводится решение задачи идентификации и диагностики нефтяных пластов с использованием метода адаптивной интерпретации ГДИС по КВД на основе интегрированных систем моделей дебита скважин и забойного давления с зависящими от времени параметрами с учетом априорной информации и экспертных оценок.

Метод адаптивной идентификации и диагностики нефтяных пластов

Основой метода адаптивной идентификации и диагностики нефтяных пластов являются интегрированная система моделей КВД [10, 11]

$$\begin{cases} y_{3,n}^* = y_{3,n} + \xi_n = f_3(t_n, b_n, q_0) + \xi_n, \\ \bar{V}_n = V_n + \eta_n = G_a(b_n, q_0) + \varepsilon_n, \quad n = \overline{1, n_k} \end{cases} \quad (1)$$

и интегрированная систем моделей дебита скважины [12]

$$\begin{cases} q_n^* = q_n + \xi_n = f_0(a_n, p_{3,n}) + \xi_n, \\ \bar{Z}_n = Z_n + \eta_n = F_a(a_n, q_n) + \eta_n, \quad n = \overline{1, n_n}, \end{cases} \quad (2)$$

Особенность моделей (1) и (2) заключается в том, что их параметры $a_n = ((\alpha_j(t_n)), j = \overline{1, m})$, $b_n = ((b_k(t_n)), k = \overline{1, p})$ – неизвестные однозначные функции $\alpha_j(t_n)$, $b_k(t_n)$ дискретного времени t_n . В моделях (1) и (2) приняты следующие обозначения: $y_{3,n}^*$, $y_{3,n}$ – фактические значения забойного давления, полученные в моменты времени $t_n \in [t_n, t_{n+1}]$ и вычисленные на основе модели КВД $f_3(t_n, b_n, q_0)$; V_n , $\bar{V}_n$ – векторы известных к моменту времени t_n дополнительных априорных данных о параметрах нефтяного пласта и скважины $\bar{V}_n = (\bar{v}_{1n}, \bar{v}_{12n}, \dots, \bar{v}_{dn})$ и априорных сведений V_n , полученных на основе моделей $G_a = (g_{a1}, g_{a2}, \dots, g_{ad})$; q_0 – значения дебита скважины в момент ее остановки t_n ; q_n^* , $p_{3,n}$ – фактические значения дебита и забойного давления скважин в момент времени нормальной эксплуатации скважин $t_n \in [t_1, t_n]$ (предполагается, что забойные давления измеряются точно либо с малой ошибкой); q_n – значения дебита скважины, полученные на основе модели $f_0(a_n, p_{3,n})$; Z_n , $\bar{Z}_n$ – векторы известных к моменту времени t_n дополнительных априорных данных $\bar{Z}_n = (\bar{z}_{1n}, \bar{z}_{12n}, \dots, \bar{z}_{pn})$ и априорных сведений Z_n , полученных на основе моделей $F_a = (f_{a1}, f_{a2}, \dots, f_{ap})$; ξ_n , η_n , ε_n – случайные величины, представляющие погрешности измерений дебита и забойных давлений скважины, ошибки дополнительных данных и экспертных оценок, неточность моделей и т. п.

Примерами моделей (1), (2) являются интегрированная система моделей КВД с учетом априорной информации о гидропроводности $\bar{\sigma}_n$ и пластовом давлении $\bar{p}_{n,n}$ [10, 11]

$$\begin{cases} P_{3,i} = f(t_i, \sigma_i, k_{p,i}, q_0) + \xi_i = \\ = P_3(t_n) + \frac{q_0}{4\pi\sigma_i} \ln(2, 25k_{p,i}t_i) + \xi_i, \\ \bar{\sigma}_n = \sigma_i + \eta_{li}, \\ \bar{p}_{n,n} = f(T, \sigma_i, k_{p,i}, q_0) + \eta_{2i}, \quad i = \overline{1, n_k}, \end{cases} \quad (3)$$

и интегрированная система моделей дебита скважины с учетом априорной информации k_n , $\bar{p}_n$ о продуктивности k_n и пластовом давлении p_n [12]

$$\begin{cases} q_n^* = k_n(p - p_{3,n}) + \xi_n, \\ \bar{k}_n = k_n + \eta_{1n}, \bar{p}_n = p_n + \eta_{2n}, \quad n = \overline{1, n_n}, \end{cases} \quad (4)$$

где σ_i , $k_{p,i} = \chi_{np}/r_{np}^2$ – гидропроводность и комплексный параметр в момент времени $t_i \in [t_n, t_{n+1}]$; T – время восстановления забойного давления до пластового; $P_{3,i}$, $P_3(t_n)$ – текущие значения забойного давления в моменты времени t_i и в момент остановки скважины. Используя обозначения $a_{1,n} = k_n p_n$, $a_{2,n} = k_n$ в модели дебита скважины (4) и $y_i^* = P_{3,i} - P_3(t_n)$ $\bar{\alpha}_{2n} = \frac{q_0}{4\pi\sigma_n} \ln(2, 25k_p)$, $\bar{\alpha}_{2n} = \frac{q_0}{4\pi\sigma_n}$ в модели КВД (3), уравнения (3), (4) не трудно привести к стандартным формам (1), (2) и далее к матричному виду (8).

Из моделей (3), (4) видно, что решение задачи идентификации по определению параметров моделей (1), (2) решает и задачу диагностики по определению параметров нефтяных пластов и скважин.

Следует отметить, что целесообразность совместного использования моделей (1), (2) и, соответственно, моделей (3), (4) возникает с необходимостью решения задачи по определению скин-фактора скважины, характеризующего степень ее загрязнения. Действительно, для определения скин-фактора скважины S , согласно формуле Дюпюи [13]

$$k_p = \frac{2\pi\sigma}{\ln(R_k / r_c) + S}, \quad (5)$$

недостаточно знать ее продуктивность k_p , необходимо определить и гидропроводность пласта $\sigma = kh/\mu$, где k – проницаемость, h – эффективная (рабочая) толщина пласта, μ – вязкость нефти, R_k – радиус контура питания скважины, r_c – радиус скважины. Известно, что наиболее достоверное определение гидропроводности пласта осуществляется при проведении гидродинамических исследований скважин по КВД [2, 3].

Рассмотрим метод решения задачи адаптивной идентификации на примере модели (2), предполагая, без потери общности, что ее параметры постоянны $a(t) = \alpha_n = \text{const}$ на текущем интервале обработки $t \in [t_n - \tau, t_n] \in [t_1, t_n]$ с номером n , который может быть определен, например, с помощью весовой функции $w_i = 1$ при $t \in [t_n - \tau, t_n]$ и $w_i = 0$ при $t \notin [t_n - \tau, t_n]$. Считаем далее, что приведенные выше допущения относительно параметров модели (2) справедливы и для параметров модели КВД (1).

С учетом приведенных допущений на параметры интегрированной системы моделей дебита

скважины (2), используя технологию интегрированных моделей [10, 11], процесс адаптивной идентификации можно представить в виде последовательного решения оптимизационных задач в моменты времени t_n

$$a_n^* (\beta_n, h_n) = \arg \min_{a_n} \Phi(a_n, \beta_n, h_n), \quad (6)$$

$$\beta_n^*, h_n^* = \arg \min_{\beta_n, h_n} J_0(a_n^* (\beta_n, h_n)), \quad (7)$$

где запись $\arg \min_x f(x)$ означает точку минимума x^* функции $f(x)$ ($f(x^*) = \min_x f(x)$); $\Phi(a_n, \beta_n, h_n) = \Phi(J_0(a_n, h_n), J_a(a_n, \beta_n))$ – комбинированный эмпирический показатель качества системы моделей (1), представляющий заданную функцию (функционал) Φ от частного показателя качества модели дебита скважины $f_0(a_n, p_{3,n})$

$$J_0(a_n) = \sum_{i=1}^n w_i(h_n) \psi_0(q_i^* - f_0(a_n, p_{3,i}))$$

и частного критерия качества моделей дополнительных априорных сведений $F_a(a_n, q_n)$

$$J_a(a_n, \beta_n) = \sum_{k=1}^p \beta_{k,n} \psi_{a,k}(\bar{z}_{k,n} - f_{a,k}(a_n, q_n)),$$

где $\beta_n = (\beta_{k,n}, k=\overline{1,m})$ – вектор управляющих параметров, определяющих значимость (вес) дополнительных априорных сведений $\bar{z}_{k,n}$; $w_i(h_n)$ – весовые функции с параметром h_n ; $\psi_0, \psi_{a,j}$ – известные функции.

Рассмотренный метод адаптивной идентификации (6), (7) позволяет синтезировать достаточно широкий спектр алгоритмов для линейных и нелинейных интегрированных систем моделей (1), (2) с переменными во времени параметрами, а также для различных показателей качества, определяемых функциями ψ_0, ψ_a . Например, для линейной по параметрам a_n дискретной интегрированной системы моделей дебита скважин (4) в матричном виде

$$\begin{cases} Y_n^* = F_0 a_n + \xi_n, \\ \bar{Z}_n = F_a a_n + \eta_n \end{cases} \quad (8)$$

и комбинированного показателя качества

$$\Phi(a_n, \beta_n) = \|Y_n^* - F_0 a_n\|_{K(h)}^2 + \|\bar{Z}_n - F_a a_n\|_{W(\beta)}^2 \quad (9)$$

задача оптимизации (6) сводится к решению систем линейных алгебраических уравнений

$$\begin{aligned} (F_0^T K(h_n) F_0 + F_a^T W(\beta_n) F_a) a_n = \\ = F_0^T K(h_n) Y_n^* + F_a^T W(\beta_n) \bar{Z}_n, \end{aligned} \quad (10)$$

где запись $\|X\|_W^2$ означает квадратичную форму $X^T W X$; $Y_n^* = (q_i^*, i=\overline{1,n})$ – вектор; $F_0 = (1, p_{3,i}, i=\overline{1,n})$ – матрица размерности $(2 \times n)$; $F_a = \text{diag} I$ – диагональная единичная матрица размерности (2×2) ; $W(\beta_n) = \text{diag}(\beta_{1,n}, \beta_{2,n})$ – диагональная матрица размерности (2×2) ; $K(h_n) = \text{diag}(w_i(h_n), i=\overline{1,n})$ – диагональная матрица весовых функций $w_i(h_n)$ с параметром h_n . Для получения системы линейных уравнений (10) достаточно взять частные производные по параметрам a_n от (9) и приравнять их к нулю.

Результаты адаптивной идентификации и диагностики

Результаты идентификации и диагностики скважины № 535 нефтяного месторождения Томской области приведены на рис. 1–6 и в таблице. На рис. 1, 2 для примера приведены суточные значения забойного давления и дебита скважины в условиях ее нормальной эксплуатации за 260 дней работы. Далее, в связи с ремонтными работами, скважина была остановлена на 305 часов. Значения забойного давления с датчика телеметрической системы на приеме насоса (в пересчете на уровень перфорации скважины) и производной забойного давления приведены на рис. 5 линиями 1 и 2 соответственно.

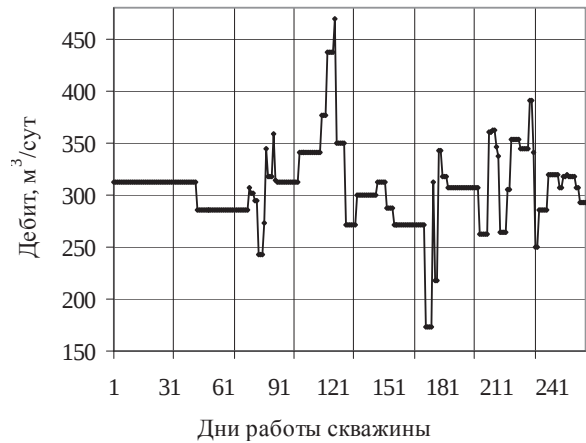


Рис. 1. Значения дебита скважины

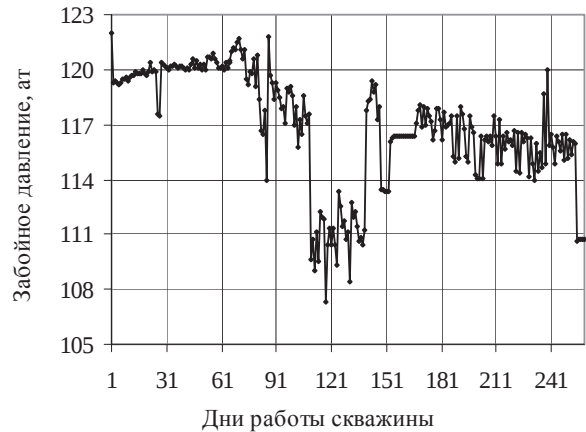


Рис. 2. Значения забойного давления скважины

На рис. 3, 4 приведены оценки коэффициента продуктивности скважины $k_n^*(u_n^*)$ и пластового давления $P_n^*(u_n^*)$ (линия 1), полученные методом адаптивной интерпретации с учетом априорной информации (10)

$$\begin{aligned} k_n^*(u_n^*) = \alpha_{2,n}^*(u_n^*), \quad P_n^*(u_n^*) = \alpha_{1,n}^*(u_n^*) / \alpha_{2,n}^*(u_n^*), \\ u_n^* = (\beta_n^*, h_n^*) \end{aligned} \quad (11)$$

с использованием интегрированной системы моделей дебита скважины (3) и экспертных оценок коэффициента продуктивности $\bar{k}_n = 4 \text{ м}^3/\text{сут}/\text{ат}$ и пластового давления $\bar{p} = 190 \text{ ат}$. Линией 2 изображены оценки коэффициента продуктивности скважины

$$k_n^*(\lambda_n^*, h_n^*) = \alpha_{2n}^*(\lambda_n^*, h_n^*),$$

$$P_{пл,n}^*(\lambda_n^*, h_n^*) = \alpha_{1n}^*(\lambda_n^*, h_n^*) / \alpha_{2n}^*(\lambda_n^*, h_n^*), \quad (12)$$

полученные по аналогии с (11) адаптивным методом идентификации с регуляризацией по А.Н. Тихонову [14] путем решения системы линейных уравнений

$$(F_0^T K(h_n) F_0 + \lambda_n I) a_n(\lambda_n, h_n) = F_0^T K(h_n) Y_n^*, \quad (13)$$

где λ_n – параметр регуляризации, I – единичная диагональная матрица. Матрицы K , F_0 и вектор Y_n^* определены в (10). Диагональная матрица $K(h_n)$ (10) выбиралась в виде

$$K(h_n) = \text{diag}(w_i(h_n)) =$$

$$= \exp(-(t_n - t_{n-i}) / h_n), i = \overline{1, n-i}.$$

Линией 3 изображены оценки метода наименьших квадратов

$$k_n^*(0, h_n^*) = \alpha_{2n}^*(0, h_n^*),$$

$$P_{пл,n}^*(0, h_n^*) = \alpha_{1n}^*(0, h_n^*) / \alpha_{2n}^*(0, h_n^*), \quad (14)$$

полученные из (13) при $w_i=1, i=\overline{1, n}, 1 < n \leq 260, \lambda_n=0$, что соответствует последовательному расширению интервала обработки до предельного значения $n_k=260$, когда интервал обработки содержит все исходные данные дебита и забойного давления скважины $q_i, p_{a,i}, i=\overline{1, 260}$.

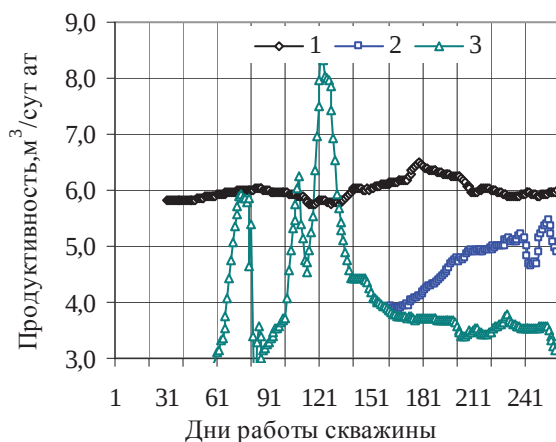


Рис. 3. Оценки продуктивности скважины

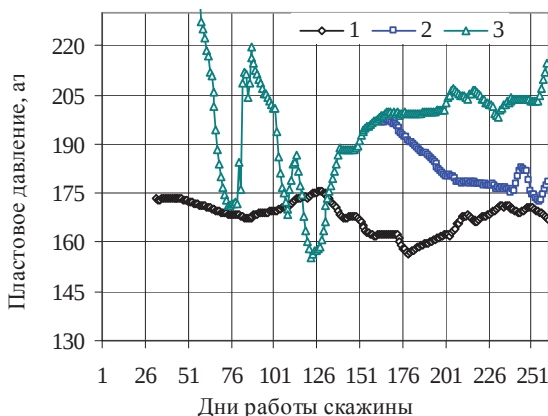


Рис. 4. Оценки пластового давления скважины

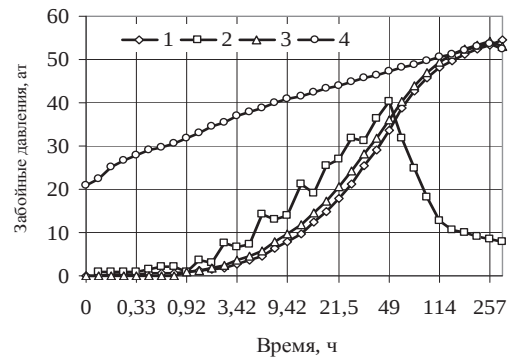


Рис. 5. Фактические (линия 1) и восстановленные (линии 3, 4) значения забойного давления скважины, линия 2 – производная забойного давления

На рис 6. приведены оценки гидропроводности (линия 1)

$$\sigma_n^*(\beta_n^*, h_n^*) = 1 / 4\pi\alpha_{2n}^*(\beta_n^*, h_n^*), \quad (15)$$

полученные на основе КВД (рис. 5, линия 1) методом адаптивной идентификации (АИ) (6), (7) путем решения системы линейных уравнений вида (10) с использованием интегрированной системы моделей КВД (4) и экспертных оценок коэффициента пластового давления $\bar{p}=190$ ат и гидропроводности $\bar{\sigma}_n=3$ м³сут/ат (в размерности, аналогичной размерности проницаемости (5)). Линией 2 изображены оценки гидропроводности

$$\sigma_n^*(\lambda_n^*, h_n^*) = 1 / 4\pi\alpha_{2n}^*(\lambda_n^*, h_n^*) \quad (16)$$

с параметром регуляризации λ_n^* , полученные (по аналогии с (12), (13)) методом АИ с регуляризацией по А.Н. Тихонову. Линией 3 изображены оценки гидропроводности, полученные по адаптивному методу наименьших квадратов при $\lambda_n^*=0$

$$\sigma_n^*(0, h_n^*) = 1 / 4\pi\alpha_{2n}^*(0, h_n^*). \quad (17)$$

Оценки управляющих параметров β_n^* , параметра регуляризации λ_n и параметра h_n^* определялись путем решения оптимизационной задачи (7) методом деформированного многогранника [15].

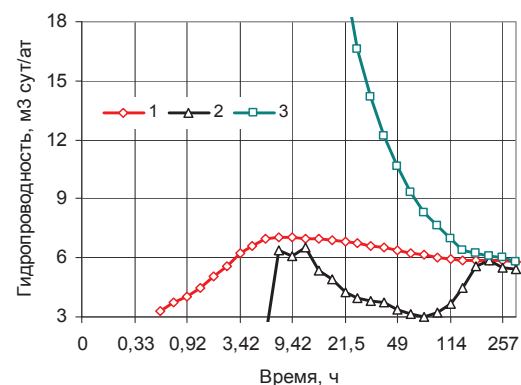


Рис. 6. Оценки гидропроводности пласта методами АИ (линии 1-3)

Из рис. 6 видно, что оценки гидропроводности практически совпадают на участке радиального притока скважины, где производная КВД, изображенная на рис. 5 линией 2, принимает постоянные

значения. Однако оценки гидропроводности с учетом априорной информации (рис. 6, линия 1) более стабильны и точны начиная с 10 часов после остановки скважины, что позволяет использовать их и для вычисления скин-фактора при кратковременных остановках скважины на случайных КВД.

Таблица. Результаты диагностики нефтяного пласта и скважины

Методы	Время, ч	Гидропроводность, м ² сут/ат	Скин-фактор скважины
«Saphir»	10	–	–
	25	–	–
	305	5,12	–2,35
АИ с учетом априорной информации	10	6,95	–1,01
	25	6,73	–1,24
	305	5,85	–2,16
АИ с регуляризацией по А.Н. Тихонову	10	6,36	–1,63
	25	3,78	–4,33
	305	5,40	–2,63
Наименьших квадратов	10	25,43	–8,28
	25	14,16	6,54
	305	5,80	–2,21

В таблице приведены оценки гидропроводности пласта и скин-фактора скважины при кратковременных, на 10 и 25 часов, и длительной – на 305 часов, остановках скважины с использованием методов адаптивной интерпретации и зарубежного программного комплекса Saphir. Оценки скин-фактора скважины S^* вычислялись с использованием формулы (5)

$$S^* = \frac{2\pi\sigma^*}{k_p^*} - \ln(R_k / r_c)$$

при радиусе контура питания скважины $R_k=400$ м и радиусе скважины $r_c=0,1$ м. Оценки гидропроводности σ^* вычислялись по формулам (15)–(17), а оценки продуктивности скважины k_p^* рассчитыва-

лись в моменты остановки скважины по формулам (11), (12), (14) соответственно.

Из таблицы видно, что метод адаптивной идентификации с учетом априорной информации обеспечивает получение более точных оценок гидропроводности нефтяного пласта и скин-фактора при кратковременных остановках скважины.

Выводы

1. Предложен метод адаптивной идентификации и диагностики, позволяющий определять параметры нефтяных пластов в процессе нормальной эксплуатации скважин и кратковременных остановок, учитывать дополнительную априорную информацию.
2. Метод адаптивной идентификации и диагностики позволяет синтезировать достаточно широкий спектр оптимальных алгоритмов идентификации параметров нефтяных пластов и скважин в зависимости от вида интегрированных моделей дебита и забойного давления, заданных показателей качества моделей и статистических характеристик случайных неконтролируемых факторов.
3. На примере обработки промысловых данных нормальной эксплуатации и гидродинамических исследований скважины нефтяного месторождения Томской области по кривой восстановления давления показано, что метод адаптивной идентификации и диагностики позволяет повысить точность оценок параметров нефтяных пластов и скин-фактора скважины.
4. Предложенные модели и алгоритмы адаптивной идентификации и диагностики нефтяных пластов могут быть использованы в системах оперативной обработки промысловых данных интеллектуальных скважин, оснащенных стационарными информационно измерительными системами.
9. Гаврилов К.С., Сергеев В.Л. Адаптивная интерпретация нестационарных гидродинамических исследований скважин в системе «пласт–скважина» методом интегрированных моделей // Известия Томского политехнического университета. – 2012. – Т. 321. – № 5. – С. 72–75.
10. Сергеев В.Л. Интегрированные системы идентификации. – Томск: Изд-во Томского политехнического университета, 2011. – 198 с.
11. Сергеев П.В., Сергеев В.Л. Идентификация гидродинамических параметров скважин на неустановившихся режимах фильтрации с учетом априорной информации // Известия Томского политехнического университета. – 2006. – Т. 309. – № 5. – С. 156–161.
12. Сергеев В.Л., Севастьянов Д.В. Интегрированные модели и алгоритмы идентификации дебита скважин нефтяных месторождений // Известия Томского политехнического университета. – 2006. – Т. 321. – № 6. – С. 149–152.
13. Мищенко И.Т. Скважинная добыча нефти. – М.: Изд-во «Нефть и газ» РГУ нефти и газа им. И.М. Губкина, 2003. – 816 с.
14. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. – М.: Наука, 1979. – 392 с.
15. Пантелеев А.В., Летова Т.А. Методы оптимизации в примерах и задачах. – М.: Высшая школа, 2002. – 544 с.

Поступила 20.12.2013 г.

СПИСОК ЛИТЕРАТУРЫ

1. Ипатов А.И., Кременецкий М.И., Гуляев Д.Н. Современные технологии гидродинамических исследований скважин и их возрастающая роль в разработке углеводородов // Нефтяное хозяйство. – 2009. – № 1. – С. 52–55.
2. Ипатов А.И., Кременецкий М.И., Гуляев Д.Н. Информационное обеспечение и технологии гидродинамического моделирования нефтяных и газовых скважин. – М.; Ижевск: Институт компьютерных исследований, 2012. – 896 с.
3. Шагиев Р.Г. Исследование скважин по КВД. – М.: Наука, 1998. – 304 с.
4. Bourdet D., Ayoub J.A., Pirard Y.M. Use of pressure derivative in well test interpretation // SPE. – 1984. – № 12777. – Р. 293–302.
5. Arps J.J. Analysis of Decline Curves. Trans. AIME. – 1945. – V. 160. – Р. 228–247.
6. Fetkovich M.J. Decline Curve Analysis Using Type Curves // JPT. – 1980. – June. – Р. 1065–1077.
7. Blasingame T.A., Lee W.J. Pressure Buildup Test Analysis – Variable-Rate Case: A New Approach // SPE-17052-PA. SPE Formation Evaluation. – 1989. – V. 4. – Iss. 02. – Р. 273–280. DOI: <http://dx.doi.org/10.2118/17052-PA>
8. Сергеев В.Л., Аниканов А.С. Метод адаптивной идентификации гидродинамических исследований скважин с учетом априорной информации // Известия Томского политехнического университета. – 2010. – Т. 317. – № 5. – С. 50–52.

UDC 519.688:622.276

METHOD OF ADAPTIVE IDENTIFICATION AND DIAGNOSIS OF OIL RESERVOIRS IN WELL NORMAL OPERATION AND SHORT STOPPAGES

Ekaterina V. Romanova,

Tomsk Polytechnic University, 30, Lenin Avenue, Tomsk, 634050, Russia.

E-mail: Romanova88EV@mail.ru

Viktor L. Sergeev,

Dr. Sc., Tomsk Polytechnic University, 30, Lenin Avenue,

Tomsk, 634050, Russia. E-mail: SergeevVL@ignd.tpu.ru

Relevance of the work is caused by the need to determine oil reservoir parameters in normal operation of the wells equipped with stationary measuring systems for hydrodynamic monitoring and control of oil field development.

Objective: To develop a method for adaptive identification and diagnosis which allows determining the parameters of oil reservoirs in well normal operation and short stoppages, observing additional a priori information.

Methods: The authors have used the theoretical and practical developments in the field of well testing, system analysis, system identification based on additional a priori information, optimization of functions and linear algebra. The identification and diagnosis problems were solved on the basis of bottom-hole pressure field data and flow rate of the wells equipped with stationary measuring systems considering expert estimates of reservoir filter parameters and using a computer program Saphir.

Results: The authors developed the method of adaptive identification and diagnosis to improve the accuracy and stability of estimates of filtration parameters for oil reservoirs when obtaining data on bottom-hole pressure and flow rate of wells in their normal operation and short stoppages. By the example of processing the field data of well hydrodynamic study on pressure recovery curve of oilfield the authors carried out the comparative analysis of the methods for determining filtration parameters of uniformly porous oil reservoir.

Key words:

Identification, diagnostics, adaptation, hydrodynamic study of wells, pressure build-up curve, a-priori information, oil reservoirs.

REFERENCES

1. Ipatov A.I., Kremenetskiy M.I., Gulyaev D.N. Sovremennye tekhnologii gidrodinamicheskikh issledovaniy skvazhin i ikh vozrastayushchaya rol v razrabotke uglevodorodov [Modern technologies of well testing and their increasing role in the development of hydrocarbon]. *Neftyanoe khozyaystvo*, 2009, no. 1, pp. 52–55.
2. Ipatov A.I., Kremenetskiy M.I., Gulyaev D.N. *Informatsionnoe obespechenie i tekhnologii gidrodinamicheskogo modelirovaniya neftnykh i gazovykh skvazhin* [The information and technology of hydrodynamic modeling of oil and gas wells]. Moscow; Izhevsk, Institut kompyuternykh issledovaniy, 2012. 896 p.
3. Shagiev R.G. *Issledovanie skvazhin po KVD* [Well surveying by pressure build-up]. Moscow, Nauka, 1998. 304 p.
4. Bourdet D., Ayoub J.A., Pirard Y.M. Use of pressure derivative in well test interpretation. *SPE*, 1984, no. 12777, pp. 293–302.
5. Arps J.J. Analysis of Decline Curves. *Trans. AIME*, 1945, vol. 160, pp. 228–247.
6. Fetkovich M.J. Decline Curve Analysis Using Type Curves. *JPT*, 1980, June, pp. 1065–1077.
7. Blasingame T.A., Lee W.J. Pressure Buildup Test Analysis – Variable-Rate Case: A New Approach. *SPE-17052-PA. SPE Formation Evaluation*, 1989, vol. 4, Iss. 02, pp. 273–280. DOI: <http://dx.doi.org/10.2118/17052-PA>
8. Sergeev V.L., Anikanov A.S. Metod adaptivnoy identifikatsii gidrodinamicheskikh issledovaniy skvazhin s ucheto apriornoy informatsii [Method of adaptive identification of well testing, taking into account a priori information]. *Bulletin of Tomsk Polytechnic University*, 2010, vol. 317, no. 5, pp. 50–52.
9. Gavrilov K.S., Sergeev V.L. Adaptivnaya interpretatsiya nestatsionarnykh gidrodinamicheskikh issledovaniy skvazhin v sisteme «plast-skvazhina» metodom integrirovannykh modeley [Adaptive interpretation of transient well test in the «layer-well» method of integrated models]. *Bulletin of Tomsk Polytechnic University*, 2012, vol. 321, no. 5, pp. 72–75.
10. Sergeev V.L. *Integrirovannye sistemy identifikatsii* [Integrated System of Identification]. Tomsk, TPU Publ. House, 2011. 198 p.
11. Sergeev P.V., Sergeev V.L. Identifikatsiya gidrodinamicheskikh parametrov skvazhin na neustanovivshikh rezhimakh filtratsii s ucheto apriornoy informatsii [Identification of well hydrodynamic parameters on unsteady filtration modes, taking into account a priori information]. *Bulletin of Tomsk Polytechnic University*, 2006, vol. 309, no. 5, pp. 156–161.
12. Sergeev V.L., Sevastyanov D.V. Integrirovannye modeli i algoritmy identifikatsii debita skvazhin neftnykh mestorozhdeniy [Integrated models and algorithms of oil field workover identification]. *Bulletin of Tomsk Polytechnic University*, 2006, vol. 321, no. 6, pp. 1149–152.
13. Mishchenko I.T. *Skvazhinnaya dobycha nefti* [Oil Production]. Moscow, Neft i gaz RGU nefti i gaza im. I.M. Gubkina, 2003. 816 p.
14. Tikhonov A.N., Arsenin V.Ya. *Metody resheniya nekorrektnykh zadach* [Methods for solving ill-posed problems]. Moscow, Nauka, 1979. 392 p.
15. Pantelev A.V., Letova T.A. *Metody optimizatsii v primerakh i zadachakh* [Optimization methods in the examples and problems]. Moscow, Vyshaya shkola Publ, 2002. 544 p.

Интеллектуальные системы

УДК 004.056.53

ПРИМЕНЕНИЕ ЦИФРОВЫХ ВОДЯНЫХ ЗНАКОВ ДЛЯ РАЗРАБОТКИ СИСТЕМ ГРАФИЧЕСКОГО ПАРОЛЯ

Шокарев Алексей Владимирович,

канд. техн. наук, доцент кафедры информационных технологий
Юргинского технологического института (филиала) Томского
политехнического университета, Россия, 652055,
г. Юрга, ул. Ленинградская, д. 26. E-mail: Shokarev_av@mail.ru

Актуальность работы состоит в повышении безопасности защищаемых ресурсов при прохождении идентификации/аутентификации пользователей средствами систем графического пароля, которую можно обеспечить с помощью применения методов стеганографии.

Цель работы заключается в повышении достоверности идентификации/аутентификации пользователей за счет применения системы графического пароля с использованием цифровых водяных знаков. Применение методов стеганографии повышает стойкость систем к большинству известных методов взлома систем с парольной аутентификацией пользователей.

Методы исследования: методы теории вероятности, математического анализа, стеганографии и экспериментальных исследований. Программная реализация метода выполнена с использованием среды программирования Borland C++.

Результаты: показана возможность создания системы графических паролей с использованием методов стеганографии, в частности цифровых водяных знаков, которые применяются в данных системах идентификации/аутентификации как одноразовые пароли для авторизации и доступа пользователя к защищаемым ресурсам. Также представлены основные блок-схемы реализации двух систем графического пароля, первая – на основе пиктограмм, вторая использует графический файл с множеством деталей. Опытным путем доказана эффективность применения приведенных систем паролей по отношению к привычным парольным методам аутентификации пользователей, в частности, пользователю не требуется запоминать сложные парольные последовательности – достаточно запомнить последовательность пиктограмм или последовательность нажатий на области графического файла, система самостоятельно меняет пароли без вмешательства пользователя в процесс, но по желанию пользователь самостоятельно может изменить последовательность пиктограмм или областей нажатий на графическом файле для входа в систему. Использование предложенных алгоритмов позволяет создать стойкие системы идентификации/аутентификации, а также уменьшить время запоминания паролей и авторизацию пользователей в системах, требующих ввода логина и пароля.

Ключевые слова:

Системы графических паролей, цифровой водяной знак, методы реализации систем графических паролей, стеганография, идентификация/аутентификация пользователей.

Пользователи имеют определенную трудность заучивания сложных, псевдослучайных паролей в течение определенного времени [1]. Большинство из них забывают пароль, который не используется регулярно, а также могут или смешать элементы различных паролей, или помнить пароль, но путать, какой системе он соответствует.

Пользователи часто уменьшают сложность и число символов в паролях, тем самым снижая безопасность систем для взлома. Безопасный пароль должен содержать не менее 8 символов, его желательно создавать генератором случайных последовательностей из символов с верхним регистром, символов с нижним регистром, цифр, а также использовать специальные символы. С такими паролями у людей возникает проблема в запоминании, и большинство пользователей игнорируют данные

рекомендации. Практика показывает, что пользователи часто выбирают короткие пароли, состоящие из имен, фамилий семьи или друзей, названия домашних животных, и даже не редко встречается слово «пароль». Чтобы не забывать пароли записывают их на бумагу либо используют тот же самый пароль для различных систем, иногда с единственной цифрой в конце [2, 3].

В связи с этими факторами и появлением мониторов и различных устройств с сенсорными экранами ведутся разработки систем графических паролей [4–6], которые создаются, чтобы избавить пользователя от сложных паролей и упростить авторизацию. Одним из недостатков этих систем является то, что большинство из них основаны на присвоении определенных символов изображению и/или координат нажатий, выбранных пользова-

телем для авторизации. Предлагаемые далее системы графических паролей на основе цифровых водяных знаков (ЦВЗ) избавлены от этого недостатка путем встраивания в графические файлы случайных символов, выработанных встроенным генератором случайных последовательностей. Пароли в приведенных системах являются одноразовыми. То есть после каждой успешной авторизации автоматически меняется последовательность символов цифровых водяных знаков, используемых в качестве паролей.

Предлагаемая модель системы ЦВЗ для разграничения доступа пользователей к защищенным ресурсам, применяемая в построении системы графического пароля [7, 8], показана на рис. 1.

Системой пользователю предлагается выбрать последовательность графических файлов, далее на все предложенные графические файлы система накладывает цифровой водяной знак W , индивидуальный для каждого графического объекта, который преобразовывается в кодере к удобному виду для встраивания в заверяемое сообщение. Алгоритм формирования такой конструкции водяного знака A представим в виде: $A=F(I, W)$, где F – функция, зависящая от I – контейнера (графический файл) и W – водяного знака [9, 10].

Затем в формирователе заверенных сообщений конструкция водяного знака A встраивается с помощью функции Z в графический контейнер, используя конфиденциальный ключ K : $Z=\Psi(A, I, K)$, где Ψ – функция, зависящая от A – конструкции водяного знака, I – контейнера (графического файла) и K – секретного ключа [11, 12].

После выбора пользователем последовательности графических объектов для своей аутентификации система передает ее по каналу связи. В канале связи на заверенное сообщение Y воздействуют нарушитель, а также случайные и преднамеренные помехи. В результате этого воздействия на приеме в устройство проверки водяных знаков поступает модифицированное сообщение Y . По алгоритму обнаружения водяного знака [13] формируется оценка водяного знака W' вида: $W'=G(Y, W, K)$, где G – функция с зависимостями от Y – модифицированное сообщение, W – водяной знак, K – секретный ключ.

Подлинность пользователя определяется в соответствии с этой оценкой [14]. Возможны решения вида $W'=1$ (подлинность сообщения подтверждена) или $W'=0$ (подлинность сообщения не подтверждена). Также возможны и другие решения вида $0,5 \leq W'_j \leq 1$ (j -й фрагмент, скорее всего, подлинный) или $0 \leq W'_j < 0,5$ (j -й фрагмент, скорее всего, навязан или искажен помехами передачи). При формировании оценки водяных знаков могут возникнуть ошибки их обнаружения получателем сообщения [15, 16].

По сравнению с криптографическими системами аутентификации система аутентификации пользователей на основе ЦВЗ имеет следующие особенности [9]:

- заверяемое сообщение и встроенный в него ЦВЗ взаимозависимы, то есть при разрушении первого разрушается и второй, а если водяной знак сохранил свою целостность, то и принятое сообщение ее не потеряло;
- при приеме искаженного фрагмента сообщения получатель может, не отказываясь от всего сообщения в целом, отказаться лишь от данного фрагмента.

В отличие от сравнительных методов методы контроля подлинности на основе водяных знаков обладают существенными достоинствами:

- высокой устойчивостью к удалению аутентификатора заверенного сообщения без разрушения самого сообщения;
- обнаружением несанкционированного копирования заверенных сообщений;
- согласованностью с источниками сообщений, обладающими существенными статистическими зависимостями и памятью, такими как изображение и звуковой сигнал.

Взяв во внимание все вышесказанное [10], можно получить несколько различных систем графического пароля, использующих стеганографические методы, которые повышают безопасность всей системы аутентификации. Первая из таких систем – система графического пароля на основе пиктограмм, вторая – система графического пароля на основе графического файла с множеством деталей. Блок-схемы работы сервера и клиентской части системы графического пароля на основе пик-

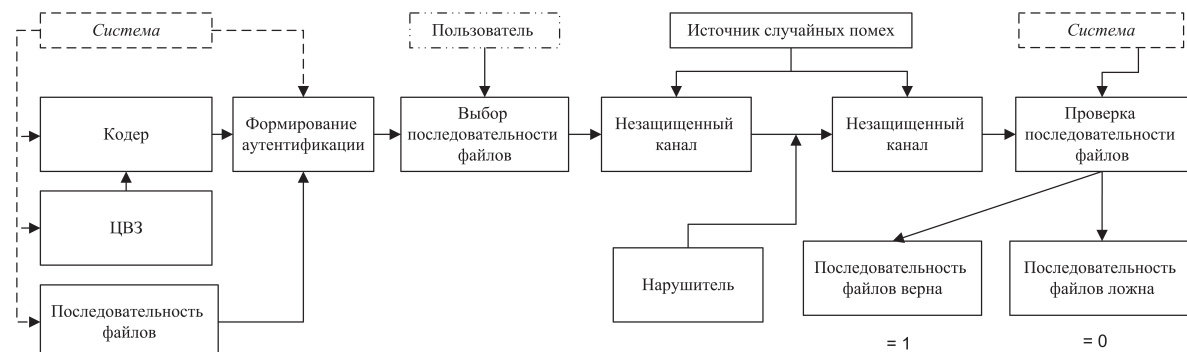


Рис. 1. Модель системы ЦВЗ для аутентификации

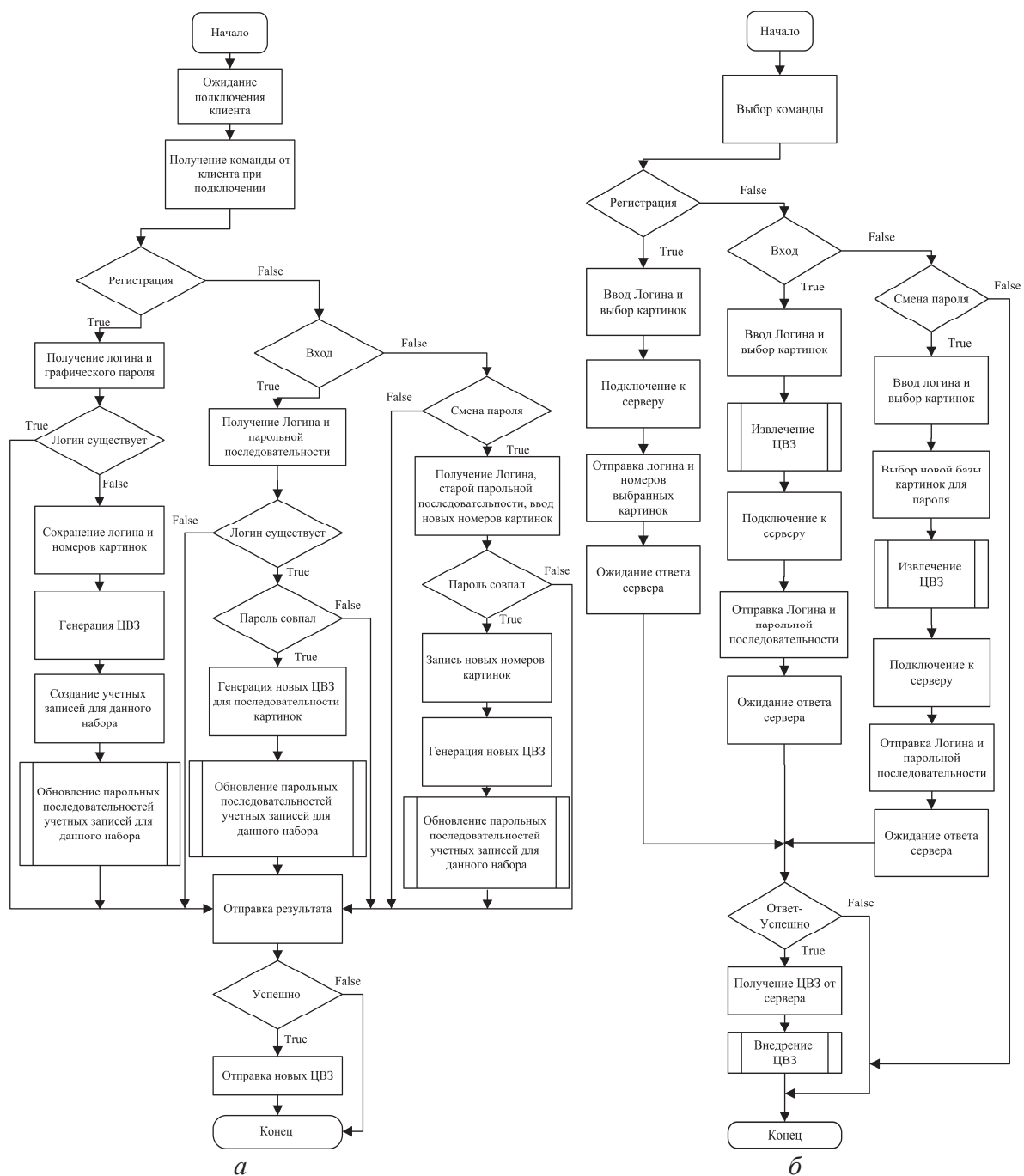


Рис. 2. Блок-схема алгоритма работы системы графического пароля на основе пиктограмм: а) сервер; б) клиент

тограмм при авторизации пользователя показаны на рис. 2.

Описание алгоритма работы сервера графического пароля на основе пиктограмм:

1. При запуске программы находится в режиме ожидания подключения клиента.
2. При подключении клиента к серверу сервер ожидает команду от клиента.
3. Действия на предъявленную клиентом команду.

Регистрация. Сервер считывает присланные клиентом логин и номера картинок. Далее сервер

сравнивает логин, предложенный клиентом, с имеющимися в базе данных учетными записями пользователей. Если такой логин уже существует, то сервер не регистрирует нового пользователя. Если учетной записи с таким логином не существует, то сервер заносит в базу данных новую запись с данным логином и предложенными номерами картинок:

- генерирует последовательность, которая будет являться ЦВЗ;
- выбирает список учетных записей, которые базируются на том же наборе картинок;

- последовательно проходит все эти записи в базе данных, меняя парольную последовательность на новую в соответствии со сгенерированными ЦВЗ и номерами картинок.

Авторизация. Сервер считывает присланные клиентом логин и парольную последовательность. Далее сервер ищет в базе данных аккаунтов учетную запись с таким логином, и если такая учетная запись существует, то сравнивает присланный пароль с хранящимся в базе данных. Если проверка прошла успешно, то сервер:

- генерирует последовательность, которая будет являться ЦВЗ;
- выбирает список учетных записей, которые базируются на том же наборе картинок;
- последовательно проходит все эти записи в базе данных, меняя парольную последовательность на новую в соответствии со сгенерированными ЦВЗ и номерами картинок.

Смена пароля. Сервер считывает присланные клиентом логин, парольную последовательность и номера картинок, соответствующих новому паролю. Далее сервер ищет в базе данных аккаунтов учетную запись с таким логином и, если такая учетная запись существует, сравнивает присланный пароль с хранящимся в базе данных. Если проверка прошла успешно, то сервер:

- заменяет в базе данных у данной учетной записи номера картинок на новые;
- генерирует последовательность, которая будет являться ЦВЗ;
- выбирает список учетных записей, которые базируются на том же наборе картинок;
- последовательно проходит все эти записи в базе данных, меняя парольную последовательность на новую в соответствии со сгенерированными ЦВЗ и номерами картинок.

4. Сервер отправляет клиенту ответ на его команду. Если запрашиваемая команда прошла успешно, то сервер отправляет сгенерированную на предыдущем этапе последовательность ЦВЗ. Алгоритм работы клиента заключается в следующем:

1. Клиент определяет команду для отправки серверу.
2. Действия клиента в соответствии с выбранной командой.

Регистрация. На данном этапе происходит ввод пользователем желаемого логина и картинок. Далее клиент подключается к серверу и передает ему введенный логин и номера картинок, после чего ожидает ответа сервера.

Авторизация. На клиенте вводится логин и выбираются картинки. По нажатию на соответствующую картинку формируется пароль на основе извлеченных из картинок скрытых данных. Клиент подключается к серверу и передает ему логин и пароль, после чего ожидает ответа от сервера.

Смена пароля. На клиенте вводится логин и формируется пароль на основе выбранных картинок. Далее клиент просит ввести новый пароль и

регистрирует уже номера введенных картинок. Клиент подключается к серверу и передает ему логин, пароль и номера картинок, после чего ожидает ответа от сервера.

3) При получении ответа от сервера регистрируется результат выполнения команды, и в случае успешного выполнения команды клиент принимает от сервера новую последовательность ЦВЗ. Клиент внедряет в используемый набор картинок присланную последовательность.

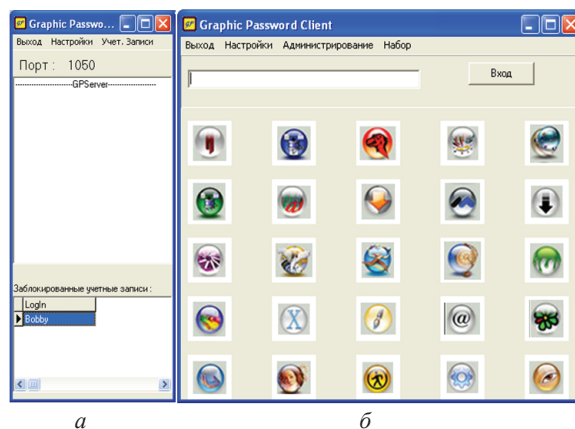


Рис. 3. Программа: а) сервер; б) программа-клиент перед началом работы

Внешний вид полученной системы графического пароля на основе пиктограмм представлен на рис. 3.

Следующей реализацией системы графического пароля является реализация с применением графического файла с множеством деталей.

Обобщенная блок-схема алгоритма работы представлена на рис. 4.

Описание алгоритма:

1. При запуске программа находится в режиме ожидания подключения клиента.
2. При подключении клиента к серверу сервер ожидает команду от клиента.
3. Действия на предъявленную клиентом команду.

Регистрация. Сервер считывает присланные клиентом логин и координаты по осям X и Y. Далее сервер сравнивает логин, предложенный клиентом, с имеющимися в базе данных учетными записями пользователей. Если такой логин уже существует, то сервер не регистрирует нового пользователя. Если учетной записи с таким логином не существует, то сервер

- заносит в базу данных новую запись с данным логином и последовательностями X и Y;
- генерирует последовательность, которая будет являться ЦВЗ;
- заносит сгенерированный ЦВЗ в базу данных.

Авторизация. Сервер считывает присланные клиентом логин и парольные последовательности X и Y. Далее сервер ищет в базе данных аккаунтов учетную запись с таким логином и, если такая учетная запись существует, сравнивает присланный пароль с

хранящимся в базе данных. В случае успешной проверки сервер выполняет следующие действия:

- генерирует последовательность, которая будет являться ЦВЗ;
- заменяет у данной учетной записи в базе данных строку, соответствующую ЦВЗ на сгенерированную.

Смена пароля. Сервер считывает присланные клиентом логин, старые и новые парольные последовательности X и Y. Далее сервер ищет в базе данных аккаунтов учетную запись с таким логином и

если такая учетная запись существует, то сравнивает присланный пароль с хранящимся в базе данных. Если проверка прошла успешно, то сервер:

- заменяет в базе данных у данной учетной записи старые последовательности, соответствующие X и Y, на новые;
- генерирует последовательность, которая будет являться ЦВЗ;
- заменяет у данной учетной записи в базе данных строку, соответствующую ЦВЗ, на сгенерированную.

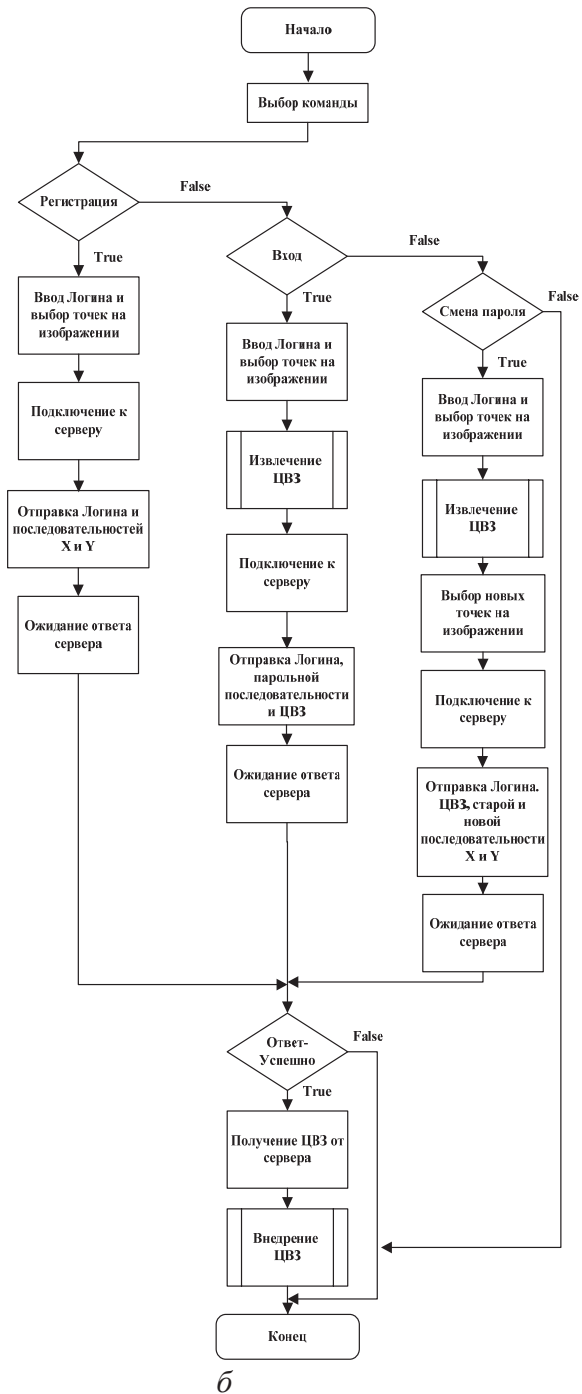
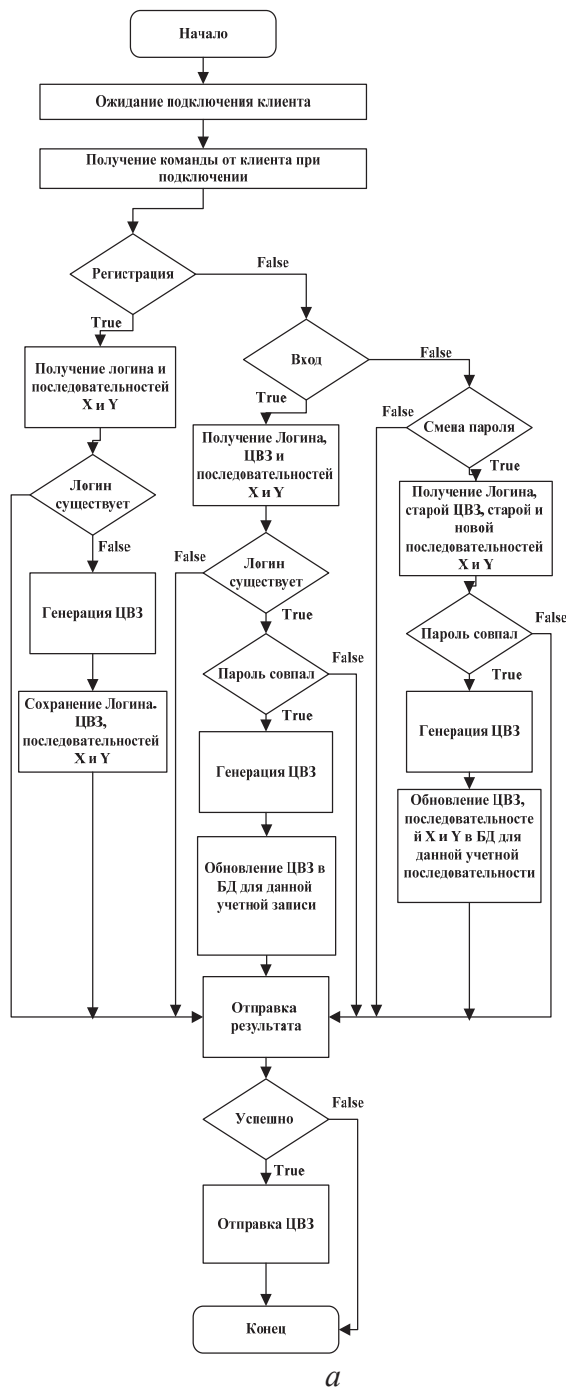


Рис. 4. Блок-схема алгоритма работы системы графического пароля на основе графического файла со множеством деталей: а) сервер; б) клиент

4. Сервер отправляет клиенту ответ на его команду. Если запрашиваемая команда прошла успешно, то сервер отправляет сгенерированную на предыдущем этапе ЦВЗ.

Данную версию графического пароля можно усилить путем применения алгоритмов теории распознавания образов. Так как в данной реализации используются картинки, то можно сравнивать эталон с предъявляемой картинкой по составляющим rgb каждого пикселя картинки. Тогда на третьем этапе работы сервера появятся дополнительные блоки, соответствующие сравнению rgb -составляющих, а в базе данных аккаунтов необходимо создать BLOB-поле для хранения эталона.

Алгоритм работы клиента:

1. Клиент определяет, какую команду необходимо отправить серверу.
2. Действия клиента в соответствии с выбранной командой.

Регистрация. На данном этапе происходит ввод пользователем желаемого логина и точек на изображении. Далее клиент подключается к серверу и передает ему введенный логин и последовательности X и Y , после чего ожидает ответа сервера.

Авторизация. На клиенте вводится логин и точки на изображении. По нажатию на соответствующую точку формируется двойной пароль на основе координат точек, указанных пользователем, также из изображения извлекаются скрытые данные. Клиент подключается к серверу и передает ему логин, пароль и ЦВЗ, после чего ожидает ответа от сервера.

Таблица. Данные, полученные опытным путем в реализациях систем графического пароля

Показатель	Пиктограммы	Изображение	Изображение с распознаванием образа
Мощность по графическим данным	минимальная: 421900 максимальная: $1,335 \cdot 10^{13}$	минимальная: $3,034 \cdot 10^{16}$ максимальная: $2,846 \cdot 10^{38}$	минимальная: $3,034 \cdot 10^{16}$ максимальная: $2,846 \cdot 10^{38}$
Мощность пространства паролей по ЦВЗ	минимальная: $7,602 \cdot 10^{17}$ максимальная: $5,275 \cdot 10^{41}$	$8,587 \cdot 10^{19}$	$8,587 \cdot 10^{19}$
Время авторизации (часы:минуты:секунды)	00:00:00,466	00:00:00,48	00:02:29,771
Среднее и максимальное время набора пароля в мин	00:00:02,505 00:00:06,199	00:00:03,129 00:00:06,029	00:00:03,129 00:00:06,029
Хранение пароля	64-битное кодирование		
Ограничения на количество попыток ввода пароля	3	3	3
Блокирование учетной записи	Да		

Смена пароля. На клиенте вводится логин и формируется пароль на основе указанных пользователем точек. Далее клиент просит ввести новый пароль и регистрирует новые точки на картинке.

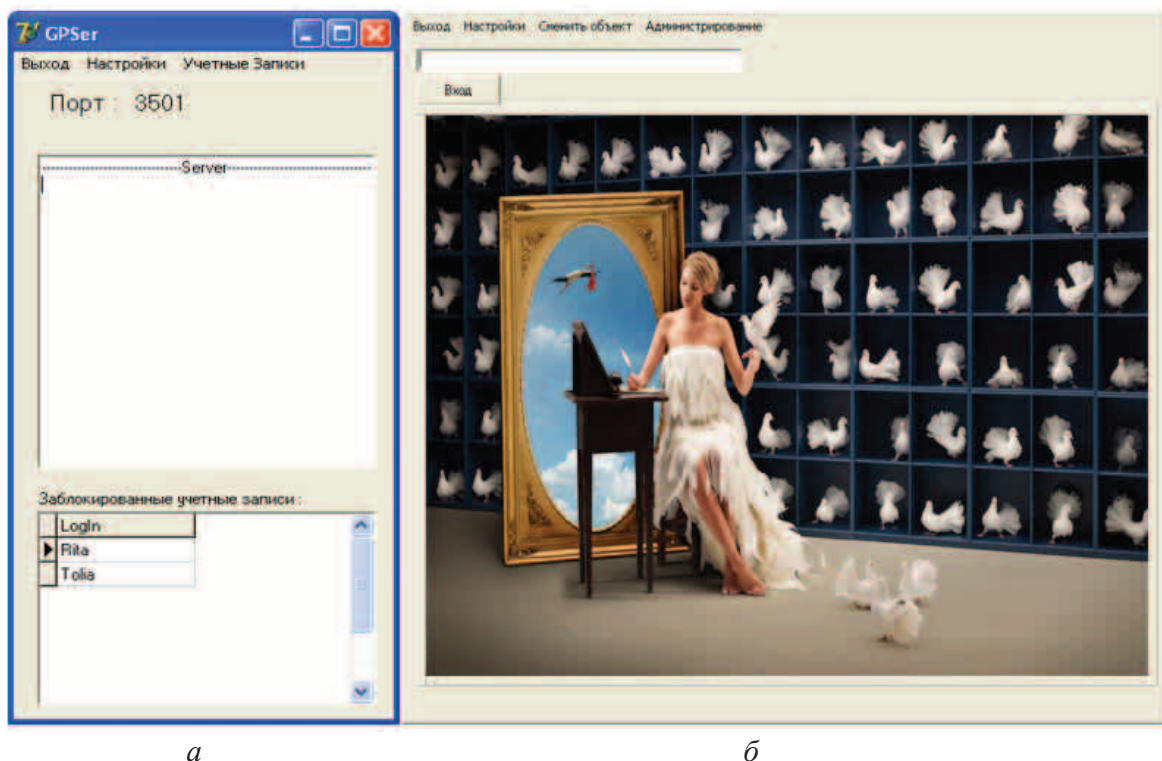


Рис. 5. Реализация системы графического пароля с применением графического файла с множеством деталей: а) сервер; б) программа-клиент

Клиент подключается к серверу и передает ему логин, ЦВЗ и парольные последовательности, после чего ожидает ответа от сервера;

3. При получении ответа от сервера регистрируется результат выполнения команды, и в случае успешного выполнения команды клиент принимает от сервера новую последовательность ЦВЗ. Клиент внедряет в используемое изображение полученную последовательность.

При усилении данной реализации графического пароля с применением распознавания образов на клиентской части алгоритм работы будет заключаться в загрузке изображения в память, последовательном переборе всех пикселей и отправке rgb-составляющих каждого пикселя на сервер.

При тестировании реализаций систем графического пароля были получены опытным путем данные, представленные в таблице.

В приведенных реализациях систем графических паролей использована защита пароля стеганографическим методом – встраивание ЦВЗ, который генерируется случайным образом из алфавита мощностью 97 символов латинского, русского алфавита, цифровых и специальных символов. Вне-

дряемые данные имеют небольшой объем, а предъявляемые к ним требования минимальны [17, 18]: заголовки вносят незначительные искажения и устойчивы к основным геометрическим преобразованиям; парольная последовательность, состоящая из набора ЦВЗ, хранится в базе данных аккаунтов в зашифрованном виде по алгоритму 64-битного кодирования, а для варианта графического пароля на основе изображения со множеством деталей зашифровываются также последовательности, соответствующие точкам входа пользователя [19].

Использование ЦВЗ в системах графических паролей показывает, что атаки на данные системы более сложные в реализации [20] и совершенно отличаются от атак на привычные и часто используемые символьные системы. Следовательно, системы графического пароля с использованием ЦВЗ имеют ряд преимуществ, таких как увеличение стойкости систем идентификации/аутентификации к взломам злоумышленниками, быстрое запоминание пароля пользователем, а также уменьшение времени авторизации в системах разграничения доступа.

СПИСОК ЛИТЕРАТУРЫ

1. Аутентификация. Теория и практика обеспечения безопасного доступа к информационным ресурсам / А.А. Афанасьев, Л.Т. Веденев, А.А. Воронцов и др. / под ред. А.А. Шелупанова, С.Л. Груздева, Ю.С. Нахаева. – М.: Горячая линия – Телеком, 2009. – 552 с.
2. Зегжда Д.П., Ивашко А.М. Основы безопасности информационных систем. – М.: Горячая линия – Телеком, 2000. – 452 с.
3. Основы информационной безопасности / А.А. Шелупанов, В.П. Лось, Р.В. Мещеряков, Е.Б. Белов. – М.: Горячая линия – Телеком, 2006. – 544 с.
4. Sobrado L., Birget J.C. Graphical passwords // The Rutgers Scholar, An Electronic Bulletin for Undergraduate Research. – 2002. – V. 4. – P. 152–158.
5. Davis D., Monroe F., Reiter M.K. On user choice in graphical password schemes. Thirteenth Usenix Security Symposium. – San Diego, CA, USA, Aug. 9–13, 2004. URL: <http://www.usenix.org/events/sec04/tech/davis.html> (дата обращения: 21.02.2006).
6. Authentication using graphical passwords: Effects of tolerance and image choice / S. Wiedenbeck, J. Waters, J.C. Birget, A. Broditskiy, N. Memon. URL: <http://clam.rutgers.edu/birget/grPsw/index.html> (дата обращения: 15.10.2006).
7. Шокарев А.В. Использование цифровых водяных знаков для аутентификации передаваемых сообщений // Вестник СибГАУ «Системная интеграция и безопасность». – 2006. – Спец. выпуск. – С. 123–127
8. Шокарев А.В., Шелупанов А.А. Использование компьютерной стеганографии для аутентификации пользователей // Научная сессия ТУСУР – 2006: Материалы докладов Всерос. научно-техн. конф. студентов, аспирантов и молодых ученых. – Томск: ТУСУР, 2006. – С. 173–176
9. Шокарев А.В., Шелупанов А.А. Аутентификация пользователей в защищенном документообороте на основе цифровых водяных знаков // Прогрессивные технологии и экономика в машиностроении: Труды IV Всерос. научно-практ. конф. – Томск: Изд. ТПУ, 2006. – Т. 2 – С. 20–22.
10. Шокарев А.В. Графические пароли с использованием методов стеганографии // Инновационные технологии и экономика в машиностроении: Труды VII Всерос. научно-практ. конф. с междунар. участием. – Томск: Изд-во Томского политехнического университета, 2009. – С. 293–299.
11. Стеганография, цифровые водяные знаки и стеганоанализ / А.В. Аграновский, А.В. Балакин, В.Г. Грибунин, С. Сапожников. – М.: Вузовская книга, 2009. – 220 с.
12. Fridrich J. Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes // Information Hiding: VI International Workshop, LNCS. – Berlin, Heidelberg: Springer-Verlag, 2004. – V. 3200. – P. 67–81.
13. Шелупанов А.А., Шокарев А.В. Теоретико-информационный и теоретико-сложностный подходы для оценки стойкости стеганографических систем // Вестник СибГАУ «Системная интеграция и безопасность». – Красноярск, 2006. – Спец. выпуск. – С. 121–123.
14. Cachin C. An information-theoretic model for steganography // Information Hiding. II International Workshop, LNCS. – Berlin, Heidelberg: Springer-Verlag, 1998. – V. 1525. – P. 306–318.
15. Ker A. General Framework for Structural Steganalysis of LSB Replacement // VII International Workshop on Information Hiding. – Berlin, Springer-Verlag, 2005. – V. 3727. – P. 296–311.
16. Барсуков В.С. Стеганографические технологии защиты документов, авторских прав и информации // Обзор специальной техники. – 2000. – № 2. – С. 31–40.
17. Craver S. On public-key steganography in the presence of an active warden // Information Hiding, II International Workshop. – Portland, Oregon: Springer, 1998. – V. 1525. – P. 355–368.
18. Грибунин В.Г., Оков И.Н., Туринцев И.В. Цифровая стеганография. – М.: Солон-Пресс, 2002. – 272 с.
19. Основы компьютерной стеганографии / А.В. Аграновский, П.Н. Девянин, Р.А. Хади, А.В. Черемушкин. – М.: Радио и связь, 2003. – 152 с.
20. Коханович Г.Ф., Пузыренко А.Ю. Компьютерная стеганография. Теория и практика. – К.: МК-Пресс, 2006. – 288 с.

Поступила 18.12.2013 г.

APPLICATION OF DIGITAL WATER MARKS FOR DEVELOPING GRAPHICAL PASSWORD SYSTEMS

Aleksy V. Shokarev,

Cand. Sc., Yurga Institute of Technology (Affiliate) of Tomsk Polytechnic University, 26, Leningradskaya street, Yurga, 652055, Russia.

E-mail: Shokarev_AV@mail.ru

The relevance of the work consists in increasing the security of protected resources with the passage of identification/authentication of users by means of graphic password, which can be achieved through the application of methods of steganography.

The main aim of the study is to improve the reliability of user identification/authentication applying a graphical password system using digital watermarking. Application of steganography methods increases system resistance to the majority of known methods of hacking systems with password authentication.

The methods used in the study: the methods of probability theory, mathematical analysis and experimental studies of steganography. Software implementation of the method was performed using a programming environment Borland C++.

The results: The paper introduces the possibility of creating a system of graphical passwords using steganography techniques, in particular digital watermarks, which are used in the systems of identification/authentication as one-time passwords for authentication and user access to protected resources. The author also introduces the basic block diagrams for implementing two graphical password systems, the first one is icon-based, the second uses an image file with many details. The author has proved empirically the effectiveness of applying the given password systems with respect to the common password authentication methods. A user just remembers the sequence of icons or sequence of keystrokes on a graphic file, but not the complex password sequences. The system changes automatically the passwords without user intervention in the process. But as an option a user can change the order of icons or areas of clicks on an image file to login. The use of the proposed algorithms allows you to create a stable system identification/authentication, and reduce time-remember passwords and authorization of users in systems that require login and password.

Key words:

Graphical password system, digital watermark, implementation techniques of graphical password systems, steganography, identification/authentication of users.

REFERENCES

1. Afanasev A.A., Vedenev L.T., Vorontsov A.A. *Autentifikatsiya. Teoriya i praktika obespecheniya bezopasnogo dostupa k informatsionnym resursam* [Authentication. Theory and practice of providing secure access to information resources]. Ed. A.A. Shelupanov, S.L. Gruzdev, Yu.S. Nakhaev. Moscow, Goryachaya liniya – Telekom Publ., 2009. 552 p.
2. Zegzhda D.P., Ivashko A.M. *Osnovy bezopasnosti informatsionnykh sistem* [Fundamentals of Information Systems Security]. Moscow, Goryachaya liniya – Telekom Publ., 2000. 452 p.
3. Shelupanov A.A., Los V.P., Meshcheryakov R.V., Belov E.B. *Osnovy informatsionnoy bezopasnosti* [Fundamentals of Information Security]. Moscow, Goryachaya liniya – Telekom Publ., 2006. 544 p.
4. Sobrado L., Birget J.C. Graphical passwords. *The Rutgers Scholar, an Electronic Bulletin for Undergraduate Research*, 2002, vol. 4, pp. 152–158.
5. Davis D., Monroe F., Reiter M.K. On user choice in graphical password schemes. *Thirteenth Usenix Security Symposium*. San Diego, CA, USA, Aug. 9–13, 2004. Available at: <http://www.usenix.org/events/sec04/tech/davis.html> (accessed 21 February 2006).
6. Wiedenbeck S., Waters J., Birget J.C., Broditskiy A., Memon N. *Authentication using graphical passwords: Effects of tolerance and image choice*. Available at: <http://clam.rutgers.edu/birget/grPssw/index.html> (accessed 15 October 2006).
7. Shokarev A.V. Ispolzovanie tsifrovyykh vodyanykh znakov dlya autentifikatsii peredavaemykh soobshcheniy [The use of digital watermarking for authentication of messages transmitted]. *Vestnik SibGAU. Sistemnaya integratsiya i bezopasnost – Herald of SibSAU. System integration and security*, 2006, Spec. Iss., pp. 123–127.
8. Shokarev A.V., Shelupanov A.A. Ispolzovanie kompyuternoy steganografii dlya autentifikatsii polzovateley [Using computer steganography for user authentication]. *Nauchnaya sessiya TUSUR. Materialy dokladov Vserossiyskoy nauchno-tekhnicheskoy konferentsii studentov, aspirantov i molodykh uchenykh* [Scientific session TUSUR-2006 Proceedings of the Russian scientific and technical conference of students and young scientists]. Tomsk, 2006, pp. 173–176.
9. Shokarev A.V., Shelupanov A.A. Autentifikatsiya polzovateley v zashchishchennom dokumentooborote na osnove tsifrovyykh vodyanykh znakov [User authentication in a secure workflow based on digital watermarking]. *Progressivnye tekhnologii i ekonomika v mashinostroenii. Trudy IV Vserossiyskoy nauchno-prakticheskoy konferentsii* [Progressive Technologies in Mechanical Engineering and Economics. Proc. IV All-Russian scientific-practical conference]. Tomsk, 2006, vol. 2, pp. 20–22.
10. Shokarev A.V. Graficheskie paroli s ispolzovaniem metodov steganografii [Graphical passwords using steganography methods]. *Innovatsionnye tekhnologii i ekonomika v mashinostroenii. VII Vserossiyskaya nauchno-prakticheskaya konferentsiya s mezhdunarodnym uchastiem* [Innovative technologies in mechanical engineering and economics. Proc. VII All-Russian scientific and practical conference with international participation]. Tomsk, 2009, pp. 293–299.
11. Agranovskiy A.V., Balakin A.V., Gribunin V.G., Sapozhnikov S. *Steganografiya, tsifrovye vodyanye znaki i steganoanaliz* [Steganography, digital watermarks, steganalysis]. Moscow, Vuzovskaya kniga Publ., 2009. 220 p.
12. Fridrich J. Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes. *Information Hiding. VI International Workshop, LNCS*. Berlin, Heidelberg, Springer-Verlag 2004. Vol. 3200, pp. 67–81.
13. Shelupanov A.A., Shokarev A.V. Teoretiko-informatsionny i teoretiko-slozhnostny podkhody dlya otsenki stoykosti steganograficheskikh sistem [Information-theoretic and complexity-theoretic approaches for assessing resistance of steganographic systems]. *Vestnik SibGAU. Sistemnaya integratsiya i bezopasnost – Herald of SibSAU. System integration and security*, 2006, Spec. Iss., pp. 121–123.
14. Cachin C. An information-theoretic model for steganography. *Information Hiding. II International Workshop, LNCS*. Berlin, Heidelberg, Springer-Verlag, 1998. Vol. 1525, pp. 306–318.

15. Ker A. General Framework for Structural Steganalysis of LSB Replacement. *VII International Workshop on Information Hiding*. Berlin, Springer-Verlag, 2005. Vol. 3727, pp. 296–311.
16. Barsukov V.S. Steganograficheskie tekhnologii zashchity dokumentov, avtorskikh prav i informatsii [Steganographic security technology for documents, copyrights and information]. *Obzor spetsialnoy tekhniki*, 2000, no. 2, pp. 31–40.
17. Craver S. On public-key steganography in the presence of an active warden. *Information Hiding, II International Workshop*. Portland, Oregon, 1998, vol. 1525, pp. 355–368.
18. Gribunin V.G., Okov I.N., Turintsev I.V. *Tsifrovaya steganografiya* [Digital steganography]. Moscow, Solon-Press Publ., 2002. 272 p.
19. Agranovskiy A.V., Devyanin P.N., Khadi R.A., Cheremushkin A.V. *Osnovy kompyuternoy steganografii* [Fundamentals of computer steganography]. Moscow, Radio i svyaz Publ., 2003. 152 p.
20. Kokhanovich G.F., Puzyrenko A.Yu. *Kompyuternaya steganografiya. Teoriya i praktika* [Computer steganography. Theory and practice]. Kiev, MK-Press Publ., 2006. 288 p.

УДК 004.4::004.85[3+5]

АВТОМАТИЧЕСКАЯ СИСТЕМА МЕТА-ОБУЧЕНИЯ С ПОДДЕРЖКОЙ ВЫБОРА ОПТИМАЛЬНОГО АЛГОРИТМА РЕШЕНИЯ ЗАДАЧИ И ВЫЧИСЛЕНИЯ ОПТИМАЛЬНЫХ ПАРАМЕТРОВ ЕГО ФУНКЦИОНИРОВАНИЯ

Орлов Андрей Александрович,

ассистент каф. промышленной электроники Томского Государственного
Университета Систем Управления и Радиоэлектроники,
634050, Россия, Томск, пр. Ленина, д. 40. Email: d1scnc@gmail.com

Актуальность исследования обусловлена необходимостью повышения эффективности работы автоматических систем интеллектуального анализа данных, основанных на мета-обучении.

Цель исследования состоит в разработке автоматической системы мета-обучения с поддержкой выбора оптимального алгоритма решения задачи и вычисления оптимальных параметров его функционирования.

Методы исследования: индуктивное моделирование, методы статистической обработки результатов.

В результате исследования проведена систематизация известных систем мета-обучения на основании выработанных классификационных признаков, учитывающих внутреннюю организацию систем. Сформулированы требования к реализации автоматической системы мета-обучения. Предложен способ построения системы мета-обучения, удовлетворяющей всем сформулированным требованиям и производящей накопление мета-знаний, построение на их основе мета-моделей, выбор оптимального алгоритма из набора доступных и вычисление оптимальных параметров его функционирования. Разработана объектно-ориентированная архитектура программной платформы для реализации любой из систем мета-обучения, представленных в систематизации. Эффективность реализованной автоматической системы мета-обучения с использованием алгоритмов методов группового учета аргументов проверена экспериментально при решении набора задач, относящихся к классу задач прогнозирования временных последовательностей (1428 временных последовательностей из тестового набора, известного под названием «M3 Competition»).

Ключевые слова:

Мета-обучение, мета-характеристики данных, мета-модель, программная платформа, объектно-ориентированный анализ и проектирование, прогнозирование временных последовательностей, метод группового учета аргументов.

Введение

Задачей Интеллектуального Анализа Данных (ИАД, в англоязычной литературе используется термин «Data Mining») является обнаружение (извлечение) в доступных исследователю исходных данных ранее неизвестных, неочевидных, но практически полезных знаний [1]. В настоящее время существует большое количество алгоритмов искусственного интеллекта (включая машинное обучение), математической статистики, оптимизации и прогнозирования и пр., применяемых для решения задачи ИАД: искусственные нейронные сети, генетические алгоритмы, деревья решений, алгоритмы нечеткой логики, корреляционный и регрессионный анализ и т. д. Каждый из существующих алгоритмов показал свою эффективность при решении разнообразных практических задач. Од-

нако в работе [2] было показано, что не существует единственного алгоритма, способного максимально эффективно решать задачу ИАД во всех возможных практических применениях, поэтому решение каждой новой практической задачи требует привлечения некоторых экспертных знаний для выбора наиболее подходящего алгоритма из числа доступных.

В работе [3] был формализован подход к проблеме выбора алгоритма (рис. 1): на основании набора мета-характеристик (meta-features, MF) $f(x) \in F$ (F – пространство мета-характеристик) для выборки данных x из пространства проблем (задач) X функция $S(f(x))$ («selection mapping») производит выбор такого алгоритма a из пространства доступных алгоритмов A таким образом, что его эффективность $p(a, x)$ («performance mapping») на

выборке данных x максимальна ($p(a, x)$ – показатель эффективности, ПЭ).

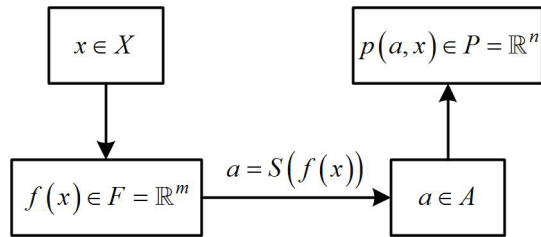


Рис. 1. Выбор алгоритма на основании мета-характеристик данных

Подход, предложенный в работе [3], предполагал, что выполнение функции «selection mapping» возложено на эксперта в области ИАД, либо в качестве $S(f(x))$ используется составленный экспертом набор правил (эвристик) выбора алгоритма. Очевидным недостатком применения подхода являлась необходимость привлечения экспертов на стадии разработки системы ИАД либо на стадии ее использования. С одной стороны это требовало дополнительных материальных и временных затрат, а с другой стороны ограничивало применение данного подхода в автоматических системах ИАД. Более того, возникала проблема эффективности накопления экспертных знаний и улучшения имеющихся эвристик выбора алгоритма, поскольку данные процессы основывались на увеличении уровня квалификации самого эксперта.

Решение описанных проблем производилось исследователями в рамках направления «мета-обучение», области ИАД, изучающей способы увеличения эффективности «обучающихся систем» ИАД путем накопления информации об эффективности обучения в таких системах (показатели эффективности обучения, их взаимосвязь с характеристиками задачи, особенности функционирования алгоритмов решения задачи и т. д.), в том числе и в автоматическом режиме [4].

Обзор существующих систем мета-обучения

До настоящего времени было предложено множество вариантов построения систем мета-обучения – например, в аналитических обзорах каждой из 10 глав только одной книги [5], посвященной мета-обучению, упомянуты более 40 систем, разрабатываемых различными научными коллективами. Несмотря на то, что в публикациях по системам мета-обучения приводится аналитический обзор и сравнение с некоторыми существующими системами, выделяются лишь некоторые внешние критерии оценки систем, как то результирующая точность, скорость принятия системой решений и т. д. Известна работа [5. С. 117–156], в которой сделана попытка описания строения систем мета-обучения, однако рассматривается лишь состав внутренних модулей без выделения общности в их организации.

В связи с этим автором было принято решение о разработке классификационных признаков, учи-

тывающих внутреннюю организацию систем, с целью систематизации существующих систем мета-обучения. Ниже описаны сформулированные классификационные признаки, согласно которым в табл. 1, 2 приводится информация для каждой из систем.

1. **Класс** разделяет приведенные системы на 4 группы по их предназначению:

- a. **Algorithm Selection** – для заданной выборки данных произвести выбор оптимального алгоритма из набора доступных (согласно задаваемому ПЭ);
- b. **Parameter Tuning** – для заданной выборки данных определить оптимальные значения параметров функционирования заданного алгоритма;
- c. **Data Mining Assistant** – выработать экспертные рекомендации, содействующие специалисту по ИАД в выборе оптимального алгоритма для заданной выборки данных;
- d. **Multi-Level Data Mining** – организовать процесс автоматического многоуровневого ИАД, при котором осуществляется не только перебор алгоритмов ИАД и их параметров, но и сравнение и выбор алгоритмов поиска оптимальных алгоритмов ИАД и оптимизации их параметров.

2. **Режим** распределяет ответственность за осуществление той или иной функции, системы мета-обучения: «Р» – ручной, когда принятие решений полностью возложено на эксперта; «А» – автоматический, когда эксперт уведомляется о решениях, принимаемых системой мета-обучения. Знаком «–» отмечены случаи, когда функция не поддерживается системой. В табл. 1, 2 рассматриваются следующие функции систем мета-обучения:

- a. **обучение** – способность системы производить накопление мета-знаний об эффективности доступных алгоритмов ИАД, мета-характеристиках данных и их взаимосвязи (в терминах работы [3]: определение вида и/или построение функции «selection mapping» S);
- b. **использование** – осуществление действий по выбору/рекомендации оптимального алгоритма ИАД и/или параметров его функционирования на основе анализа поставленной задачи – анализа заданной выборки данных (в терминах работы [3] для выбора алгоритма: вычисление мета-характеристик данных f и применение S для выбора алгоритма a);
- c. **выбор мета-характеристик** (значимых) – определение подмножества заданных мета-характеристик, отражающих лишь значимые свойства выборок данных, определяющих корректный выбор алгоритма для решения поставленной задачи. Иными словами, это исключение нерепрезентативных мета-характеристик;

- d. **настройка параметров алгоритмов** – определение таких значений параметров функционирования алгоритмов, при которых достигается максимизация ПЭ при решении поставленной задачи;
 - e. **проверка эффективности** (лучшего алгоритма) – оценка допустимости применения выбранного системой алгоритма для решения поставленной задачи. Это необходимо для детектирования случаев, когда все доступные системе мета-обучения алгоритмы неудовлетворительно решают поставленную задачу, поэтому выбор лучшего из них не гарантирует ожидаемое высокое значение ПЭ.
3. **Вид мета-характеристик** определяет набор групп мета-характеристик, используемых системой для анализа различных свойств заданной выборки данных:
- a. **КОЛ** – количественные: длина выборки, число переменных, число классов в задаче классификации и т. п.;
 - b. **ОБЛ** – область происхождения выборки данных (финансы, промышленность, информационные технологии и т. д.);
 - c. **СТАТ** – статистические: среднее, минимальное и максимальное значения, дисперсия, корреляция и т. п.;
 - d. **ГЕОМ** – (для задач классификации) – характеристики пересечений классов по входным данным, делимость, диаметры классов и расстояния между классами в пространстве параметров и т. д.;
 - e. **ПОВ** – поведение временной последовательности: характеристики тренда, периодичность, автокорреляция, число точек изменения направления и т. п.;
 - f. **ЭКСП** – экспертные знания: особенности области применения, описание входных и выходных данных и т. д.;
 - g. **ХАОС** – показатель Ляпунова, размерность фрактального пространства и т. д.;
 - h. **ОРИ** – ориентирование, landmarking – запуск относительно быстрых и простых алгоритмов ИАД, значение ПЭ которых есть мета-характеристика данных;
 - i. **ГРАФ** – характеристики графов: число вершин, ребер, показатели связности, оценка решения сверху и снизу, статистические показатели и др.;
 - j. **ШАБ** – (для алгоритмов) – содержание в алгоритме ИАД определенных структурных шаблонов обработки данных.
4. **Показатель эффективности алгоритма и системы** указывает способ расчета ПЭ отдельных доступных алгоритмов и системы в целом, включая известные критерии качества решения задач ИАД.
5. **Содержание базы мета-данных** иллюстрирует способ представления мета-знаний в виде хранимой информации для последующего использования.
6. **Мета-модель** содержит информацию о входных и выходных величинах мета-модели (разделены знаком «->»).
7. **Способы получения оптимального алгоритма и настройки оптимальных параметров алгоритма** определяет процедуру поиска оптимального алгоритма из доступных и/или вычисления оптимальных параметров его функционирования: **ВЫБ** – выбор из фиксированного множества, **ПЕР** – перебор элементов вычисляемого набора значений, **ОПТ** – решение задачи оптимизации.
8. **Задача** – краткая характеристика практического применения, на котором разработчиками показана эффективность системы мета-обучения. Помимо количественных характеристик указан вид задачи: **КС** – классификация, **КЗ** – кластеризация, **РЕ** – регрессия, **ПР** – прогнозирование, **РГ** – раскраска графов, **УК** – задача об упаковке в контейнеры (Bin Packing Problem, BPP).
- Следует отметить, что в табл. 1, 2 представлены системы мета-обучения, действующие с применением возможностей вычислительной техники. Аналитический обзор других систем мета-обучения (например, экспертных систем мета-обучения, систем построения онтологий и пр.) был выполнен в работе [5. С. 117–156]. Знаком «?» и наклонной штриховкой отмечены ячейки, информации по которым недостаточно.
- На основании анализа табл. 1, 2 можно сделать следующие выводы.
1. Большинство систем мета-обучения (17 из 18) поддерживают **автоматический режим обучения**, что избавляет от необходимости трудоемкого ручного формирования базы мета-знаний. Также большинство систем, относящихся к классам Algorithm Selection и Multi-Level Data Mining, поддерживают **автоматический режим использования** (11 из 12). Это позволяет включать данные системы в состав автоматических систем ИАД.
 2. Используются 10 видов мета-характеристик данных, причем наиболее часто в задаче классификации используются виды КОЛ и СТАТ (соответственно, 7 и 5 из 11), а в задаче прогнозирования – ПОВ и СТАТ (3 и 2 из 3). Это подтверждает, что репрезентативность различных мета-характеристик данных определяется видом задачи, решаемой системой мета-обучения, и подчеркивает актуальность выбора значимых мета-характеристик. Однако **автоматический режим выбора мета-характеристик** поддерживается лишь в 4 системах из 18.
 3. 11 систем из 18 осуществляют формирование мета-модели на основе накопленных при обучении мета-знаний, причем только 3 системы строят модели, позволяющие **прогнозировать значения ПЭ** на основании мета-характеристик заданной выборки данных, а значит, и производить **оценку качества решения**.

4. 8 систем из 18 позволяют производить **вычисление оптимальных параметров** функционирования алгоритмов, причем 6 из них – **в автоматическом режиме**. При этом только 3 системы из 6 основываются на **анализе свойств заданной выборки** данных (на расчете мета-характеристик). Остальные системы опираются на значение внешнего ПЭ, что увеличивает вероятность получения смещенных оценок оптимальных параметров функционирования алгоритмов (в том числе вследствие явления «переобучения» – «overfitting»).
5. Только одна система оценивает эффективность полученного оптимального алгоритма и допустимость его применения для решения поставленной задачи, остальные системы осуществляют выбор и/или генерируют рекомендации безотносительно предполагаемой эффективности работы выбранного алгоритма (качества решения задачи). Это ограничивает применение данных систем в отсутствии контроля со стороны эксперта.

На основании данных выводов сформулируем **требования к реализации автоматической системы мета-обучения**, лишенной упомянутых недостатков.

Требование А: Поддержка автоматических режимов обучения, использования, выбора значимых мета-характеристик, настройки параметров алгоритмов и проверки эффективности.

Требование В: Накопление мета-знаний и построение на их основе мета-моделей, отражающих закономерности в процессах обучения системы и алгоритмов.

Требование С: Расчет оптимальных параметров функционирования и выбор оптимального алгоритма на основании анализа свойств заданной выборки данных (мета-характеристик) и предсказания производительности алгоритмов.

Требование Д: Поддержка принятия решений о допустимости применения найденного оптимального алгоритма и его оптимальных параметров для решения поставленной задачи.

В табл. 1, 2 для каждой существующей системы проведен анализ ее соответствия сформулированным требованиям, причем степень соответствия с точки зрения возможности и трудоемкости доработки каждой системы отмечена соответствующим цветом ячейки (примечание к табл. 2). Из анализа следует, что доработка каждой существующей системы является либо трудоемкой, либо приводящей к существенному изменению ее структуры. Таким образом, разработка системы мета-обучения, удовлетворяющей всем сформулированным требованиям, является актуальной.

Организация системы мета-обучения

С целью построения системы мета-обучения, удовлетворяющей сформулированным выше требованиям, в работе [21] автором был предложен способ, представленный на рис. 2.

Выбор алгоритма a из пространства доступных алгоритмов A и набора его оптимальных параметров функционирования q^a из пространства параметров Q^a производится функцией $L(f(x))$ на основании набора мета-характеристик $f(x) \in F$ (F – пространство мета-характеристик) для выборки данных x из пространства проблем X таким образом, что эффективность $p(a(q^a), x)$ алгоритма a на выборке данных x максимальна ($p(a(q^a), x)$ – ПЭ).

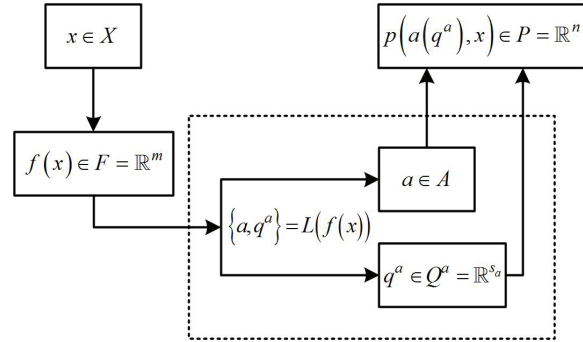


Рис. 2. Выбор алгоритма и параметров его функционирования на основании мета-характеристик данных

Рассмотрим подробнее процесс обучения предлагаемой системы мета-обучения.

Пусть система располагает информацией о некотором наборе из K выборок данных x_k , где $k=1, \bar{K}$. Для каждой выборки данных значение мета-характеристик рассчитывается как $f_k = f(x_k)$, причем $f_k = [f_k^1, f_k^2, \dots, f_k^m]$ – вектор из m мета-характеристик данных (m – размерность F).

Системе доступен набор из G алгоритмов a_g , где $g=1, \bar{G}$.

Система также обладает информацией об эффективности каждого из доступных ей алгоритмов a_g : $q_{i,j}^g$, где $i \in I_g \subseteq I = \{1, 2, \dots, G\}$ – информация о производительности алгоритма a_g может быть доступна не для всех выборок x_k ; $j=1, J_i^g$, J_i^g – количество наборов параметров $q_{i,j}^g$ алгоритма a_g , для которых доступно значение показателя эффективности на выборке x_i . При этом каждый набор параметров $q_{i,j}^g = [q_{i,j}^{g,1}, q_{i,j}^{g,2}, \dots, q_{i,j}^{g,s_g}]$ – вектор из s_g параметров алгоритма a_g (разные алгоритмы могут иметь различное число параметров s_g). Составим матрицу мета-характеристик M^F , матрицы параметров алгоритма $M_{g,i}^Q$ и матрицы показателей эффективности $M_{g,i}^P$:

$$M^F = \begin{bmatrix} f_1^1 & \dots & f_1^m \\ \dots & \dots & \dots \\ f_K^1 & \dots & f_K^m \end{bmatrix}, \quad M_{g,i}^Q = \begin{bmatrix} q_1^1 & \dots & q_1^{s_g} \\ \dots & \dots & \dots \\ q_{J_i^g}^1 & \dots & q_{J_i^g}^{s_g} \end{bmatrix},$$

$$M_{g,i}^P = \begin{bmatrix} p_1^1 & \dots & p_1^n \\ \dots & \dots & \dots \\ p_{J_i^g}^1 & \dots & p_{J_i^g}^n \end{bmatrix}.$$

С учетом введенных обозначений для каждого алгоритма a_g и выборки данных x_i составим матрицы доступной мета-информации для выборки данных x_i :

Таблица 1. Обзор систем мета-обучения. Класс «Algorithm Selection»

Класс системы	Система мета-обучения	Мета-обучение											Задача			
		Режим обучения		Режим использования	Вид мета-характеристик	Режим выбора MF	Показатель эффективности алгоритма	Содержание базы мета-данных (содержание записей)	Мета-модель	Способ получения оптимального алгоритма	Режим настройки параметров алгоритмов	Способ настройки параметров алгоритмов	Показатель эффективности системы	Режим проверки эффективности	Количество выборок данных	Количество MF
Algorithm Selection	NOEMON Prudencio, Ludermir [6]	A	A	ПОВ	P	MAE	MF,Algo -> MAE	MF -> Binary Compare (MLP)	ВЫБ Ранжи-рование	-	- (def)	SRC	P	ПП		
														645	5	3
	Prudencio, Ludermir [7]	A	A	СТАТ	P	?	MF -> вектор рангов алгоритмов	MF -> Binary Compare (MLP)	ВЫБ Ранжи-рование	-	-	Average Error Rate	P	PE		
														50	10	?
	Romero, Olmo, Ventura [8]	A	A	КОЛ, сложности, ОБЛ	P	5 ПЭ: Sen, Prec, F-M, Кар, AUC	MF -> вектор рангов алгоритмов	-	Nearest Neighbour, Euclidean distance on MF	-	- (def)	F-Measure (precision, recall)	P	KC		
														32	16	19
	Matijas, Suykens, Krajcar [9]	A	A	СТАТ, ПОВ	A	MASE + NRMSE	MF -> вектор рангов алгоритмов	MF->rank (объединение 7 алгоритмов классификации)	ВЫБ Ранжи-рование	-	- (def)	MASE NRMSE MAPE	P	ПП		
														65 + 4	13	7
	Cacoveanu, Vidrighin, Potolea [10]	A	A	КОЛ СТАТ ОРИ	P	CA	MF,Algo -> CA	-	ВЫБ наиболее похожих случаев (kNN)	-	-	CA	P	KC		
														13 + 2	11	7
	Cruz-reyes et. al. [11]	A	A	СТАТ	P	Число контейнеров	MF,Algo -> Group/algo (kNN)	MF->Group (Discriminant Analysis, C4.5, SOM)	ВЫБ Применение мета-модели – классификатора	-	-	Число контейнеров	P	УК		
														243 0	5	7
	Wang, Smith-Miles, Hyndman [12]	A	A	СТАТ ПОВ ХАОС	P	?(FA)	MF,Algo -> ranking	MF->ranking (SOM)	ВЫБ (DT)	-	-	?(FA)	A	ПП		
														315	13	4
	Abdelmessih, Shafait, Reif, Goldstein [13]	A	A	КОЛ СТАТ ОРИ	P	CA	MF -> CA	MF -> CA (SVM, RMSE)	ВЫБ (на основе прогноза ПЭ)	A	ПЕР (GS)	-(RMSE для мета-моделей)	P	KC		
														90	7	7
	Molina, Luna, Romero, Ventura [14]	A	P	КОЛ	P	CA	MF, Params -> Variation Accuracy	MF+Params -> Variation Accuracy (DT)	-	P	ПЕР	Variation of Accuracy	P	KC		
														14	3	1
	Musliu, Schwengerer [15]	A	A	ГРАФ	P	Минимальное число цветов	MF -> Число цветов	MF -> ranking (6 AK)	ВЫБ (применение AK)	-	- (def)	SSR	P	ПГ		
														859	8 + 82	6

Таблица 2. Обзор систем мета-обучения. Классы «Parameter Tuning», «Data Mining Assistant» и «Multi-level Data Mining»

Класс системы	Система мета-обучения	Мета-обучение												Задача		
		Режим обучения	Режим использования	Вид мета-характеристик	Режим выбора MF	Показатель эффективности алгоритма	Содержание базы мета-данных (содержание записей)	Мета-модель	Способ получения оптимального алгоритма	Режим настройки параметров алгоритмов	Способ настройки параметров алгоритмов	Показатель эффективности системы	Режим проверки эффективности	Количество выборок данных	Количество MF	Количество алгоритмов
Parameter Tuning	Konen et al. [16]	A	P	–	A	WCE	–	–	–	A	ОПТ	WCE	P	KC 3 - 1		
	Soares, Bradzil [17]	A	P	КОЛ СТАТ	P	NMSE	MF, Params -> NMSE	MF -> ranking (SOM)	–	A	ПЕР (kNN)	SRC	P	PE 42 14 + 3 1		
	Reif, Shafait, Dengel [18]	A	A	КОЛ СТАТ ОРИ	A	CA	MF -> Optimal Params (GS)	–	–	A	ПЕР (CBR) + ОПТ (GA)	CA	P	KC 102 14 1 + 1		
Data Mining Assistant	Sleeman, Rissakis et al. [19]	P	P	КОЛ ЭКСП	P	CA; Calcula- tion time; complexi- ty etc.	MF -> ПЭ	набор пра- вил выбора алгоритма	ВЫБ согласно набору правил	P	ОПТ (диа- ло- говый ре- жим)	ПЕР	P	KC 1 7 10		
	Kalousis, Hilario [20]	A	A	СТАТ	P	CE	MF, Algo -> Class.Error	–	ВЫБ наиболее похожих (CBR)	–	– (def)	Strict Error Loose Error	P	KC/[K3]/ [PE] 103 - 10		
	Hilario, Nguyen, Do, Woznica, Kalousis [5. C. 179–224]	A	P	Данные: СТАТ ГЕОМ Алго- ритм: ШАБ	A	CA	Data MF, Algo MF -> CA	Data MF, Algo MF -> CA	ВЫБ (на осно- ве про- гноза ПЭ)	–	–	CA	P	KC 65 18 7		
Multi-level Data Mining	Kordik, Cerny [5. C. 179–224]	A	A	–	–	CA RMSE	–	–	ВЫБ Лучший по ПЭ	A	ОПТ (GA)	CA RMSE	P	KC / PE 15 / 18 - 1*		
	Jankowski, Grabczewski [5. C. 1–76]	A	A	КОЛ	–	CA	–	–	ВЫБ Лучший по ПЭ	A	ПЕР ОПТ	CA	P	KC 8 - 54		
Цвет	Соответствие сформулированным требованиям															
X	Соответствует. Доработка не требуется															
X	Не соответствует. Доработка возможна и не приведет к изменению структуры системы															
X	Не соответствует. Доработка либо чрезвычайно трудоемкая либо приведет к существенному изменению структуры системы															

MAE – Mean Absolute Error, RMSE – Root Mean Square Error, NMSE – Normalized Mean Square Error, CA – Classification Accuracy (процент верных классификаций), CE – Classification Error (процент ошибок классификации), WCE – Weighted Classification Error (взвешенная ошибка классификации), FA – Forecasting Accuracy (точность прогноза), SRC – Spearman Rank Correlation (мера совпадения с идеальным вариантом ранжирования), SSR – Selection Success Rate (процент случаев верного выбора лучшего алгоритма), Algo – алгоритм, Params – параметры функционирования алгоритма, AK – алгоритм классификации, GS – Grid Search (перебор всех значений в заданных пределах с заданным шагом), MLP – Multi-layer perceptron, kNN – k-nearest-neighbor algorithm (метод k ближайших соседей), SOM – Self-Organizing Map (самоорганизующаяся карта), SVM – Support Vector Machine (метод опорных векторов), CBR – Case-Based Reasoning (основанный на правилах выбор наиболее похожего случая), GA – Genetic Algorithm (генетический алгоритм), DT – Decision Tree (дерево решений).

$$M_{g,j}^{left} = [M_i^F | M_{g,i}^Q] = \begin{bmatrix} f_i^1 & \dots & f_i^m & q_1^1 & \dots & q_1^{s_g} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ f_i^1 & \dots & f_i^m & q_{j_i^g}^1 & \dots & q_{j_i^g}^{s_g} \end{bmatrix}$$

$$M_{g,j}^{right} = M_{g,j}^P = \begin{bmatrix} P_1^1 & \dots & P_1^n \\ \dots & \dots & \dots \\ P_{j_i^g}^1 & \dots & P_{j_i^g}^n \end{bmatrix}.$$

Объединение данных матриц для всех $i \in I_g$ для каждого алгоритма a_g позволит получить матрицы мета-информации M_g^{left} и M_g^{right} , составляющие выборку данных для обучения мета-моделей вида $p_g = \mu^g(f, q)$, способных прогнозировать значение ПЭ для алгоритма a_g при заданных значениях мета-характеристик f и параметрах алгоритма q . Мета-модели $\mu^g(f, q)$ являются решением задачи идентификации вида

$$\mu^g = \arg \min_{\mu \in \Omega} Z(M_g^P, \mu(M^F, M_g^Q)) =$$

$$= \arg \min_{\mu \in \Omega} Z(M_g^{right}, \mu(M_g^{left})),$$

где Ω – множество всех мета-моделей; $Z(M^{true}, M^\mu)$ – критерий оценки качества мета-моделей.

Таким образом, в общем виде функция $L(f)$ может быть записана следующим образом:

$$\{a, q\} = L(f) =$$

$$= \left\{ \begin{array}{l} a^{g_{opt}}, q^{g_{opt}} \mid g_{opt} = \arg \max_{g=1..G} \mu^g(f, q_g^{opt}), \\ q_g^{opt} = \arg \max_{q_g \in R^{s_g}} \mu^g(f, q_g) \end{array} \right\}.$$

В работах [18] и [22] было указано, что большой размер пространства поиска R^{s_g} параметров q_g^{opt} при наличии обычно небольшого объема доступной мета-информации (J_i^g записей для каждой пары алгоритма a_g и выборки данных x_i) приводит к ухудшению качества прогнозирующих мета-моделей. Одним из возможных решений данной проблемы является обучение дополнительных мета-моделей вида $q_g^{opt} = \eta^g(f)$, производящих расчет оптимальных параметров для каждого алгоритма a_g на основании значений мета-характеристик данных f :

$$\eta^g = \arg \min_{\eta \in \Theta} Y(M_g^{Q_{opt}}, \eta(M^F)),$$

где $Y(M^{true}, M^\eta)$ – критерий оценки качества мета-модели, причем

$$M_{g,i}^{Q_{opt}} = \left\{ M_{g,i,j_{opt}}^{Q_{opt}} \mid j_{opt} = \arg \max_{j=1..J_i^g} p_i^g \right\}.$$

Таким образом, функция примет вид:

$$\{a, q\} = L^\eta(f) =$$

$$= \left\{ a^{g_{opt}}, q^{g_{opt}} \mid g_{opt} = \arg \min_{g=1..G} \mu^g(f, q_g^{opt}), q_g^{opt} = \eta^g(f) \right\}.$$

Схема, соответствующая данному модифицированному способу (в сравнении с предложенным на рис. 2), представлена на рис. 3.

Программная платформа для реализации системы мета-обучения

Архитектура программной платформы для реализации систем мета-обучения, полученная в результате объектно-ориентированного анализа и проектирования, представлена на рис. 4.

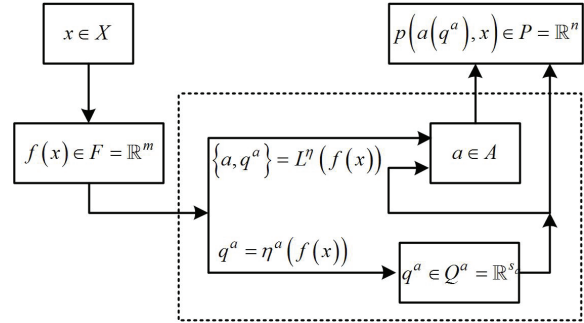


Рис. 3. Выбор алгоритма и параметров его функционирования на основании мета-характеристик данных на основе двух видов мета-моделей

Алгоритм (Algorithm) – объекты данного класса осуществляют некоторую последовательность действий по решению поставленной задачи (метод solve()) с целью получения результата, выраженного абстракцией Result. Исходная информация к решению задачи передается в виде выборки данных Data, а семантика данных определяется условием конкретной решаемой задачи. Работа алгоритма регулируется набором параметров функционирования params (q^a), специфичных для каждого алгоритма (a).

Ответственностью класса Criterion является оценка качества решения задачи на основании имеющихся результатов и исходных данных. В соответствии с введенными ранее обозначениями, оценка представляет собой вектор числовых значений p . MetaFeature – интерфейс расчета мета-характеристик (f^m) для задаваемой выборки данных (x).

Модель (Model) – это сущность, отражающая некоторую взаимосвязь или закономерность в данных (Data). Структура и параметры модели определяют ее уникальность среди других моделей. Модель может являться результатом работы некоторого алгоритма (ModelAlgorithm на рис. 4) по решению им поставленной задачи (например, задачи структурной и параметрической идентификации системы, классификации, прогнозирования и пр.).

Ответственностью класса MetaLearner является организация процесса мета-обучения. Объекты класса MetaLearner имеют доступ к некоторому задаваемому набору алгоритмов (algorithms), которые могут быть использованы для решения (solve()) поставленной задачи, а также к набору критериев оценки качества решения задачи (criteria). В процессе работы производится накопление мета-данных (metadata), которые используются для построения мета-моделей (MetaModel, наследованный от Model) с использованием задаваемой реализации интерфейса алгоритмов построения моделей (ModelAlgorithm). Полученные мета-мо-

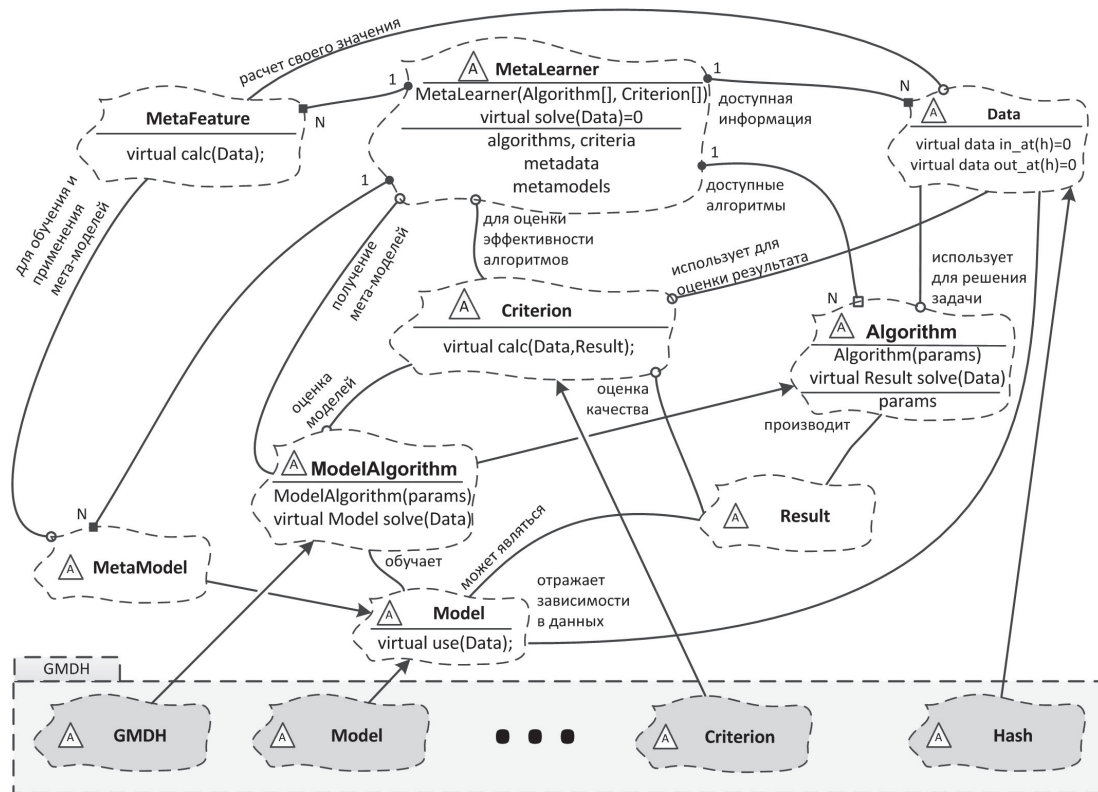


Рис. 4. Объектно-ориентированная архитектура программной платформы для реализации системы мета-обучения

дели (metamodels) используются для выбора оптимального алгоритма и набора оптимальных параметров его функционирования для решения поставленной задачи.

Предлагаемая объектно-ориентированная архитектура системы мета-обучения была реализована в рамках единой системы с авторской программной платформой для реализации алгоритмов метода группового учета аргументов (МГУА), подробное описание которой приведено в статье [23], а пример ее применения для решения задачи прогнозирования нестационарной временной последовательности – в статье [21]. На рис. 4 для иллюстрации механизма интеграции программных платформ приведены классы программной платформы МГУА (выделены серым, рамкой «GMDH»; сохранены наименования, используемые в работе [23]).

Реализация системы мета-обучения на основании предложенной объектно-ориентированной архитектуры удовлетворяет всем сформулированным ранее требованиям.

Требование А: автоматическое функционирование системы обеспечивается использованием в MetaLearner алгоритмов решения поставленной задачи и алгоритмов обучения мета-моделей, не требующих присутствия эксперта. В частности, при проведении экспериментов автором использовались алгоритмы метода МГУА.

Требование В: накопление мета-знаний в MetaLearner происходит путем добавления мета-данных о каждом решении поставленной задачи (вы-

зов метода solve()). При этом при увеличении объема доступных мета-данных происходит дообучение (переобучение) моделей metamodels.

Требование С удовлетворяется реализацией в MetaLearner общего или модифицированного способов, предложенных ранее и изображенных на рис. 2, 3.

Требование D удовлетворяется реализацией в MetaLearner модуля самодиагностики: чрезмерно низкое прогнозируемое значение ПЭ для выбранного алгоритма расценивается модулем как недопустимость применения алгоритма для решения задачи. При этом использование алгоритмов МГУА для поиска мета-моделей упрощает реализацию данного модуля в связи с тем, что они позволяют получить значение *внешнего* критерия качества для обученных мета-моделей.

Экспериментальная проверка эффективности предлагаемой системы мета-обучения применительно к решению задачи прогнозирования

Для проверки эффективности предлагаемой системы мета-обучения использовался тестовый набор временных последовательностей, известный под названием «М3 Competition». Данный набор содержит 3003 временные последовательности из разных областей человеческой деятельности: промышленность, экономика, финансы, демография и пр. Целью создания «М3 Competition» было предоставление исследователям в области прогнозирования возможности проверки разрабатываемых ими алгоритмов прогнозирования и сравнения с

существующими аналогами: на сегодняшний день в базе данных IIF (International Institute of Forecasters) содержатся прогнозы, полученные с использованием 24 современных методов прогнозирования, начиная от «наивного прогнозирования» (Naive) и заканчивая искусственными нейронными сетями (ANN) [24].

Значения каждой временной последовательности x_t следуют с фиксированным шагом по времени Δt , равным одному году (645 последовательностей), кварталу (756 последовательностей) или месяцу (1428 последовательностей) [25]. Для каждой последовательности задается число контрольных точек в конце, которые **не могут участвовать в обучении** прогнозирующих моделей, но **должны быть использованы для оценки** точности прогнозирования.

Для экспериментов были выбраны месячные данные: 1428 последовательностей. Для удобства сравнительного анализа результатов прогнозирования все используемые последовательности были нормализованы к интервалу $[-1; 1]$.

Были определены 8 мета-характеристик временных последовательностей, подробная информация по которым приведена в табл. 3. Для каждой из мета-характеристик приведен вид согласно классификации, введенной ранее в аналитическом обзоре систем мета-обучения, а также способ их вычисления. На рис. 5 представлены гистограммы

распределений значений мета-характеристик, рассчитанные для всех 1428 выборок данных (N – число выборок, значение мета-характеристики которых попадает в соответствующий интервал гистограммы), а на рис. 6 визуализированы значения линейного коэффициента их взаимной корреляции (более темный цвет означает более коррелированные значения, знаком «+» и «-» отмечены, соответственно, положительные и отрицательные значения корреляции). Гистограммы распределений для каждой мета-характеристики имеют различный вид, а абсолютное значение корреляции в среднем составляет 0,29, что свидетельствует о том, что мета-характеристики отражают независимые свойства временных последовательностей.

В качестве алгоритма прогнозирования применялся один из «прогнозирующих модулей», описанных автором в работе [21], соответствующий типу с названием «Sliding Window» и характеризующийся параметром T – размером скользящего окна, определяющим количество точек предыстории, использующихся прогнозирующей моделью для построения прогноза на следующий шаг. Данный параметр был выбран в качестве единственного параметра функционирования алгоритма q^a . В экспериментах значение T полагалось равным 1, 2, 3, 4, 5, 6 и 7. Построение прогнозирующих моделей производилось с использованием следующих алгоритмов МГУА:

Таблица 3. Описание применяемых мета-характеристик временных последовательностей

Метка	Вид	Описание и способ вычисления	
length	КОЛ	Длина временной последовательности (число точек)	
noise	СТАТ	Средняя величина абсолютного изменения за шаг: $f^{noise} = \frac{1}{T} \sum_{t=1}^{T-1} x_{t+1} - x_t $	
trend	ПОВ	Коэффициент наклона прямой линейного тренда $f^{trend} = a$. Расчет коэффициента a – из уравнения $x_t = at + b$ по методу наименьших квадратов	
period	ПОВ	Период гармонического сигнала, рассчитываемый согласно критерию баланса ординат из предположения наличия единственной гармоники во входной временной последовательности [26. С. 76–78]	
land1in	ОРИ	Значение CR на обучающей выборке данных: $f^{land1in} = CR_{0, \lfloor T/2 \rfloor}^{combi-1}$	Осуществлялось построение прогнозирующей модели с использованием комбинаторного алгоритма метода МГУА [27. С. 32–38] с ограничением на 1-ю (land1, линейная модель) и 7-ю (land2, нелинейная модель) степень полинома. Производился поиск модели вида $x_t = \phi(t)$. Обучающая выборка составляла первую половину точек временной последовательности, проверочная – вторую. В качестве внешнего критерия использовался критерий регулярности [27. С. 84]: $CR_{T1, T2} = \frac{1}{T2 - T1} \sum_{t=T1}^{T2-1} (x_t - \hat{x}_t)^2,$ где $\hat{x}_t$ – выход модели, $T1$ и $T2$ – начальный и конечный индексы части временной последовательности
land1out	ОРИ	Значение CR на проверочной выборке данных $f^{land1out} = CR_{\lfloor T/2 \rfloor + 1, T}^{combi-1}$	
land2in	ОРИ	Значение CR на обучающей выборке данных: $f^{land2in} = CR_{0, \lfloor T/2 \rfloor}^{combi-7}$	
land2out	ОРИ	Значение CR на проверочной выборке данных $f^{land2out} = CR_{\lfloor T/2 \rfloor + 1, T}^{combi-7}$	

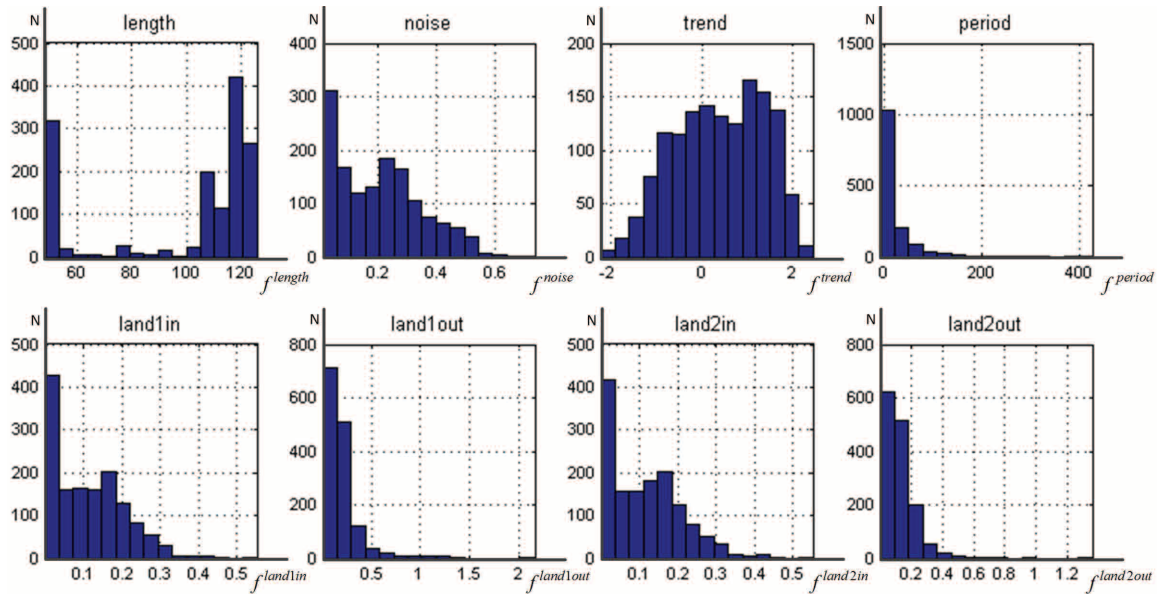


Рис. 5. Гистограммы распределений значений мета-характеристик для месячных временных последовательностей «МЗ Competition»

- Комбинаторный алгоритм; ограничение на степень полинома 1; обучающая выборка – первые 2/3 точек временной последовательности, проверочная – последующие 1/3 точек; внешний критерий селекции моделей – CR. Далее используется метка «COMBI-row1».
- Алгоритм дважды многорядной полиномиальной сети; ограничение на размер сети 2×2 , 3×3 , 4×4 , 5×5 (число слоев $\times$ число нейронов в слое); двухвходовые нейроны; комбинаторный алгоритм обучения нейронов, максимальная степень полиномиального частного описания нейрона 2; минимальное изменение внешнего критерия от слоя к слою 10^{-5} ; обучающая выборка – первые 2/3 точек временной последовательности, проверочная – последующие 1/3 точек; внешний критерий селекции моделей – CR. Далее используются метки «PNN-2 $\times$ 2», «PNN-3 $\times$ 3» и т. п.

Таким образом, для каждой временной последовательности осуществлялся запуск $7 \times (1+4) = 35$ прогнозирующих модулей для получения прогноза на контрольной выборке данных. В качестве ПЭ алгоритмов ($p(a(q^*), x)$) использовался критерий CR, и среднее его значение на контрольных выборках составило 3,52.

Согласно предложенному ранее модифицированному способу (рис. 3), в системе мета-обучения производилось построение двух видов мета-моделей.

- $q_g^{opt} = \eta^g(f)$, где размер вектора мета-характеристик f равнялся 8, $q_g^{opt} - 1$, а число строк матрицы $M_g^{q, opt}$ составляло 1428.
- $\mu^g(f, q_g^{opt})$, где размеры векторов f и q_g^{opt} составляли 8 и 1 соответственно, а число строк матрицы $M^F - 1428$.

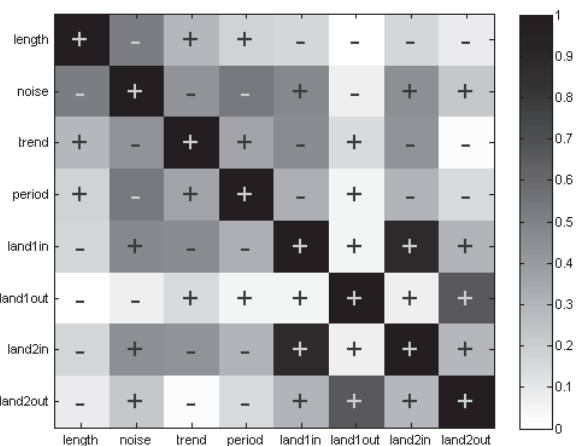


Рис. 6. Величины линейного значения взаимной корреляции значений мета-характеристик для месячных временных последовательностей «МЗ Competition»

Построение мета-моделей осуществлялось с помощью алгоритма дважды многорядной полиномиальной сети МГУА с ограничением на размер 10×10 и внешним критерием CR (разделение выборки на обучающую и проверочную – пополам; двухвходовые нейроны, комбинаторный алгоритм обучения нейронов, максимальная степень полиномиального частного описания 2).

Результатом работы модуля самодиагностики (реализация требования D), применявшегося в том числе для контроля качества мета-знаний, явилось исключение из обучающих выборок 0,1...2,7 % информации по причине несоответствия ее минимальным ожидаемым значениям ПЭ для поставленной задачи прогнозирования ($CR > 10^3$).

Также стоит отметить, что в процессе построения моделей вида $q_g^{opt} = \eta^g(f)$ алгоритмом МГУА бы-

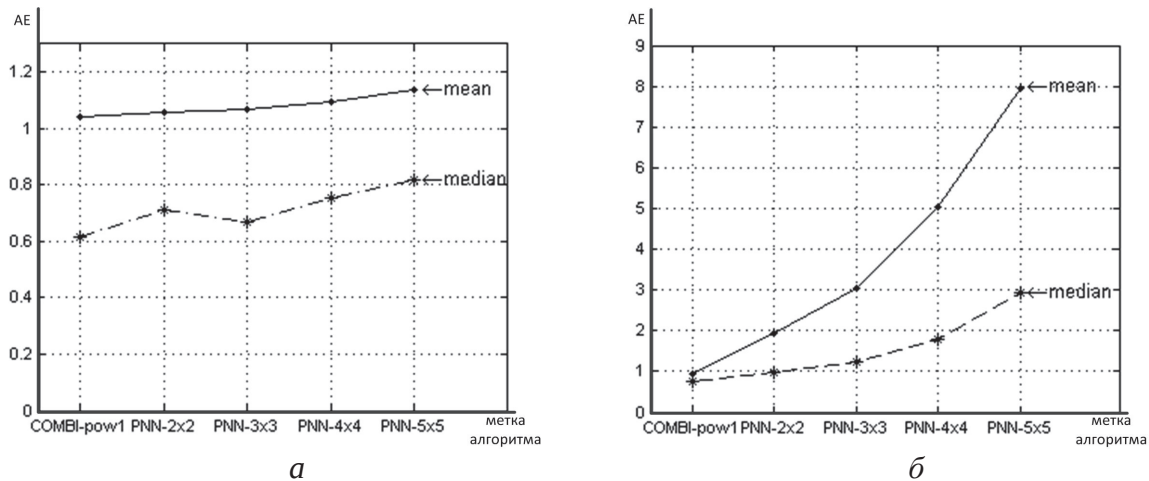


Рис. 7. Средние и медианные значения абсолютных ошибок мета-моделей вида: а) $q_g^{opt} = \eta^g(f)$; б) $\mu^g(f, q_g^{opt})$

ли автоматически исключены мета-характеристики «length» (для 3 мета-моделей из 5), «period» (3 из 5), «landlout» (5 из 5) и «land2out» (4 из 5) как нерепрезентативные. Для моделей вида $\mu^g(f, q_g^{opt})$ – только «land2in» (3 из 5).

На рис. 7, а приведены среднее (mean) и медианное (median) значения абсолютной погрешности (AE) мета-моделей вида $q_g^{opt} = \eta^g(f)$ для доступных системе алгоритмов прогнозирования. Можно видеть, что медианное значение лежит в пределах [0,6; 0,8], что является удовлетворительным показателем для модели, вычисляющей прогноз оптимального значения параметра, лежащего в диапазоне [1; 7]. На рис. 7, б показаны средние (mean) и медианные (median) значения абсолютной погрешности мета-моделей вида $\mu^g(f, q_g^{opt})$. На графике наблюдается возрастание ошибки мета-модели с увеличением сложности прогнозирующих моделей, генерируемых соответствующим алгоритмом. Тем не менее, погрешность находится на приемлемом уровне [1; 8] для показателя, принимающего значение [0; 110].

Проверка эффективности системы мета-обучения проводилась путем передачи очередной временной последовательности системе с целью решения задачи построения прогноза на заданное число шагов вперед. Система производила выбор оптимального алгоритма путем применения обученных мета-моделей обоих видов согласно модифицированному способу, описанному ранее. Среднее значение CR на контрольной выборке для системы составило 1,95, что на 44,7 % меньше, чем среднее значение CR отдельных алгоритмов. Стоит отметить, что среднее значение CR для системы мета-обучения, использующей идеально точные мета-

модели (мета-модели со средним значением абсолютного отклонения, равным нулю, т. е. имеющие нулевую ошибку прогнозирования), составило 1,55, что лишь на 20,4 % лучше полученного результата.

Заключение

Известные на сегодняшний день аналитические обзоры систем мета-обучения выделяют лишь некоторые внешние критерии оценки систем либо рассматривают состав внутренних модулей без выделения общности в организации систем. В данной работе сформулированы классификационные признаки, учитывающие внутреннюю организацию систем, и проведена систематизация существующих систем мета-обучения. Разработана объектно-ориентированная архитектура программной платформы, позволяющая реализовать любую из систем мета-обучения, представленных в систематизации. Предложен способ построения системы мета-обучения, реализованный в рамках программной платформы и удовлетворяющий всем сформулированным требованиям. Проведенные эксперименты по решению задачи прогнозирования с использованием набора алгоритмов МГУА показали эффективность разработанной системы мета-обучения по выбору оптимального алгоритма прогнозирования и расчету оптимальных параметров его функционирования (улучшение значения критерия качества на 44,7 %). Направлением дальнейших работ является анализ поведения алгоритмов выбора оптимального алгоритма решения задачи с целью дальнейшего повышения эффективности системы.

СПИСОК ЛИТЕРАТУРЫ

1. U. Fayyad, Piatetsky-Shapiro G., Smyth P. From Data Mining to Knowledge Discovery in Databases // *AI Magazine: International Journal*. – 1996. – V. 17. – № 3. – P. 37–54.
2. Wolpert D.H. The Lack of a Priori Distinctions between Learning Algorithms // *Neural Computation: International Journal*. – 1996. – V. 8. – Iss. 7. – P. 1341–1390.
3. Rice J. The Algorithm Selection Problem // *Advances in Computers*. – 1976. – V. 15. – P. 65–118.
4. Vilalta R., Drissi Y. A Perspective View and Survey of Meta-Learning // *Artificial Intelligence Review: International Journal*. – 2002. – V. 18. – Iss. 2. – P. 77–95.
5. Meta-Learning in Computational Intelligence. *Studies in Computational Intelligence Series* / Ed. N. Jankowski, W. Duch, K. Grabczewski. – Berlin: Springer, 2011. – V. 358. – 359 p.
6. Prudencio R.B.C., Ludermir T.B. Selecting and ranking time series models using the NOEMON approach // *Proceedings of the Joint International Conference on Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP 2003)*. – Istanbul, Turkey, 2003. – P. 654–661.
7. Prudencio R.B.C., Ludermir T.B. Active Meta-Learning with Uncertainty Sampling and Outlier Detection // *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN 2008)*. – Hong Kong, 2008. – P. 346–351.
8. Romero C., Olmo J.L., Ventura S. A meta-learning approach for recommending a subset of white-box classification algorithms for Moodle datasets // *Proceeding of the VI International Conference on Educational Data Mining (EDM 2013)*. – Memphis, Tennessee, USA, 2013. – P. 268–271.
9. Matijas M., Suykens J.A.K., Krajcar S. Load forecasting using a multivariate meta-learning system // *Expert Systems with Applications: International Journal*. – 2013. – V. 40. – Iss. 11. – P. 4427–4437.
10. Căcovăanu S., Vidrighin C., Potolea R. Evolutional meta-learning framework for automatic classifier selection // *Proceedings of the V International Conference on Intelligent Computer Communication and Processing (IEEE ICCP 2009)*. – Cluj-Napoca, Romania, 2009. – P. 27–30.
11. Algorithm Selection: From Meta-Learning to Hyper-Heuristics / L. Cruz-Reyes, C. Gomez-Santillan, J. Perez-Ortega, V. Landero, M. Quiroz, A. Ochoa // *Intelligent Systems*. – InTech, 2012. – Ch. 6. – 27 p. URL: <http://www.intechopen.com/books/intelligent-systems/algorithm-selection-from-meta-learning-to-hyper-heuristics> (дата обращения: 07.02.2014).
12. Wang X., Smith-Miles K., Hyndman R. Rule induction for forecasting method selection: Meta-learning the characteristics of univariate time series // *Neurocomputing: International Journal*. – 2009. – V. 72. – Iss. 10–12. – P. 2581–2594.
13. Landmarking for Meta-Learning using RapidMiner / S.D. Abdelmessih, F. Shafait, M. Reif, M. Goldstein // *Proceedings of the RapidMiner Community Meeting and Conference*. – Dortmund, Germany, 2010. – 6 p. URL: http://www.dfki.de/web/forschung/iwi/publikationen/renamFileForDownload?filename=Sarah-Landmarking-RCOMM10.pdf&file_id=uploads_781 (дата обращения: 13.12.2013).
14. Meta-Learning Approach for Automatic Parameter Tuning: A Case Study with Educational Datasets / M.M. Molina, J.M. Luna, C. Romero, S. Ventura // *Proceeding of the V International Conference on Educational Data Mining (EDM 2012)*. – Chania, Greece, 2012. – P. 180–183.
15. Musliu N., Schwengerer M. Algorithm Selection for the Graph Coloring Problem // *Proceedings of the VII International Conference on Learning and Intelligent Optimization (LION 7)*. – Catania, Italy, 2013. – P. 389–403.
16. Tuned data mining: A benchmark study on different tuners / W. Konen, T. Bartz-beielstein, P. Koch, M. Friese, O. Flasch, B. Naujoks // *Proceedings of the XIII Annual Conference on Genetic and Evolutionary Computation (GECCO'11)*. – ACM New York, USA, 2011. – P. 1995–2002.
17. Soares C., Brazdil P.B. Selecting parameters of SVM using meta-learning and kernel matrix-based meta-features // *Proceedings of the ACM symposium on Applied computing (SAC'06)*. – New York, USA, 2006. – P. 564–568.
18. Reif M., Shafait F., Dengel A. Meta-learning for evolutionary parameter optimization of classifiers // *Machine Learning: International Journal*. – 2012. – V. 87. – Iss. 3. – P. 357–380.
19. Consultant-2: pre- and post-processing of machine learning applications / D. Sleeman, M. Rissakis, S. Craw, N. Graner, S. Sharma // *International Journal of Human-Computer Studies*. – 1995. – V. 43. – Iss. 1. – P. 43–63.
20. Kalousis A., Hilario M. Representational issues in meta-learning // *Proceedings of the XX International Conference on Machine Learning (ICML-03)*. – Washington, DC, USA, 2003. – P. 313–320.
21. Orlov A.A. Non-Stationary Time Series Forecasting on Basis of Analysis and Prediction of Forecasting Models Efficiency // *Proceedings of the IV International Conference on Inductive Modeling (ICIM-2013)*. – Kyiv, 2013. – P. 192–199.
22. Gomes T.A.F., Prudencio R.B.C., Soares C. Combining meta-learning and search techniques to select parameters for support vector machines // *Neurocomputing: International Journal*. – 2012. – V. 75. – Iss. 1. – P. 3–13.
23. Орлов А.А. Принципы построения архитектуры программной платформы для реализации алгоритмов метода группового учета аргументов (МГУА) // *Управляющие системы и машины: Международный журнал*. – 2013. – № 2. – С. 65–71.
24. Makridakis S., Hibon M. M3-Competition // *INSEAD*. – Fontainebleau, France, 1999. – 41 p. URL: <http://www.insead.edu/facultyresearch/research/doc.cfm?did=1094> (дата обращения: 23.03.2014).
25. Makridakis S., Hibon M. The M3-Competition: results, conclusions and implications // *International Journal of Forecasting*. – 2000. – V. 16. – Iss. 4. – P. 451–476.
26. Ивахненко А.Г. Индуктивный метод самоорганизации моделей сложных систем. – Киев: Наукова Думка, 1982. – 296 с.
27. Madala H.R., Ivakhnenko A.G. *Inductive Learning Algorithms for Complex System Modeling*. – Boca Raton; Ann Arbor; London; Tokyo: CRC Press, 1994. – 368 p.

Поступила 31.03.2014 г.

UDC 004.4::004.85[3+5]

AUTOMATIC META-LEARNING SYSTEM SUPPORTING SELECTION OF OPTIMAL ALGORITHM FOR PROBLEM SOLVING AND CALCULATION OF OPTIMAL PARAMETERS OF ITS FUNCTIONING

Andrey A. Orlov,

Tomsk State University of Control Systems and Radioelectronics,
40, Lenin avenue, Tomsk, 634050, Russia. E-mail: d1scnc@gmail.com

The relevance of the work is caused by necessity of increasing efficiency of automatic data mining systems based on meta-learning. **The main aim of the study** is to design an automatic meta-learning system supporting selection of optimal algorithm for problem solving and calculation of optimal parameters of its functioning.

The methods used in the study: inductive modeling, methods of statistical analysis of results.

Results: The known meta-learning systems were integrated based on produced classification features taking into account internal structure of systems. The author has stated the requirements for implementation of the automatic meta-learning system and has offered the way to build a meta-learning system satisfying all stated requirements and accumulating meta-knowledge, building meta-models on its basis, selecting optimal algorithm from a set of available ones and calculating optimal parameters of its functioning. The object-oriented architecture of a software framework for implementation of any meta-learning system presented in the systematization was developed. The efficiency of the implemented automatic meta-learning system using algorithms of group method of data handling was experimentally examined being applied to solution of problems related to the short-term time series forecasting (1428 time series from the testing set known as «M3 Competition»).

Key words:

Meta-learning, meta-features of data, meta-model, software framework, object-oriented analysis and design, forecasting of time series, group method of data handling.

REFERENCES

1. Fayyad U., Piatetsky-Shapiro G., Smyth P. From Data Mining to Knowledge Discovery in Databases. *AI Magazine: International Journal*, 1996, vol. 17, no. 3, pp. 37–54.
2. Wolpert D.H. The Lack of A Priori Distinctions Between Learning Algorithms. *Neural Computation: International Journal*, 1996, vol. 8, Iss. 7, pp. 1341–1390.
3. Rice J. The Algorithm Selection Problem. *Advances in Computers*, 1976, vol. 15, pp. 65–118.
4. Vilalta R., Drissi Y. A Perspective View and Survey of Meta-Learning. *Artificial Intelligence Review: International Journal*, 2002, vol. 18, Iss. 2, pp. 77–95.
5. *Meta-Learning in Computational Intelligence. Studies in Computational Intelligence Series*. Ed. N. Jankowski, W. Duch, K. Grabczewski. Berlin, Springer, 2011. Vol. 358. 359 p.
6. Prudencio R.B.C., Ludermit T.B. Selecting and ranking time series models using the NOEMON approach. *Proceedings of the Joint International Conference on Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP 2003)*. Istanbul, Turkey, 2003. pp. 654–661.
7. Prudencio R.B.C., Ludermit T.B. Active Meta-Learning with Uncertainty Sampling and Outlier Detection. *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN 2008)*. Hong Kong, 2008. pp. 346–351.
8. Romero C., Olmo J.L., Ventura S. A meta-learning approach for recommending a subset of white-box classification algorithms for Moodle datasets. *Proceeding of the VI International Conference on Educational Data Mining (EDM 2013)*. Memphis, Tennessee, USA, 2013. pp. 268–271.
9. Matijas M., Suykens J.A.K., Krajcar S. Load forecasting using a multivariate meta-learning system. *Expert Systems with Applications: International Journal*, 2013. vol. 40, Iss. 11, pp. 4427–4437.
10. Cacoveanu S., Vidrighin C., Potolea R. Evolutional meta-learning framework for automatic classifier selection. *Proceedings of the V International Conference on Intelligent Computer Communication and Processing (IEEE ICCP 2009)*, 2009. pp. 27–30.
11. Cruz-Reyes L., Gomez-Santillan C., Perez-Ortega J., Landero V., Quiroz M., Ochoa A. Algorithm Selection: From Meta-Learning to Hyper-Heuristics. *Intelligent Systems. InTech*, 2012. Ch. 6, 27 p. Available at: <http://www.intechopen.com/books/intelligent-systems/algorithm-selection-from-meta-learning-to-hyper-heuristics> (accessed 07 February 2014).
12. Wang X., Smith-Miles K., Hyndman R. Rule induction for forecasting method selection: Meta-learning the characteristics of univariate time series. *Neurocomputing: International Journal*, 2009, vol. 72, Iss. 10–12, pp. 2581–2594.
13. Abdelmessih S.D., Shafait F., Reif M., Goldstein M. Landmarking for Meta-Learning using RapidMiner. *Proceedings of the RapidMiner Community Meeting and Conference*. Dortmund, Germany, 2010. 6 p. Available at: http://www.dfki.de/web/forschung/iwi/publikationen/renamFileForDownload?filename=Sarah-Landmarking-RCOMM10.pdf&file_id=uploads_781 (accessed 13 December 2013).
14. Molina M.M., Luna J.M., Romero C., Ventura S. Meta-Learning Approach for Automatic Parameter Tuning: A Case Study with Educational Datasets. *Proceeding of the V International Conference on Educational Data Mining (EDM 2012)*. Chania, Greece, 2012. pp. 180–183.
15. Musliu N., Schwengerer M. Algorithm Selection for the Graph Coloring Problem. *Proceedings of the VII International Conference on Learning and Intelligent Optimization (LION 7)*. Catania, Italy, 2013. pp. 389–403.
16. Konen W., Bartz-beielstein T., Koch P., Friese M., Flasch O., Naujoks B. Tuned data mining: A benchmark study on different tuners. *Proceedings of the XIII Annual Conference on Genetic and Evolutionary Computation (GECCO '11)*. ACM New York, USA, 2011. pp. 1995–2002.
17. Soares C., Brazdil P.B. Selecting parameters of SVM using meta-learning and kernel matrix-based meta-features. *Proceedings of the ACM symposium on Applied computing (SAC '06)*. New York, USA, 2006. pp. 564–568.
18. Reif M., Shafait F., Dengel A. Meta-learning for evolutionary parameter optimization of classifiers. *Machine Learning: International Journal*, 2012, vol. 87, Iss. 3, pp. 357–380.
19. Sleeman D., Rissakis M., Craw S., Graner N., Sharma S. Consultant-2: pre- and post-processing of machine learning applications. *International Journal of Human-Computer Studies*, 1995, vol. 43, Iss. 1, pp. 43–63.

20. Kalousis A., Hilario M. Representational issues in meta-learning. *Proceedings of the XX International Conference on Machine Learning (ICML-03)*. Washington, DC, USA, 2003. pp. 313–320.
21. Orlov A.A. Non-Stationary Time Series Forecasting on Basis of Analysis and Prediction of Forecasting Models Efficiency. *Proc. the IV International Conference on Inductive Modelling (ICIM-2013)*. Kyiv, 2013. pp. 192–199.
22. Gomes T.A.F., Prudencio R.B.C., Soares C. Combining meta-learning and search techniques to select parameters for support vector machines. *Neurocomputing: International Journal*. – Elsevier Science Publishers, Amsterdam, 2012, vol. 75. Iss. 1. pp. 3–13.
23. Orlov A.A. Printsipy postroeniya arkhitektury programnoy platformy dlya realizatsii algoritmov metoda gruppovogo ucheta argumentov (MGUA) [Principles of constructing software framework architecture to implement the algorithms of group method of data handling (GMDH)]. *Upravlyayushchie Sistemy i Mashiny: Mezhdunarodnyi Zhurnal – Control Systems and Computers: International Journal*, 2013, no. 2. pp. 65–71.
24. Makridakis S., Hibon M. M3-Competition. *INSEAD*. Fontainebleau, France, 1999. 41 p. Available at: <http://www.insead.edu/facultyresearch/research/doc.cfm?did=1094> (accessed 23 March 2014).
25. Makridakis S., Hibon M. The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, 2000, vol. 16, Iss. 4, pp. 451–476.
26. Ivakhnenko A.G. *Induktivnyy metod samoorganizatsii modeley slozhnykh system* [Inductive method of self-organization of models of complex systems]. Kiev, Naukova Dumka, 1982. 296 p.
27. Madala H.R., Ivakhnenko A.G. *Inductive Learning Algorithms for Complex System Modeling*. Boca Raton; Ann Arbor; London; Tokyo, CRC Press, 1994. 368 p.

УДК 004.822; 004.942; 620.9:001.891.57

ИНТЕГРАЦИЯ СЕМАНТИЧЕСКИХ МОДЕЛЕЙ В ИССЛЕДОВАНИЯХ ПРОБЛЕМ ЭНЕРГЕТИЧЕСКОЙ БЕЗОПАСНОСТИ

Массель Алексей Геннадьевич,

канд. техн. наук, ст. науч. сотр. Института систем энергетики

им. Л.А. Мелентьева (ИСЭМ) СО РАН,

Россия, 664033, г. Иркутск, ул. Лермонтова, 130. E-mail: amassel@gmail.com

Тюрюмин Вадим Олегович,

аспирант Института систем энергетики им. Л.А. Мелентьева (ИСЭМ)

СО РАН Россия, 664033, г. Иркутск, ул. Лермонтова, 130.

E-mail: vadim.tyuryumin@gmail.com

Актуальность работы определяется, с одной стороны, важностью проблем энергетической безопасности и их исследований, с другой – необходимостью развития интеллектуальных инструментальных средств исследований и их интеграции с традиционными математическими моделями.

Цель работы: усовершенствование двухуровневой технологии исследований проблем энергетической безопасности за счет интеграции семантических моделей и расширения онтологического пространства знаний о предметной области онтологиями событий.

Методы исследования: В предложенной ранее двухуровневой технологии исследований проблем энергетической безопасности на верхнем (первом) уровне выполняется качественный анализ (экспресс-анализ) с применением методов и средств семантического (онтологического, когнитивного и событийного) моделирования, на нижнем (втором) – количественный анализ на основе численных расчетов с использованием традиционных программных комплексов. Когнитивные модели используются для моделирования угроз энергетической безопасности, событийные модели – для моделирования вариантов развития чрезвычайных ситуаций в энергетике. Совместное использование когнитивного и событийного моделирования позволяет получить более объективную оценку ситуации. Вводятся онтологии событий, используемые для перехода от когнитивных к событийным моделям. Рассматривается использование Joiner-сетей для последующего анализа и обработки событийных моделей.

Результаты: Предложены правила генерирования онтологий событий. Описан алгоритм автоматизированного перехода от когнитивных карт к событийным моделям на основе онтологий (с использованием онтологий событий) как один из этапов усовершенствования двухуровневой технологии.

Приведены разработанные онтологии, когнитивные и событийные карты и построенная на их основе событийная модель с использованием Joiner-сетей.

Ключевые слова:

Энергетическая безопасность, семантическое моделирование, онтологическое, когнитивное и событийное моделирование, онтологии событий, Joiner-сети.

Введение

В Институте систем энергетики им. Л.А. Мелентьева СО РАН проводятся комплексные исследования, важную роль в которых играют исследования проблем энергетической безопасности.

Энергетическая безопасность (ЭБ) рассматривается как часть национальной безопасности, а именно как защищенность граждан, общества, государства, экономики от угроз дефицита в обеспечении их обоснованных потребностей топливно-энерге-

тическими ресурсами (ТЭР) приемлемого качества [1]. Важными для исследований проблем ЭБ являются понятия чрезвычайной ситуации (ЧС) и угроз ЭБ, под которыми понимаются события, неблагоприятные для энергетики. Анализ развития и последствий ЧС направлен на установление причин возникновения ЧС, характера ее развития и масштаба последствий. Основная сложность проведения таких исследований заключается в их многовариантном характере. Для преодоления этой сложности коллективом, в котором работают авторы (лаборатория Информационных технологий в энергетике ИСЭМ СО РАН, возглавляемая д.т.н. Л.В. Массель), была предложена двухуровневая технология исследований проблем ЭБ, на верхнем (первом) уровне которой выполняется этап качественного анализа (экспресс-анализ) с применением методов и средств семантического (онтологического, когнитивного и событийного) моделирования, а на нижнем (втором) – количественный анализ на основе численных расчетов с использованием традиционных программных комплексов. Это позволяет существенно сократить количество рассматриваемых вариантов проведения вычислительного эксперимента, проводимого на втором уровне, и снизить нагрузку на эксперта [2]. Современное состояние двухуровневой технологии и инструментальных средств ее поддержки подробно рассмотрено в [3]. В статье рассматривается задача интеграции и совместного использования в исследованиях проблем ЭБ онтологических, когнитивных и событийных моделей.

Семантическое моделирование

Под *семантической моделью* в обобщенном виде понимается информационная модель, отражающая понятия предметной области и отношения между ними [4]. Авторы рассматривают семантическое моделирование на примере онтологических, когнитивных и событийных моделей [5].

Под *онтологическим моделированием* понимается построение онтологий как в графическом, так и формализованном виде. Онтологии определяют как базу знаний специального вида, или как «спецификацию концептуализации» предметной области [6]. Последнее означает процесс классификации базовых терминов предметной области с определением основных понятий (концептов) и установлением связей между ними. В свою очередь, процесс спецификации заключается в описании онтологии в графическом виде или на одном из формальных языков (XML, RDFS, OWL и др.) [7]. Для работы с экспертами авторы используют графическое представление онтологий; для хранения онтологий используется их представление на языке XML.

Под *когнитивным моделированием* понимается построение когнитивных моделей, или, иначе, когнитивных карт (ориентированных графов), в которых вершины соответствуют факторам (концептам), а дуги – связям между факторами, (положительным или отрицательным), в зависимости от

характера причинно-следственного отношения [8]. В простейшем случае веса связей могут иметь значения +1 или –1 либо принимать нечеткие значения из отрезка $[-1, 1]$ или некоторой лингвистической шкалы, что в наибольшей степени соответствует качественному анализу [9].

Под *событийным моделированием* понимается построение поведенческих моделей, причем в качестве объектов моделирования могут выступать как люди, так и технические объекты. Сущность событийного метода моделирования заключается в отслеживании на модели последовательности событий в том же порядке, в каком они происходили бы в реальной системе. Задаваемые моделью последовательности реализации событий – цепочки событий – описывают сценарии реакции системы на возникновение инициирующего события, стоящего в начале цепочки. В результате событийная модель позволяет получить множество альтернативных сценариев развития заданной ситуации в системе, что и является основной целью событийного моделирования [10].

Методика совместного использования онтологического, когнитивного и событийного моделирования в исследованиях ЭБ с использованием инструментальных средств их поддержки включает в себя этап перехода от когнитивных карт к событийным моделям в виде событийных карт [11–13]. На текущий момент этот переход осуществляется вручную с помощью эксперта, что занимает довольно много времени и не позволяет оперативно перейти к количественному анализу полученных результатов и проведению вычислительных экспериментов. Для решения этой проблемы и обеспечения интеграции программных средств когнитивного и событийного моделирования в рамках вышеупомянутой методики предлагается автоматизировать этап перехода от когнитивных карт к событийным моделям с использованием онтологий.

Использование онтологии событий для перехода от когнитивных карт к событийным моделям

Ввиду того, что когнитивные карты отображают причинно-следственные связи факторов и силу их взаимовлияния, а событийные модели показывают развитие ситуации во времени, становится очевидной невозможность непосредственного сопоставления факторов на когнитивной карте и событий в событийной модели. Данная проблема может быть решена путем использования онтологий в качестве классификаторов концептов предметной области. На практике это означает, что каждому фактору когнитивной карты ставится в соответствие концепт из онтологического пространства знаний, который фиксирует принадлежность рассматриваемого фактора к некоторой группе объектов топливно-энергетического комплекса (ТЭК) или процессов в нем. На основании анализа предметной области были выделены следующие группы факторов, используемые при построении когнитивных карт:

1. Негативные факторы (угрозы ЭБ).
2. Объекты ТЭК (объекты добычи/производства, переработки, транспортировки и хранения ТЭР, потребители ТЭР).
3. Мероприятия по устранению причин и следствий угроз ЭБ.
4. Показатели, отображающие некоторые числовые характеристики объектов или процессов. Показатели всегда связаны с некоторым объектом ТЭК и выделяются в отдельную группу, как часто встречающийся факторы в когнитивных картах (иллюстративным примером является модель влияния потребления электроэнергии в регионе, предложенная Р. Аксельродом [14]).

Как уже было отмечено в предыдущем разделе, факторам, соответствующим объектам ТЭК, невозможно прямо сопоставить некоторое событие на событийной карте, поскольку оно появляется на этапе перехода от когнитивных к событийным моделям при рассмотрении причинно-следственных отношений факторов. Для преодоления этой сложности авторами предлагается использовать *онтологию событий*.

Онтологии событий рассматриваются в данном случае как онтологии, описывающие фрагменты знаний о возможных событиях, способных возникнуть на объектах ТЭК при воздействии на них негативных факторов или, наоборот, факторов, улучшающих их состояние. Таким образом, онтологии событий расширяют онтологическое пространство знаний о некотором объекте ТЭК. На рис. 1 представлен фрагмент такой онтологии для объектов добычи углеснабжающей системы (АВП – аварии, взрывы, пожары).

Для осуществления автоматизации перехода от когнитивных карт к событийным моделям авторами предлагается набор *правил* генерирования событий на основе принадлежности фактора к той или иной группе, согласно онтологии и характеру причинно-следственного отношения, связывающего его с другими факторами, с учетом весового коэффициента связи:

1. Угрозе ЭБ или иному негативному явлению в системах энергетики можно сопоставить соответствующее событие.

2. Для фактора, соответствующего объекту ТЭК, выбирается событие из онтологии событий, связанной с этим фактором.
3. Управляющим воздействиям можно сопоставить соответствующие события.
4. Показателю можно сопоставить событие, связанное с изменением его значения (например, для фактора «Уровень потребления ТЭР» таким событием может являться «Увеличение/уменьшение уровня потребления ТЭР»).

Кроме вышеперечисленных, существует еще одно общее правило логического следствия: при анализе угроз ЭБ и построении событийных моделей в начале цепочки событий ставятся превентивные мероприятия (если присутствуют) или ЧС, затем их негативные последствия (могут быть извлечены из онтологии событий) в системах энергетики. Далее указываются ликвидационные мероприятия, направленные на устранение последствий ЧС. После их выполнения при необходимости может быть выполнен повторный обход когнитивной карты с генерированием соответствующих событий на основании направлений связей и их весов. Представляется, что последний вариант построения событийной модели будет особенно полезен при работе с динамическими когнитивными картами.

Алгоритм автоматизированного перехода от когнитивных карт к событийным моделям с использованием онтологий

На основании вышеперечисленных правил авторами предлагается алгоритм автоматизированного перехода от когнитивных карт к событийным моделям с использованием онтологий в качестве классификаторов концептов:

1. Осуществляется выбор одного или нескольких инициирующих событий. В соответствии с правилом логического следствия и правилами генерирования событий (предыдущий раздел) таким событием становится угроза ЭБ или превентивные мероприятия.
2. Из множества факторов выбирается один элемент (текущий фактор), связанный отношением следствия с исходным фактором.

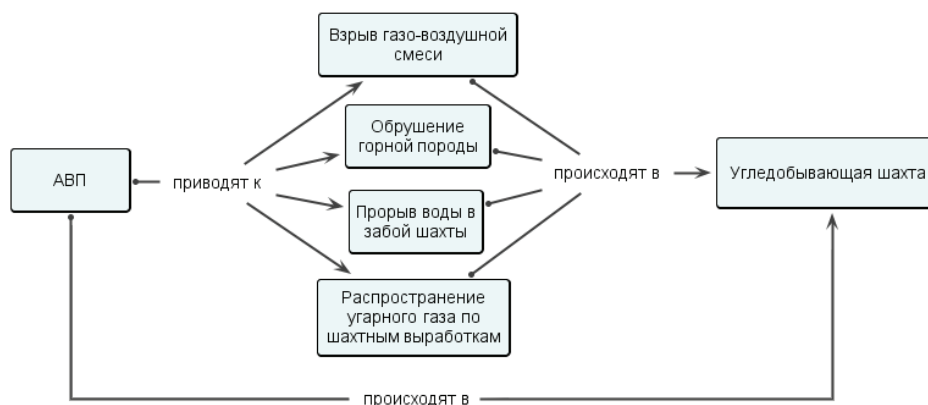


Рис. 1. Фрагмент онтологии событий для объекта добычи углеснабжающей системы

3. Осуществляется сбор информации для последующего генерирования события: анализируется вес связи между факторами и принадлежность текущего фактора к той или иной их группе согласно системе онтологий.
4. На основании информации, полученной на предыдущем шаге алгоритма, фактору сопоставляется некоторое событие, которое занимает свое место в их причинно-следственной цепочке.
5. Пункты 2–4 выполняются для каждого элемента множества факторов когнитивной карты до тех пор, пока не останутся нерассмотренными факторы, обозначающие ликвидационные мероприятия.
6. Если на когнитивной карте присутствуют факторы, соответствующие ликвидационным мероприятиям, то они рассматриваются как исходные факторы, а соответствующие им события продолжают построенную ранее их причинно-следственную цепочку. Пункты 2–4 выполняются для каждого элемента множества факторов когнитивной карты. Когда рассмотрены все факторы, работа алгоритма заканчивается.

Алгоритм в виде блок-схемы представлен на рис. 2. Нумерация блоков соответствует вышеприведенным пунктам алгоритма.

В качестве примера, демонстрирующего работу вышеописанного алгоритма, выбрана группа угроз ЭБ техногенного происхождения «Аварии, взрывы, пожары» («АВП»). Реализация угроз такого типа проявляется в разрушении объектов ТЭК и, прежде всего, объектов добычи/производства энергоресурсов, транспортных магистралей и потребителей энергоресурсов. Эти события, в свою очередь, могут повлечь за собой истощение запасов энергоресурсов с угрозой их дефицита. Мероприятия, призванные ликвидировать последствия реализации ЧС, направлены на скорейшее восстановление объектов ТЭК и связей между ними для поддержания нормативного уровня запасов энергоресурсов.

Согласно методике совместного использования онтологического, когнитивного и событийного моделирования, на первом этапе разрабатывается онтология или метаонтология, описывающая взаимосвязи основных концептов предметной области (рис. 3). На рисунке показано, что выделяются шесть типов угроз ЭБ: техногенные, природные, экономические, социально-политические, управленческо-правовые, а также угрозы кибербезопасности. Мероприятия, которые повышают уровень ЭБ, могут быть как превентивными (предотвращающими угрозы ЭБ), так и ликвидационными

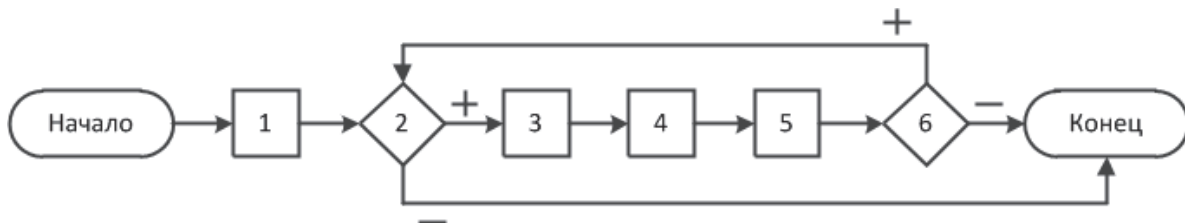


Рис. 2. Блок-схема алгоритма перехода от когнитивных карт к событийным моделям на основе онтологий

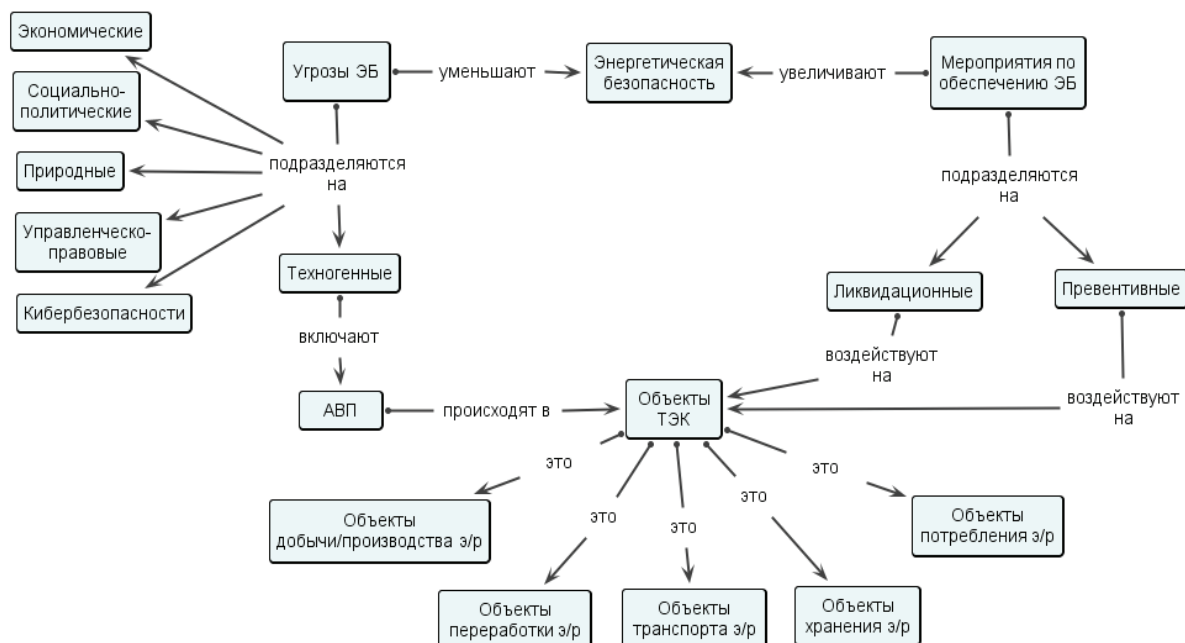


Рис. 3. Фрагмент метаонтологии, используемой для описания угрозы «АВП»

ми (если угрозу возникновения чрезвычайной ситуации предотвратить не удалось, ЧС произошла и надо устранять ее последствия).

На втором этапе выполняется когнитивное моделирование угрозы с описанием основных зависимостей между негативными факторами, показателями ТЭК и мероприятиями по обеспечению энергетической безопасности (рис. 4).

На основе имеющейся метаонтологии описания угрозы «АВП» и когнитивной карты строится набор онтологий событий (в данном случае метаонтологий), позволяющих перейти от описания фактора к соответствующему событию в событийной модели. Примеры метаонтологий событий в объектах ТЭК при реализации в них угрозы «АВП» и ликвидационных мероприятий представлены на рис. 5.

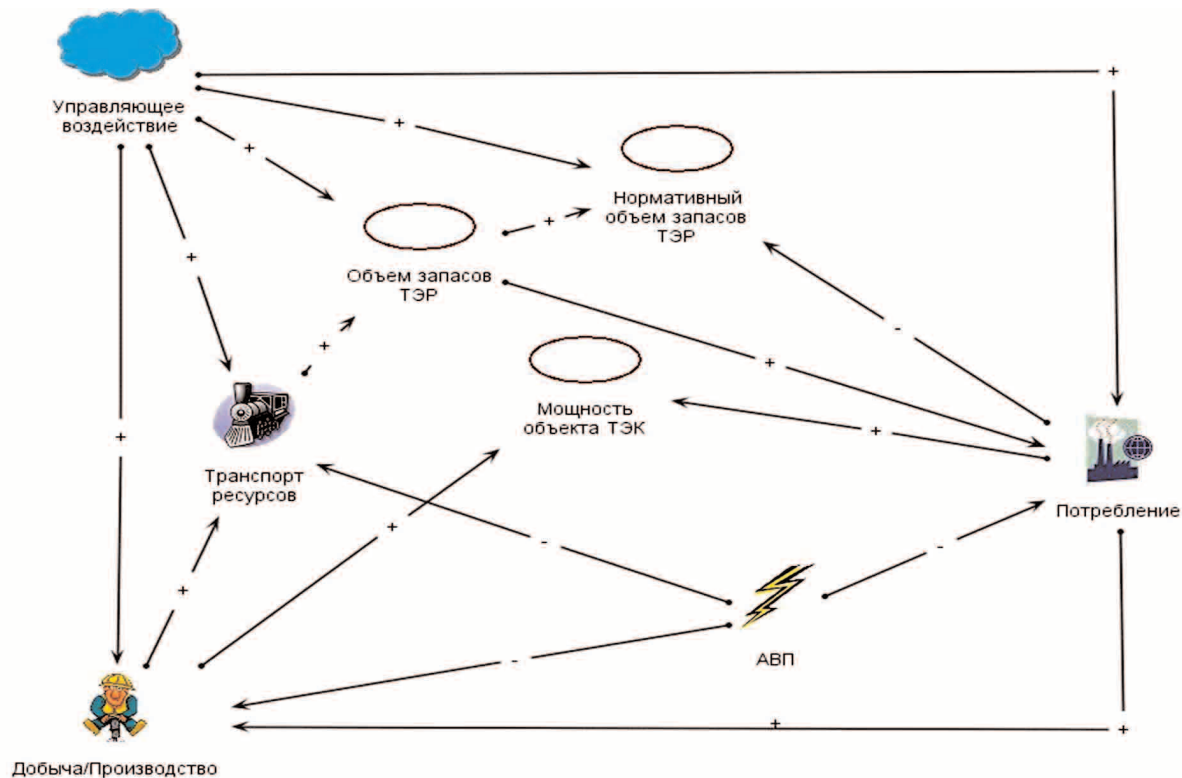


Рис. 4. Когнитивная карта угрозы «АВП»

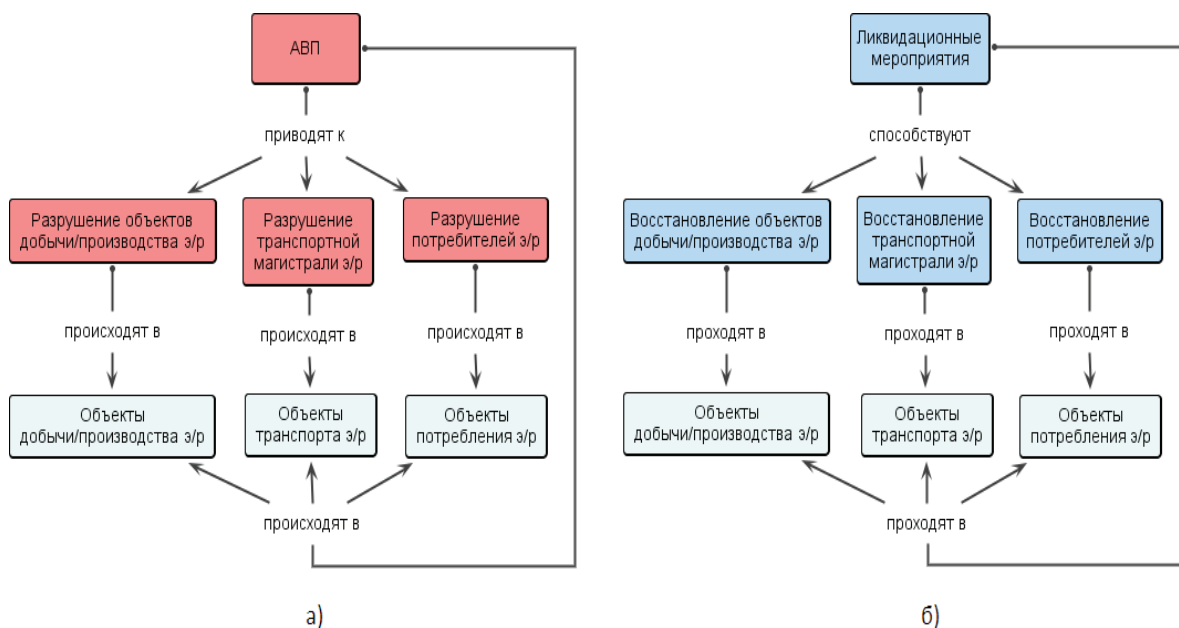


Рис. 5. Примеры метаонтологий событий объектов ТЭК: а) реализация угрозы «АВП»; б) реализация ликвидационных мероприятий

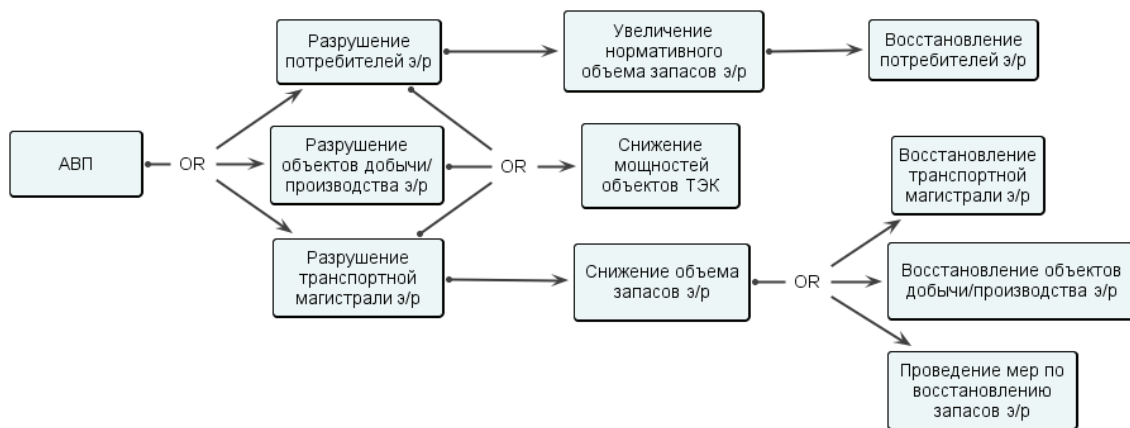


Рис. 6. Событийная карта угрозы «АВП»

Далее для формирования сценариев реализации мероприятий с помощью предложенного авторами алгоритма разрабатывается событийная модель, инициирующим событием в которой является угроза «АВП». События генерируются при переходе от одного фактора к другому с учетом веса и направления связи, а также построенных ранее метаонтологий событий (рис. 5). Результат представлен на рис. 6.

Полученная таким образом событийная модель может быть отредактирована исследователем и использована для формирования многовариантных сценариев проведения вычислительного эксперимента.

Использование Joiner-сетей для последующего анализа и обработки событийных моделей

Авторы предлагают использовать когнитивные модели для моделирования угроз ЭБ. Для рассмотрения вариантов развития ЧС предлагается ис-

пользовать событийное моделирование (определенно выше). Совместное использование когнитивного и событийного моделирования позволяет получить более объективную оценку ситуации.

Алгоритм автоматизированного перехода от когнитивных карт к событийным моделям с использованием онтологий в качестве классификаторов концептов позволяет обойти ряд трудностей. Для лучшего восприятия информации экспертом предлагается использовать аппарат Joiner-сетей, который в дальнейшем позволяет нам выполнять анализ полученных событийных карт.

Аппарат Joiner-сетей является одной из разновидностей алгебраических сетей, предложенной проф. МФТИ Л.Н. Столяровым и развиваемой его учениками [10, 15–18]. Joiner-сети (Joiner-Nets – JN) можно рассматривать как расширение сетей Петри, ориентированное на построение поведенческих моделей. В основе теории JN лежит описание логики

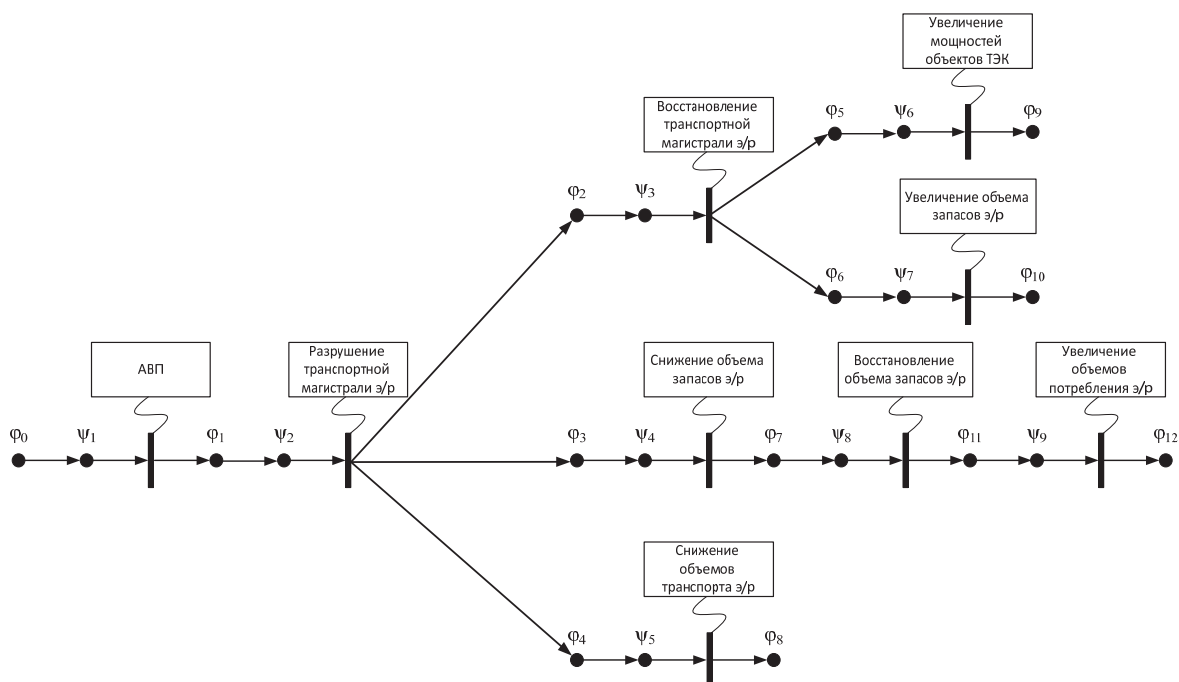


Рис. 7. Пример JN, построенной по событийной карте угрозы «АВП»

взаимодействия асинхронных процессов в виде набора пусковых и флаговых функций, состоящих из булевых функций. Особенностью JN является то, что они предусматривают как графическое представление, так и описание в виде логических формул, обработку которых можно автоматизировать.

Таблица. Список логических уравнений

Пусковые функции	Флаговые функции
$\psi_1(t+1)=\varphi_0(t) \cdot \varphi_1(t)$	$\varphi_0(t+1)=0; \varphi_1(t+1)=1$
$\psi_2(t+1)=\varphi_1(t) \cdot \varphi_2(t) \times \varphi_3(t) \cdot \varphi_4(t)$	$\varphi_1(t+1)=0; (\varphi_2(t+1)=1) \vee (\varphi_3(t+1)=1) \vee (\varphi_4(t+1)=1)$
$\psi_3(t+1)=\varphi_2(t) \cdot \varphi_5(t) \times \varphi_6(t)$	$\varphi_2(t+1)=0; (\varphi_5(t+1)=1) \vee (\varphi_6(t+1)=1)$
$\psi_4(t+1)=\varphi_3(t) \cdot \varphi_7(t)$	$\varphi_3(t+1)=0; \varphi_7(t+1)=1$
$\psi_5(t+1)=\varphi_4(t) \cdot \varphi_8(t)$	$\varphi_4(t+1)=0; \varphi_8(t+1)=1$
$\psi_6(t+1)=\varphi_5(t) \cdot \varphi_9(t)$	$\varphi_5(t+1)=0; \varphi_9(t+1)=1$
$\psi_7(t+1)=\varphi_6(t) \cdot \varphi_{10}(t)$	$\varphi_6(t+1)=0; \varphi_{10}(t+1)=1$
$\psi_8(t+1)=\varphi_7(t) \cdot \varphi_{11}(t)$	$\varphi_7(t+1)=0; \varphi_{11}(t+1)=1$
$\psi_9(t+1)=\varphi_{11}(t) \cdot \varphi_{12}(t)$	$\varphi_{11}(t+1)=0; \varphi_{12}(t+1)=1$

На рис. 7 представлен пример фрагмента JN, построенной по событийной карте угрозы «АВП» (рис. 6).

Список логических уравнений, описывающих передачу возбуждений в Joiner-сети, представлен

ной на рис. 7, представлен в таблице, Joiner-сеть соответствует событийной карте, показанной на рис. 6.

Заключение

В статье рассмотрена интеграция семантических моделей в исследованиях проблем ЭБ. Предложен алгоритм автоматизированного перехода от когнитивных карт к событийным моделям с использованием онтологий, который расширяет методику совместного использования онтологического, когнитивного и событийного моделирования для ситуационного анализа в исследованиях ЭБ.

Алгоритм, описанный в работе, использует систему онтологий в качестве классификатора концептов исследуемой предметной области. В дополнение к традиционным онтологиям авторы предлагают использовать онтологии событий, описывающие фрагменты знаний о возможных событиях в объектах ТЭК при воздействии на них негативных факторов, или, наоборот, факторов, улучшающих их состояние. Таким образом, используя онтологическое пространство знаний, возможно обеспечить лучшую интеграцию методов и программных средств когнитивного и событийного моделирования.

Работа выполнена при частичной финансовой поддержке грантов РФФИ № 12-07-00359, 13-07-140, 14-07-116, а также грантов Программы Президиума РАН № 229 и интеграционного проекта СО РАН № 18Б.

СПИСОК ЛИТЕРАТУРЫ

1. Энергетическая безопасность России / Бушуев В.В., Воропай Н.И., Мастепанов А.М., Шафраник Ю.К. – Новосибирск: Сибир. издат. фирма РАН «Наука», 1998. – 302 с.
2. Массель А.Г. Двухуровневая технология вычислительных экспериментов в исследованиях проблем энергетической безопасности // Информационные и математические технологии в науке и управлении: Труды XVI Байкальской Всероссийской конференции. – Иркутск: ИСЭМ СО РАН, 2011. – Ч. III. – С. 70–76.
3. Массель Л.В., Массель А.Г. Интеллектуальные вычисления в исследованиях направлений развития энергетики // Известия Томского политехнического университета. – 2012. – Т. 321. – № 5. – С. 135–141.
4. Массель Л.В. Интеграция семантического и математического моделирования в исследованиях проблем энергетической безопасности // Моделирование-2012: Труды Междунар. конф. – Киев: ИПМЭ НАН Украины, 2012. – С. 270–273.
5. Массель Л.В., Массель А.Г. Семантические технологии на основе интеграции онтологического, когнитивного и событийного моделирования // OSTIS-2013: Матер. III Междунар. научно-техн. конф. – Беларусь, Минск: БГУИР, 2013. – С. 247–250.
6. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СПб.: Питер, 2001. – 384 с.
7. Тузовский А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии). – Томск: Изд-во НТЛ, 2005. – 260 с.
8. Трахтенгерц Э.А. Компьютерная поддержка принятия решений. – М.: СИНТЕГ, 1998. – 376 с.
9. Kosko B. Fuzzy Cognitive Maps // International Journal of Man-Machine Studies. – 1986. – № 24. – Р. 65–75.
10. Столяров Л.Н. Философия событийного моделирования на примере сценария энергетической катастрофы // Информационные технологии в науке, образовании, телекоммуникации и бизнесе: Труды Междунар. конф. – Гурзуф, 2010. – С. 197–200.
11. Массель А.Г. Методологический подход к организации интеллектуальной поддержки исследований проблемы энергетической безопасности // Информационные технологии. – 2010. – № 9. – С. 32–36.
12. Массель Л.В. Применение онтологического, когнитивного и событийного моделирования для анализа развития и последствий чрезвычайных ситуаций в энергетике // Проблемы безопасности и чрезвычайных ситуаций. – 2010. – № 2. – С. 34–43.
13. Аршинский В.Л. Методический подход к событийному моделированию в исследованиях энергетической безопасности // Информационные и математические технологии в науке и управлении: Труды XV Байкальской Всеросс. конф. – Иркутск: ИСЭМ СО РАН, 2010. – Ч. III. – С. 120–129.
14. Axelrod R. Structure of decision. – Princeton, New Jersey: Princeton University Press, 1976. – 404 p.
15. Новик К.В. Сеть автоматов для моделирования асинхронного взаимодействия процессов: автореф. дис. ... канд. физ.-мат. наук. – М., 2006. – 22 с.
16. Столяров Л.Н., Новик К.В. Joiner-сеть для моделирования взаимодействующих параллельных процессов // Моделирование процессов управления: Сб. научных трудов. – М.: Изд-во Моск. физ.-тех. ин-та, 2004. – С. 81–97.
17. Столяров Л.Н., Новик К.В. Реализация параллельных процессов с помощью сетей Joiner-net // Информационные и математические технологии: Сб. научных трудов. – Иркутск: ИСЭМ СО РАН, 2004. – С. 11–14.
18. Анисимов М.М. Управление событийными сетями // Информационные и математические технологии в науке и управлении: Труды XIV Байкальской Всеросс. конф. – Иркутск: ИСЭМ СО РАН, 2009. – Ч. 3. – С. 238–240.

Поступила 09.04.2014 г.

SEMANTIC MODELS INTEGRATION IN RESEARCH OF ENERGY SECURITY PROBLEMS

Aleksei G. Massel,

Cand. Sc., Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences, 130, Lermontov street, Irkutsk, 664033, Russia.

E-mail: amassel@gmail.com

Vadim O. Tyuryumin,

Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences, 130, Lermontov street, Irkutsk, 664033, Russia.

E-mail: vadim.tyuryumin@gmail.com

The urgency of the work is determined, on the one hand, by the importance of energy security issues and its research, on the other – by the need to develop intelligent research tools and their integration with traditional mathematical models.

The main aim of the study: improvement of two-level technology of researching energy security problems by integrating semantic models and expanding ontological space of knowledge on object domain by event ontologies.

The methods used in the study: On the top (first) level of the previously proposed two-level technology of energy security problems research a qualitative analysis (express analysis) is performed using the methods and means of semantic (ontological, cognitive and event) simulation, on the lower (second) level a quantitative analysis is performed on the basis of numerical calculations using traditional software systems. Cognitive models are used to simulate the energy security threats, event model are used to simulate emergencies development in energy sector. Joint implementation of cognitive and event modeling allows obtaining more objective evaluation of the situation. The paper introduces the event ontologies used for transition from cognitive to event-driven models and considers application of Joiner-nets for analyzing and processing event models.

The results: The authors have proposed the rules for generating events ontologies. The paper describes the algorithm for automated transition from cognitive maps to event-driven models based on ontology (using events ontologies) as one of the stages of improving two-level technology. The article introduces the developed ontologies, cognitive and event maps and event model using Joiner-nets built on their basis.

Key words:

Energy security, Semantic modeling, Ontological, Cognitive and Event modeling, Event ontology, Joiner-nets.

REFERENCES

1. Bushuev V.V., Voropay N.I., Mastepanov A.M., Shafranik Yu.K. *Energeticheskaya bezopasnost Rossii* [Energy security of Russia]. Novosibirsk, Nauka, 1998. 302 p.
2. Massel A.G. Dvukhurovnevaya tekhnologiya vychislitelnykh eksperimentov v issledovaniyakh problem energeticheskoy bezopasnosti [Two-level technology of computing experiments in energy security problems research]. *Informatsionnye i matematicheskie tekhnologii v nauke i upravlenii. Trudy XVI Baykalskoy Vserossiyskoy konferentsii* [Mathematical and Informational Technologies in Science and Management. Proc. XVI Baykal Russian Conference]. Irkutsk, ISEM SB RAS, 2011. P. III, pp. 70–76.
3. Massel L.V., Massel A.G. Intellectualnye vychisleniya v issledovaniyakh napravleniy razvitiya energetiki [Intelligent computing in studying directions of power engineering development]. *Bulletin of the Tomsk Polytechnic University*, 2012, vol. 321, no. 5, pp. 135–141.
4. Massel L.V. Integratsiya semanticheskogo i matematicheskogo modelirovaniya v issledovaniyakh problem energeticheskoy bezopasnosti [Integration of semantic and mathematical modeling when studying the problems of energy security]. *Modelirovanie-2012. Trudy mezhdunarodnoy konferentsii* [Proceedings of the International Conference Simulation-2012]. Kiev, IPME NAS, 2012. pp. 270–273.
5. Massel L.V., Massel A.G. Semanticheskie tekhnologii na osnove integratsii ontologicheskogo, kognitivnogo i sobyitnogo modelirovaniya [Semantic technologies based on integration of ontological, cognitive and event simulation]. *OSTIS-2013. Materialy III Mezhdunarodnoy nauchno-tekhnicheskoy konferentsii* [Proc. III International Scientific and Technical Conference. OSTIS-2013]. Belarus, Minsk, BSUIR, 2013. pp. 247–250.
6. Gavrilova T.A., Khoroshevsky V.F. *Bazy znaniy intellektualnykh sistem* [Knowledge base of intelligent systems]. St. Petersburg, Piter, 2001. 384 p.
7. Tuzovsky A.F., Chirikov S.V., Yampolsky V.Z. *Sistemy upravleniya znaniyami (metody i tekhnologii)* [Knowledge management systems (methods and techniques)]. Tomsk, NTL Publ., 2005. 260 p.
8. Trakhtengerts E.A. *Kompyuternaya podderzhka prinyatiya resheniy* [Computer support of decision-making]. Moscow, SINTEG Publ., 1998. 376 p.
9. Kosko B. Fuzzy Cognitive Maps. *International Journal of Man-Machine Studies*, 1986, no. 24, pp. 65–75.
10. Stolyarov L.N. Filosofiya sobyitnogo modelirovaniya na primere stseneriya energeticheskoy katastrofy [Philosophy of event simulation by the example of energy catastrophe script]. *Informatsionnye tekhnologii v nauke, obrazovanii, telekommunikatsii i biznese. Trudy Mezhdunarodnoy konferentsii* [Proc. of the International Conference. Information Technologies in science, education, telecommunications, and business]. Ukraine, Gursuf, 2010. pp. 197–200.
11. Massel A.G. Metodologicheskii podkhod k organizatsii intellektualnoy podderzhki issledovaniy problemy energeticheskoy bezopasnosti [Methodological approach to organization of intellectual support of researching the energy security problem]. *Informatsionnye tekhnologii – Information Technologies*, 2010, no. 9, pp. 32–36.
12. Massel L.V. Primenenie ontologicheskogo, kognitivnogo i sobyitnogo modelirovaniya dlya analiza razvitiya i posledstviy chrezvychaynykh situatsiy v energetike [Application of ontological, cognitive and event simulation to analyze the development and consequences of emergency situations in power industry]. *Problemy bezopasnosti i chrezvychaynykh situatsiy – Problems of safety and emergencies*, 2010, no. 2, pp. 34–43.

13. Arshinsky V.L. Metodicheskiy podkhod k sobytiynomu modelirovaniyu v issledovaniyakh energeticheskoy bezopasnosti [Methodical approach to event simulation in energy security research]. *Informatsionnye i matematicheskie tekhnologii v nauke i upravlenii. Trudy XV Baykalskoy Vserossiyskoy konferentsii* [Proc. XV Baikal Russian Conference. Mathematical and Informational Technologies in Science and Management]. Irkutsk, ESI SB RAS, 2010. P. III, pp. 120–129.
14. Axelrod R. *Structure of decision*. Princeton, New Jersey, Princeton University Press, 1976. 404 p.
15. Novik K.V. *Set avtomatov dlya modelirovaniya asinkhronnogo vzaimodeystviya protsessov. Avtoreferat dis. kand. nauk* [Network automats for modelling asynchronous communication processes. Abstract. Cand diss.]. Moscow, 2006. 22 p.
16. Stolyarov L.N., Novik K.V. Joiner-set dlya modelirovaniya vzaimodeystviyushchikh parallelnykh protsessov [Joiner-interacting network for modelling concurrent processes]. *Informatsionnye i matematicheskie tekhnologii. Sbornik nauchnykh trudov* [Information and mathematical techniques. Scientific papers]. Moscow, Moscow Institute of Physics and Technology Press, 2004. pp. 81–97.
17. Stolyarov L.N., Novik K.V. Realizatsiya parallelnykh protsessov s pomoshchyu setey Joiner-net [Implementation of parallel processes with a help of Joiner-net]. *Informatsionnye i matematicheskie tekhnologii. Sbornik nauchnykh trudov* [Informational and Mathematical technologies. Scientific papers]. Irkutsk, ESI SB RAS, 2004. pp. 11–14.
18. Anisimov M.M. Upravlenie sobytiynymi setyami [Event-management networks]. *Informatsionnye i matematicheskie tekhnologii v nauke i upravlenii. Trudy XIV Baykalskoy Vserossiyskoy konferentsii* [Proc. XIV Baikal All Russia Conference. Information technology and mathematical science and management]. Irkutsk, ESI SB RAS, 2009. P. 3, pp. 238–240.

УДК 004.9

НЕЙРОСЕТЕВОЙ ПАРАЛЛЕЛЬНЫЙ АЛГОРИТМ СЛЕЖЕНИЯ ЗА ОБЪЕКТОМ В РЕАЛЬНОМ ВРЕМЕНИ

Тарков Михаил Сергеевич,

канд. техн. наук, ст. науч. сотр. лаборатории физических основ интегральной микроэлектроники Института физики полупроводников им. А.В. Ржанова СО РАН, Россия, 630090, Новосибирск, пр. Ак. Лаврентьева, 13. E-mail: tarkov@isp.nsc.ru

Дубынин Сергей Владимирович,

магистрант Новосибирского государственного университета, Россия, 630090, г. Новосибирск, ул. Пирогова, 2. E-mail: dubyninsergey@gmail.com

Актуальность исследования обусловлена необходимостью разработки программных средств слежения за объектами в реальном масштабе времени.

Цель работы: Создание алгоритма слежения за объектом в кадре в реальном масштабе времени.

Методы исследования: Параллельная реализация сигмоидальной нейронной сети на графическом процессоре, замеры временных характеристик параллельного алгоритма и его оптимизация.

Результаты: Предложена реализация на графическом процессоре (GPU) нейросетевого алгоритма слежения за объектом, спецификой которого является использование при обучении нейронной сети задачника, устанавливающего однозначное соответствие обрабатываемого кадра в видеопотоке координатам центра объекта в кадре. Благодаря использованию GPU удается решить задачу слежения в реальном масштабе времени (25 кадров в секунду) при размерах обрабатываемого кадра до 1280×960. Алгоритм основан на использовании многослойного персептрона и имеет ряд параметров, которые определены экспериментально. Одним из таких параметров является число нейронов скрытого слоя. В связи с реализацией алгоритма на GPU рассмотрены числа нейронов, кратные 16. В экспериментах установлено, что 16 и 32 нейрона не могут обеспечить даже малой степени запоминания образов, 48 нейронов справлялись с обучением только на малых обучающих выборках, 64 нейрона обеспечили хорошую степень запоминания образов и скорость работы. Дальнейшее увеличение числа нейронов приводит только к уменьшению скорости работы нейронной сети и ее обучения.

Также заслуживает внимания частота, с которой нужно брать кадры из видеозаписи, чтобы эффективно обучить нейронную сеть. Экспериментально установлено, что на частоте выборки одного кадра из десяти сумма максимальных отклонений по обеим координатам равна 50 при размерах объекта 300×300; дальнейшее увеличение частоты кадров лишь замедляет процесс обучения, не давая существенного выигрыша в качестве.

Получены ускорения процесса слежения в 10 раз по сравнению с центральным процессором персонального компьютера. Процесс обучения нейронной сети ускорился в среднем только в 2 раза. Это обусловлено необходимостью транспонирования матриц весов при реализации обучения нейронной сети на GPU.

Для реализации параллельного алгоритма использована программно-аппаратная архитектура CUDA, позволяющая производить вычисления с использованием графических процессоров NVIDIA, поддерживающих технологию GPGPU (произвольных вычислений на видеокартах). Для предварительной обработки изображений и вывода информации использовалась библиотека компьютерного зрения OpenCV.

Ключевые слова:

Слежение за объектом, нейронная сеть, параллельные вычисления, графический процессор, CUDA.

Введение

В настоящее время разработано большое количество приложений для видеобработки с использованием графических процессоров [1–9], которые обеспечивают массовый параллелизм обработки данных в реальном масштабе времени. Одной из самых распространенных задач в этой сфере является слежение за объектами [3–10]. Использование алгоритмов слежения за объектами используется для различных целей: выявление определенных движущихся целей и слежение за ними в военной технике; фиксация номерных знаков автомобилей, превышающих скорость; наложение различных визуальных эффектов на видеозапись и прочее.

Для реализации слежения за объектами разработано множество методов и алгоритмов, но зачастую они являются узкоспециализированными и устойчивы лишь на определенном типе видеозаписей. В хороших условиях (при четких изображениях, при низкой скорости перемещения объекта и прочее) эти алгоритмы работают достаточно хорошо, но при возникновении помех, при увеличении скорости объекта и уменьшении его размеров, алгоритмы дают сбой. Ко всему прочему, алгоритмы слежения за объектами являются достаточно трудоемкими, что вынуждает сжимать обрабатываемый кадр или как-то иначе упрощать обрабатываемую информацию. В связи с этим и возникает проблема разработки эффективных робастных алгоритмов слежения за объектами.

В решении задачи слежения на графических процессорах активно используются нейронные сети [6–8] и другие математические модели [4, 5, 9]. В данной работе предложена реализация на графическом процессоре нейросетевого алгоритма слежения за объектом [10], спецификой которого является использование при обучении нейронной сети заданника, задающего однозначное соответствие обрабатываемого кадра в видеопотоке координатам центра объекта в кадре. Благодаря использованию графического ускорителя удастся решить задачу слежения в реальном масштабе времени без уменьшения размеров обрабатываемого кадра.

Постановка задачи

Существует множество различных систем слежения за объектами. Эти системы используют разные алгоритмы и работают на различных входных данных. Наиболее эффективные реализации используют сложное дорогостоящее оборудование: несколько видеокамер, запись цветного видео или видео в инфракрасном спектре. С одной стороны, цветное изображение дает возможность использовать больше различных алгоритмов, но, с другой стороны, эти алгоритмы достаточно трудоемки и не всегда могут корректно работать (например, при слабом освещении). Алгоритмы для монохромных изображений могут использовать более доступную технику, но являются менее эффективными и зачастую используют небольшие разрешения кадра.

Алгоритм [10] работает с монохромными изображениями малого разрешения (320×240), которое во время предобработки данных снижается до 80×60 . Однако можно создать быстрый алгоритм, работающий с большими разрешениями кадра в реальном масштабе времени на достаточно дешевом оборудовании. Ключ к такому решению лежит в использовании графических карт как устройств, позволяющих выполнять массивно-параллельные вычисления [11, 12].

Общая задача заключается в слежении за объектом на изображениях и видео, то есть в определении координат центра объекта на основе информации, получаемой из изображения. Так как одной из целей работы является слежение за объектом на видео в реальном масштабе времени, то на алгоритм определения координат центра объекта накладывается ограничение по быстродействию: на обработку одного кадра должно уходить не более $1/25$ секунды. При этом требуется обрабатывать кадр целиком без потерь информации. Отметим, что в [10] для повышения скорости работы алгоритма использовался лишь каждый четвертый пиксель изображения.

Нейронная сеть и её обучение

Для обработки изображений использовалась сигмоидальная сеть прямого распространения с одним скрытым слоем [13–18]. Искусственная нейронная сеть представляет собой систему взаимосвязанных простых процессоров – нейронов. Каждый нейрон получает входные сигналы w_i , $i=0,1,\dots,N$ и порождает выходной сигнал $y=f(u)$, где $f(u)$ – нелинейная функция активации;

$u = \sum_{i=0}^N w_i x_i$ – активация нейрона, w_i , $i=0,1,\dots,N$ – весовые коэффициенты нейрона, w_0 – величина порога, $x_0=1$.

Обучение нейрона заключается в выборе весовых коэффициентов w_i таким образом, что выходной сигнал y совпадает с требуемой величиной d . Обучение с учителем использует набор обучающих примеров, то есть множество пар вида (x, d) , где x – вектор входных сигналов.

Для униполярного сигмоидального нейрона функция активации задана выражением

$$f(u) = \frac{1}{1 + \exp(-\beta u)},$$

где β – параметр функции активации. В трехслойной сигмоидальной сети, первый слой нейронов содержит входные сигналы, скрытый слой нейронов получает входные сигналы и преобразует их, чтобы передать выходному слою нейронов. Выходной слой активизируется и порождает выходные сигналы нейронной сети.

Благодаря дифференцируемости функции активации для обучения нейронной сети можно использовать градиентные методы оптимизации. В частности используется метод наискорейшего

спуска (метод обратного распространения ошибки), согласно которому модификация вектора весов производится в направлении отрицательного градиента целевой функции

$$E(w) = \frac{(y - d)^2}{2},$$

то есть

$$w(t+1) = w(t) - \alpha \nabla E(w(t)),$$

где $\alpha \in (0, 1]$ – коэффициент (шаг) обучения.

Перед обработкой производилась нормировка яркости монохромных изображений в диапазон $[0, 1]$. В результате исследования поведения сети при разном числе нейронов в скрытом слое принято решение использовать 64 нейрона, что обеспечивает достаточную скорость работы алгоритма и его точность. В выходном слое присутствует всего два нейрона, каждый из которых дает на выходе одну из координат искомого объекта. Для обучения использовался алгоритм обратного распространения ошибки с коррекцией величины шага. В качестве задачника использовался набор изображений с известными координатами центра объекта.

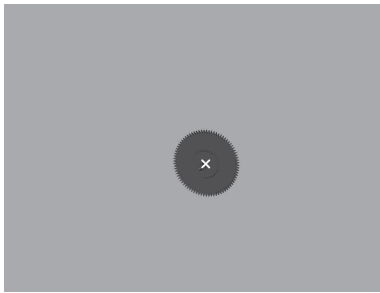


Рис. 1. Пример обучающего изображения

Для создания задачника использовалась программа Autodesk Maya 2011 [19], в которой создана трехмерная модель шестеренки (рис. 1).

Для этой модели с помощью формул, зависящих от номера кадра, заданы следующие параметры: координаты объекта в кадре, углы поворота объекта по трем осям и размер объекта. Затем с помощью программы получены изображения с объектом и видеозапись для тестирования процесса слежения за объектом. Для создания файла с координатами центра объекта на разных изображениях создана программа, вычисляющая координаты центра по формулам, описывающим его движение, и записывающая эти координаты в файл.

Для реализации параллельного алгоритма использована программно-аппаратная архитектура CUDA [11, 12], позволяющая производить вычисления с использованием графических процессоров NVIDIA, поддерживающих технологию GPGPU (произвольных вычислений на видеокартах). Для предварительной обработки изображений и вывода информации по ходу обучения использовалась библиотека компьютерного зрения OpenCV [20].

Реализация нейронной сети на основе технологии CUDA

CUDA (Compute Unified Device Architecture) – это интегрированная среда, позволяющая разрабатывать программы на языке C/C++, которые запускают параллельное исполнение специальных функций ядра на графической карте, поддерживающей технологию CUDA. Эта графическая карта в данном контексте называется устройством, а компьютер, на котором установлено устройство, называется хостом (host). Функции ядра выполняются параллельно нитями, которые объединяются в блоки одинакового размера. Блоки и нити внутри блоков формируют сетку, которая в ходе выполнения функции отображается на мультипроцессоры GPU (Graphic Processing Unit) и их (скалярные) процессоры, соответственно.

Функции ядра могут использовать различные типы памяти устройства: регистры, разделяемая (shared), текстурная и константная памяти. Текстурная и константная памяти являются малыми по объему, но быстрыми памятьми устройства. Локальная и глобальная памяти устройства работают значительно медленнее, но вмещают значительно больше данных (обычно до 2-х Гбайт). Регистры и локальная память доступны только текущей нити, разделенная память доступна каждому блоку, а константная, текстурная и глобальная памяти доступны всем нитям. Передача данных между блоками возможна только через глобальную память.

Поскольку нейроны одного слоя могут выполнять вычисления независимо друг от друга, принято решение реализовать параллельные версии основных функций нейронной сети и подстройки ее весов. Созданы дополнительные функции, обеспечивающие взаимосвязь параллельного алгоритма с интерфейсом и последовательной версией. Например, программа позволяет обучить нейронную сеть с использованием CUDA, а обрабатывать выбранные пользователем изображения с помощью последовательной реализации алгоритма и наоборот.

Функцию, выполняемую нейронной сетью, можно разбить на три части:

- 1) Входной вектор x умножить на матрицу весов скрытого слоя W_1 , после чего к результату прибавить вектор смещений скрытого слоя b_1 :

$$a_1 = W_1 \cdot x + b_1. \quad (1)$$

- 2) К полученному вектору a_1 применить функцию активации f :

$$u = f(a_1). \quad (2)$$

- 3) Полученный вектор u умножить на матрицу весов выходного слоя W_2 и прибавить вектор смещений b_2 нейронов выходного слоя:

$$y = W_2 \cdot u + b_2. \quad (3)$$

В результате получаем три процедуры, содержащие большое число операций, которые можно выполнить параллельно. Для этих процедур написаны функции ядра GPU, которые запускаются параллельно на множестве нитей.

В основе функций (1) и (3) лежит операция произведения матрицы на вектор. Функция (2) проста и не требует оптимизации. В изначальной версии реализации функции (1) каждая нить производит перемножение одной строки матрицы W_1 на вектор x . Число нитей фиксировано и соответствует числу нейронов скрытого слоя. Каждая нить выполняет операции умножения и сложения. Обращения нитей к глобальной памяти GPU значительно замедляют их работу. К тому же профилировщик Compute Visual Profiler показал, что при этом не используется свойство GPU, позволяющее объединять запросы к памяти: когда следующие друг за другом нити обращаются в следующие друг за другом ячейки памяти, эти обращения могут быть объединены в одно (warp) и вместо группы обращений происходит по сути одно. Максимальное число нитей, входящих в warp, равно 16 (32 для более новых моделей GPU).

Чтобы устранить указанные недостатки, решено:

- 1) Увеличить число нитей, чтобы каждая нить перемножала только векторы из 16 чисел. В результате увеличивается полезная нагрузка на GPU, больше нитей выполняется параллельно, появляется необходимость новой функции ядра, суммирующей результаты работы нитей.
- 2) Использовать вместо глобальной памяти разделяемую память (shared memory). Эта память выделяется каждому блоку нитей и может использоваться всеми нитями блока. Для этого надо загрузить фрагмент глобальной памяти, используемый всеми нитями блока, в разделяемую память. Каждая нить делает лишь одно обращение к соответствующей ячейке глобальной памяти, скопировав значение в разделяемую память, а остальные данные нить сможет получить из разделяемой памяти.
- 3) Транспонировать матрицу W_1 , чтобы обращения к элементам матрицы, находящимся в глобальной памяти, объединялись в warp. Изначально одна нить работала с вектором-строкой, элементы которой располагаются в разных сегментах памяти. Транспонирование матрицы позволяет нити работать с вектором-столбцом, что объединяет подряд идущие нити в warp.

В результате этих модификаций суммарное время выполнения функции ядра для первого блока по всем запускам в течение работы программы уменьшилось с 30 % времени работы GPU до 1 % плюс 3 % на суммирующую функцию ядра, появившуюся в ходе изменений пункта 1. Все загрузки из глобальной памяти и выгрузки в глобальную память GPU являются объединенными (coalesced). Ветвления, вызванные циклом функции, и запуск блоков нитей (warps) занимают крайне мало времени.

В функции подстройки весов определенные блоки операций организованы в функции ядра, выполняющиеся параллельно на множестве нитей. Однако специфика этих блоков не позволяет

сильно ускорить их и вынуждает использовать транспонирование матриц перед вызовом функции подстройки весов и после нее, чтобы на следующей итерации функция работы нейронной сети получила транспонированные матрицы весов. Без транспонирования наблюдается замедление работы функции подстройки весов.

Алгоритм имеет ряд параметров, которые определены экспериментально. Одним из таких параметров является число нейронов скрытого слоя. В связи со спецификой параллельной реализации алгоритма рассмотрены числа нейронов, кратные 16 (размер warp равен 16). В экспериментах установлено, что 16 и 32 нейрона не могут обеспечить даже малой степени запоминания образов, 48 нейронов справлялись с обучением только на малых обучающих выборках, 64 нейрона обеспечили хорошую степень запоминания образов и скорость работы. Дальнейшее увеличение числа нейронов приводит только к уменьшению скорости работы нейронной сети и ее обучения.

Кроме вышеуказанных параметров, внимания также заслуживает частота, с которой нужно брать кадры из видеозаписи, чтобы эффективно обучить нейронную сеть. Экспериментально установлено (рис. 2), что:

- 1) на частоте выборки одного кадра из десяти сумма максимальных отклонений по обеим координатам равна 50 при размерах объекта 300×300 ;
- 2) дальнейшее увеличение частоты кадров лишь замедляет процесс обучения, не давая существенного выигрыша в качестве.

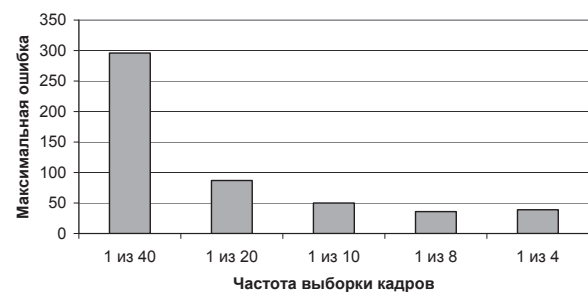


Рис. 2. Зависимость ошибки определения координат объекта от частоты выборки кадров

Тестирование параллельной и последовательной реализаций проводилось на компьютере со следующими характеристиками: CPU – AMD Athlon 7750, 2 ядра по 2,7 ГГц, GPU – NVIDIA GeForce 9800 GT. 512MB 256 bit, количество потоковых процессоров – 112. Разработка параллельной версии велась с использованием Cuda Toolkit 4.1 и Cuda Toolkit 4.0 (использован профилировщик этой версии). Главным параметром, по которому проводилось сравнение, является скорость работы нейронной сети, то есть быстродействие основной функции слежения за объектом.

Из рис. 3 следует, что параллельная реализация нейронной сети на GPU позволяет увеличить линейные размеры обрабатываемых изображений

в 4 раза (с 320×240 до 1280×960). Из рис. 3, 4 следует, что процесс обработки нейронной сетью последовательности кадров ускорился в среднем в 10 раз. Процесс обучения ускорился в среднем только в 2 раза (рис. 5). Это обусловлено необходимостью транспонирования матриц весов при реализации обучения нейронной сети на GPU.

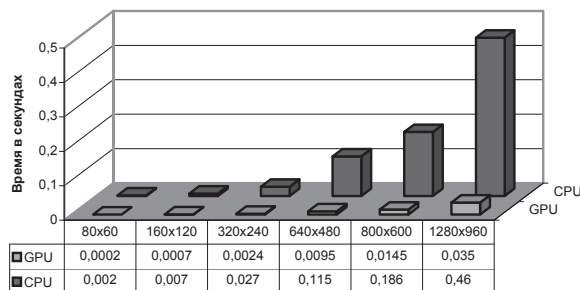


Рис. 3. Времена работы нейронной сети на CPU и с использованием GPU

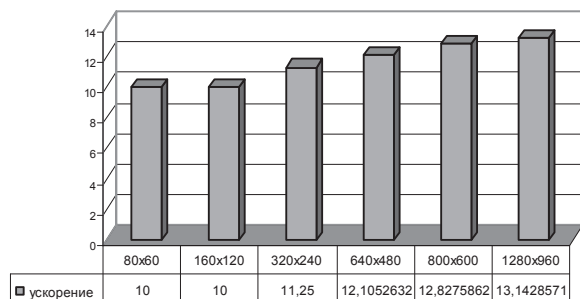


Рис. 4. Ускорение параллельной реализации алгоритма по сравнению с последовательной

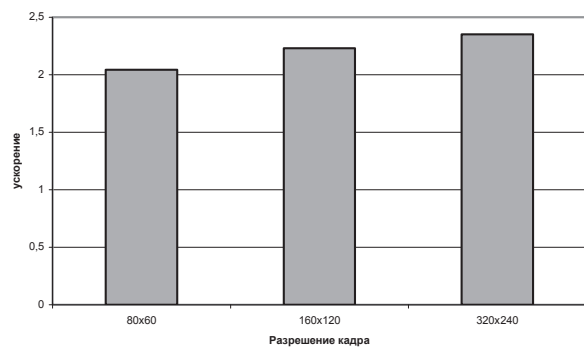


Рис. 5. Ускорение обучения нейронной сети при различных разрешениях кадра

Заключение

Реализованы алгоритмы слежения за объектами в реальном масштабе времени, основанные на нейронных сетях с обучением алгоритмом обратного распространения ошибки. Предложен алгоритм слежения за объектами, использующий массивно-параллельные вычисления с использованием графического процессора, проведены исследования и оптимизация параметров алгоритмов. Получены ускорения процесса слежения в 10 раз и процесса обучения в 2 раза. Экспериментально определены максимальные разрешения кадра, пригодные для слежения в реальном времени, и оптимальная частота взятия кадров из видеозаписи в обучающую выборку. Возможно развитие работы в направлении создания алгоритмов, обучающихся в реальном времени, то есть уже во время слежения за объектом.

СПИСОК ЛИТЕРАТУРЫ

1. Real-time image segmentation on a GPU / A. Abramov, T. Kulvicius, F. Worgotter, B. Dellen // Facing the Multicore-Challenge. Lecture Notes in Computer Science. – Berlin: Springer, 2010. – V. 6310. – P. 131–142.
2. NeuFlow: Dataflow Vision Processing System-on-a-Chip / P.-H. Pham, D. Jelaca, C. Farabet, B. Martini, Y. LeCun, E. Culurciello // Proc. of 2012 IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS). – West Lafayette, IN, USA, 5–8 Aug., 2012. – P. 1044–1047.
3. Di Salvo R., Pino C. Image and Video Processing on CUDA: State of the Art and Future Directions // MACMESE'11: Proc. of the XIII WSEAS International conference on Mathematical and computational methods in science and engineering. – Catania, Italy, 2011. – P. 60–66.
4. Fully automatic extraction of salient objects from videos in near real-time / K. Akamine, K. Fukuchi, A. Kimura, S. Takagi // Computer Vision and Pattern Recognition. – 2010. – С. 1–23. URL: <http://arxiv.org/pdf/1008.0502.pdf> (дата обращения: 30.03.2014).
5. Real-time GPU color-based segmentation of football players / M.A.M. Laborda, E.F.T. Moreno, J.M. del Rincon, J.E.H. Jara-ba // Journal of Real-Time Image Processing. – 2012. – V. 7. – Iss. 4. – P. 267–279.
6. OpenCL Implementation of a Color Based Object Tracking / M. Jocić, D. Obradović, Z. Konjović, D. Tertei // Proc. of III International Conference on Information Society Technology (ICIST). – Novi Sad, Serbia, 2013. – P. 7–11.
7. Ugolotti R., Nashed Y.S.G., Cagnoni S. Real-Time GPU Based Road Sign Detection and Classification // Proc. of PPSN. Lecture Notes in Computer Science. – Berlin: Springer, 2012. – V. 7491. – P. 1. – P. 153–162.
8. Pleshkova S. Spiking Neural Networks for Real-Time Infrared Images Processing in Thermo Vision Systems // Recent Researches in Circuits and Systems. – 2012. – P. 183–187. URL: <http://hgpu.org> (дата обращения: 30.03.2014).
9. Ferreira J.F., Lobo J., Dias J. Bayesian real-time perception algorithms on GPU. Real-time implementation of Bayesian models for multimodal perception using CUDA // Journal of Real-Time Image Processing. – 2011. – V. 6. – Iss. 3. – P. 171–186.
10. Design and Implementation of a Neural Network for Real-Time Object Tracking / J. Ahmed, M.N. Jafri, J. Ahmad, M.I. Khan // World Academy of Science, Engineering and Technology. – 2005. – № 6. – P. 209–212.
11. Боресков А.В., Харламов А.А. Основы работы с технологией CUDA. – М.: ДМК Пресс, 2010. – 232 с.
12. Сандерс Дж., Кэндрот Э. Технология CUDA в примерах: введение в программирование графических процессоров. – М.: ДМК Пресс, 2011. – 232 с.
13. Осовский С. Нейронные сети для обработки информации. – М.: Финансы и статистика, 2002. – 344 с.
14. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. – М.: Горячая линия – Телеком, 2004. – 452 с.
15. Хайкин С. Нейронные сети: полный курс. – М.: Вильямс, 2006. – 1104 с.

16. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. – Новосибирск: Наука, 1996. – 276 с.
17. Галушкин А.И. Нейронные сети: основы теории. – М.: Горячая линия – Телеком, 2010. – 496 с.
18. Тарков М.С. Нейрокомпьютерные системы. – М.: Интернет-Ун-т Информ. Технологий: Бином. Лаборатория знаний, 2006. – 142 с.
19. Maya – 3D Animation – Autodesk. URL: <http://usa.autodesk.com/maya/> (дата обращения: 30.03.2014).
20. Bradski G., Kaehler A. Learning OpenCV. – USA, Sebastopol, CA: O'Reilly Media, Inc., 2008. – 556 p.

Поступила 31.03.2014 г.

UDC 004.9

NEURAL NETWORK PARALLEL ALGORITHM FOR REAL-TIME OBJECT TRACKING

Mikhail S. Tarkov,

Cand. Sc., Laboratory of Physical Foundations of Integrated Microelectronics,
A.V. Rzhanov Institute of Semiconductor Physics SB RAS, 13, Lavrentiev avenue,
Novosibirsk, 630090, Russia. E-mail: tarkov@isp.nsc.ru

Sergey V. Dubynin,

Novosibirsk State University, 2, Pirogov street, Novosibirsk,
630090, Russia. E-mail: dubyninsergey@gmail.com

The urgency of the discussed issue is caused by the need to provide software for tracking objects in real time.

The main aim of the study: to create an object-tracking algorithm in the frame in real time.

The methods used in the study: parallel implementation of the sigmoid neural network on the GPU, measuring the temporal characteristics of the parallel algorithm and its optimization.

The results: The authors have proposed implementation of a neural network algorithm on graphic processor (GPU) for tracking an object in a video frame. The specific character of the algorithm is the use of a training set which establish correspondence between the video frame and the object center coordinates in this frame when training a neural network. Owing to GPU application the tracking problem can be solved in real time (25 frames per second) at the processed frame sizes up to 1280×960.

The algorithm is based on the use of multilayer perceptron and has a number of parameters, which are determined experimentally. One of such parameters is the number of the hidden layer neurons. Due to the algorithm implementation on GPU the authors considered the number of neurons multiple 16. It was determined experimentally that 16 and 32 neurons cannot provide even a small degree of memorizing images, 48 neurons cope with learning only small training samples, and 64 neurons provided a good degree of memorizing images and speed. Further increase in the number of neurons results only in reducing speed of the neural network functioning and its training.

The frequency which is required for taking pictures from a video to train effectively a neural network is worth noticing as well. It is found out experimentally that at a sampling rate of one frame of ten, the sum of the maximum deviations in coordinates is 50, when the object size is 300×300; further increase of the frame rate slows down the process of training without significant gain in quality.

The authors obtained the tracking accelerating by 10 times in comparison with the CPU of a personal computer. The neural network training is accelerated only 2 times on average. This is caused by the need to transpose the weight matrices when implementing the neural network training on the GPU.

To implement the parallel algorithm, the hardware and software architecture CUDA is used. It allows computation on graphics processors NVIDIA, supporting GPGPU technology (general purpose computations on GPU). For preliminary image processing and data output the computer vision library OpenCV is used.

Key words:

Object tracking, neural network, parallel computing, GPU, CUDA.

REFERENCES

1. Abramov A., Kulvicius T., Worgotter F., Dellen B. Real-time image segmentation on a GPU. *Lecture Notes in Computer Science*. Berlin, Springer, 2010, vol. 6310, pp. 131–142.
2. Pham P.-H., Jelaca D., Farabet C., Martini B., LeCun Y., Culurciello E. NeuFlow: Dataflow Vision Processing System-on-a-Chip. *Proc. of 2012 IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS)*. West Lafayette, IN, USA, 2012, 5–8 Aug., 2012, pp. 1044–1047.
3. Di Salvo R., Pino C. Image and Video Processing on CUDA: State of the Art and Future Directions. *MACMESE'11: Proc. of the XI–II WSEAS International conference on Mathematical and computational methods in science and engineering*. Catania, Italy, 2011, pp. 60–66.
4. Akamine K., Fukuchi K., Kimura A., Takagi S. Fully automatic extraction of salient objects from videos in near real-time. *Computer Vision and Pattern Recognition*, 2010, pp. 1–23. Available at: <http://arxiv.org/pdf/1008.0502.pdf> (accessed 30 March 2014).
5. Laborda M.A.M., Moreno E.F.T., Del Rincon J.M., Jaraba J.E.H. Real-time GPU color-based segmentation of football players. *Journal of Real-Time Image Processing*, 2012, vol. 7, Iss. 4, pp. 267–279.
6. Jocić M., Obradović D., Konjović Z., Tertei D. OpenCL Implementation of a Color Based Object Tracking. *Proc. of III International Conference on Information Society Technology (ICIST)*. Novi Sad, Serbia, 2013, pp. 7–11.
7. Ugolotti R., Nashed Y.S.G., Cagnoni S. Real-Time GPU Based Road Sign Detection and Classification. *Proc. of PPSN. Lecture*

- Notes in Computer Science*. Berlin, 2012, vol. 7491, pp. 153–162.
8. Pleshkova S. Spiking Neural Networks for Real-Time Infrared Images Processing in Thermo Vision Systems. *Recent Researches in Circuits and Systems*, 2012, pp. 183–187. Available at: <http://hgpu.org> (accessed 30 March 2014).
9. Ferreira J.F., Lobo J., Dias J. Bayesian real-time perception algorithms on GPU. Real-time implementation of Bayesian models for multimodal perception using CUDA. *Journal of Real-Time Image Processing*, 2011, vol. 6, Iss. 3, pp. 171–186.
10. Ahmed J., Jafri M.N., Ahmad J., Khan M.I. Design and Implementation of a Neural Network for Real-Time Object Tracking. *World Academy of Science, Engineering and Technology*, 2005, no. 6, pp. 209–212.
11. Borekov A.V., Kharlamov A.A. *Osnovy raboty s tekhnologiyey CUDA* [The basics of working with technology CUDA]. Moscow, DMK Press, 2010. 232 p.
12. Sanders Dzh., Kendrot E. *Tekhnologiya CUDA v primerakh: vvedenie v programmirovaniye graficheskikh protsessorov* [CUDA by Example: An introduction to general-purpose GPU programming]. Moscow, DMK Press, 2011. 232 p.
13. Osovskiy S. *Neyronnye seti dlya obrabotki informatsii* [Neural networks for information processing]. Moscow, Finansy i statistika, 2002. 344 p.
14. Rutkovskaya D., Pilinskiy M., Rutkovskiy L. *Neyronnye seti, geneticheskie algoritmy i nechetkie sistemy* [Neural networks, genetic algorithms and fuzzy systems]. Moscow, Goryachaya liniya – Telekom, 2004. 452 p.
15. Khaykin S. *Neyronnye seti: polny kurs* [Neural networks: A comprehensive foundation]. Moscow, Vilyams Publ., 2006. 1104 p.
16. Gorban A.N., Rossiev D.A. *Neironnye seti na personalnom kompyutere* [Neural networks on PC]. Novosibirsk, Nauka Publ., 1996. 276 p.
17. Galushkin A.I. *Neyronnye seti: osnovy teorii* [Neural networks: basic theory]. Moscow, Goryachaya liniya – Telekom, 2010. 496 p.
18. Tarkov M.S. *Neyrokomp'yuternyye sistemy* [Neurocomputer systems]. Moscow, Internet Universitet Informatsionnykh tekhnologiy, Binom. Laboratoriya znaniy, 2006. 142 p.
19. *Maya – 3D Animation – Autodesk*. Available at: <http://usa.autodesk.com/maya/> (accessed 30 March 2014).
20. Bradski G., Kaehler A. *Learning OpenCV*. USA, Sebastopol, CA, O'Reilly Media, Inc., 2008. 556 p.

УДК 004.931

ПРИМЕНЕНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ВЫДЕЛЕНИЯ И РАСПОЗНАВАНИЯ АВТОМОБИЛЬНЫХ НОМЕРНЫХ ЗНАКОВ НА ИЗОБРАЖЕНИЯХ СО СЛОЖНЫМ ФОНОМ

Друки Алексей Алексеевич,аспирант кафедры вычислительной техники Института кибернетики ТПУ,
Россия, 634050, г. Томск, пр. Ленина, д. 30. E-mail: druki2008@yandex.ru

Актуальность исследования обусловлена необходимостью разработки методов, алгоритмов и программ, обеспечивающих повышение эффективности распознавания автомобильных номерных знаков на изображениях со сложным фоном.

Цель работы: повышение эффективности распознавания автомобильных номерных знаков на изображениях со сложным фоном за счёт создания методов, алгоритмов и программ, инвариантных к смещениям, поворотам, смене ракурса и прочим искажениям входных данных.

Методы исследования: для решения поставленных задач использованы методы искусственного интеллекта, идентификации и распознавания образов на изображениях, теории искусственных нейронных сетей, сверточные нейронные сети, эволюционные алгоритмы, математическое моделирование, теория вероятности и математическая статистика с использованием программного обеспечения Visual Studio и MatLab.

Результаты: разработана программная система, позволяющая распознавать автомобильные номерные знаки на изображениях со сложным фоном. Для выделения области расположения символов на изображении предложена сверточная нейронная сеть, состоящая из 7 слоев. Для выделения отдельных символов используется алгоритм, основанный на гистограммах средней интенсивности пикселей. Для распознавания символов реализована сверточная нейронная сеть, состоящая из 6 слоев. Представленная программная система позволяет распознавать автомобильные номерные знаки под большими углами наклона по вертикали, горизонтали и на плоскости при достаточно высоком быстродействии.

Ключевые слова:

Обработка изображений, искусственный интеллект, распознавание символов, нейронные сети, гистограммы средней интенсивности.

Введение

Одной из самых сложных задач в информационных технологиях является обработка и распознавание объектов на изображениях. О важности этой проблемы говорит тот факт, что исследования по распознаванию объектов, анализу изображений и речи включены в перечень приоритетных направлений развития науки и техники и критических технологий федерального уровня [1].

Современные методы распознавания символов на изображениях используются для решения широкого круга задач, таких как распознавание текста, изображений маркировки на поверхностях различных объектов и др.

Процесс распознавания символов на изображениях со сложным фоном можно разделить на 3 этапа:

1. Выделение области расположения символов на изображениях.
2. Выделение отдельных символов.
3. Распознавание символов.

В настоящее время такие технологии реализуются тремя традиционными методами – структурным, признаковым и шаблонным. Каждый из этих методов ориентирован на свои условия применения, для которых они являются эффективными. Вместе с тем всем этим методам присущи недостатки. При регистрации изображений наибольшие искажения, влияющие на результат распознавания, вносят аффинные и проекционные искажения, возникающие в связи с изменением угла регистрации, изменением масштаба, погодными условиями. Также наличие посторонних объектов на изображениях со сложным фоном существенно

снижают надежность распознавания методами, используемыми в современных системах распознавания автомобильных номерных знаков на изображениях и видеопоследовательностях [2, 3].

Эти недостатки особенно ярко проявились при масштабной эксплуатации программно-технологических систем, использующих в своей основе эти методы. Практически у всех систем распознавания символов на изображениях точностные характеристики резко падают и становятся ниже технологически приемлемых при различных искажениях входного изображения. Вместе с тем технологические условия получения информации о маркировке не позволяют полностью устранить эти искажения [4, 5].

Анализ методов выделения и распознавания объектов на изображениях показал, что для решения данной задачи эффективно использовать искусственные нейронные сети в связи с тем, что они являются слабо чувствительными к искажениям входного сигнала, а также обеспечивают возможность получения классификатора, хорошо моделирующего сложную функцию распределения изображений символов, тем самым увеличивая скорость и точность решения задачи по сравнению с остальными методами.

Разработка сверточной нейронной сети для выделения области расположения символов на изображениях

Наиболее часто в задачах распознавания и идентификации изображений используются классические нейросетевые архитектуры (многослой-

ный перцептрон, сети с радиально-базисной функцией и др.), но из анализа данных работ и экспериментальных исследований следует, что применение классических нейросетевых архитектур в данной задаче неэффективно по следующим причинам:

- изображения имеют большую размерность, соответственно возрастает размер нейронной сети [6];
- большое количество параметров увеличивает вместимость системы и соответственно требует большей тренировочной выборки, увеличивает время и вычислительную сложность процесса обучения;
- для повышения эффективности работы системы желательно применять несколько нейронных сетей (обученные с различными начальными значениями синаптических коэффициентов и порядком предъявления образов), но это увеличивает вычислительную сложность решения задачи и время выполнения [7, 8];
- отсутствует инвариантность к изменениям масштаба изображения, ракурсов съёмки камеры и других геометрических искажений входного сигнала [1].

Поэтому для решения задачи выделения области расположения символов были выбраны сверточные нейронные сети, т. к. они обеспечивают частичную устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и прочим искажениям.

Каждый слой сверточной нейронной сети представляет собой набор плоскостей, состоящих из нейронов. Нейроны одной плоскости имеют одинаковые синаптические коэффициенты, ведущие ко всем локальным участкам предыдущего слоя. Каждый нейрон слоя получает входы от некоторой области предыдущего слоя (локальное рецептивное поле), т. е. входное изображение предыдущего слоя сканируется небольшим окном и пропускается сквозь набор синаптических коэффициентов, а результат отображается на соответствующий нейрон текущего слоя. Таким образом, набор плоскостей представляет собой карты характеристик, и каждая плоскость находит «свой» участки изображения в любом месте предыдущего слоя. Размер локального рецептивного поля выбирается самостоятельно в процессе разработки нейронной сети [9].

Слои делятся на 2 типа: сверточные и подвыборочные. В сверточных слоях при сканировании рецептивных поля частично наслаиваются друг на друга по принципу черепицы, в подвыборочных слоях области соседних нейронов не перекрываются. Подвыборочный слой уменьшает масштаб плоскостей путём локального усреднения значений выходов нейронов, таким образом достигается иерархическая организация. Последующие слои извлекают более общие характеристики, меньше зависящие от искажений изображения. После прохождения нескольких слоев карта признаков выражается в вектор [10, 11].

Постепенно нейронная сеть обучается выделять ключевые характеристики автомобильных номерных знаков в поступающих на вход изображениях.

Для выделения области расположения символов на изображениях была разработана сверточная нейронная сеть, состоящая из 7 слоёв (рис. 1).

Входной слой размером 28×44 нейронов состоит из 1232 нейронов, не несет какой-либо функциональной нагрузки и служит лишь для подачи входного образа в нейронную сеть.

Следом за входным слоем находится первый скрытый слой, который является свёрточным. Этот слой состоит из 6 свёрточных плоскостей. Размер каждой плоскости этого слоя равен $24 \times 40 = 960$ нейронов.

Второй скрытый слой является подвыборочным, также состоит из 6 плоскостей, каждая из которых имеет синаптическую маску размером 2×2 . Размер каждой плоскости этого слоя $12 \times 20 = 240$ нейронов, что вдвое меньше, чем размер плоскости предыдущего слоя.

Третий скрытый слой является слоем свёртки. Он состоит из 18 плоскостей размером $16 \times 8 = 128$ нейронов.

Четвертый скрытый слой является подвыборочным и состоит из 18 плоскостей размером $4 \times 12 = 48$ нейронов.

Пятый скрытый слой состоит из 18 простых сигмоидальных нейронов, по одному на каждую плоскость предыдущего слоя. Роль этого слоя состоит в обеспечении классификации, после того, как выполнено извлечение особенностей и сокращение размерности входа. Каждый нейрон этого слоя полностью связан с каждым нейроном только одной плоскости предыдущего слоя.

Шестой слой является выходным слоем. Он состоит из одного нейрона, который полностью связан со всеми нейронами предыдущего слоя.

В соответствии с решаемой задачей в структуре нейронной сети достаточно одного выхода. Выходное значение нейронной сети находится в интервале $[-1; +1]$, что соответственно означает наличие или отсутствие автомобильного номерного знака на классифицируемом изображении [12, 13].

Таким образом, при сканировании входного изображения отклики нейронной сети образуют максимумы в местоположениях автомобильных номерных знаков. Отклики находятся в диапазоне $[-1; +1]$, в соответствии с выбранной активационной функцией [14, 15].

Размер синаптической маски в сверточных слоях равен 5×5 нейронов, в подвыборочных слоях – 2×2 нейрона.

Использование принципа объединения синаптических коэффициентов дает эффект уменьшения количества настраиваемых параметров нейронной сети.

Размер свёрточной плоскости определяется в соответствии со следующим выражением:

$$w_c = w_u - K + 1, \quad (1)$$

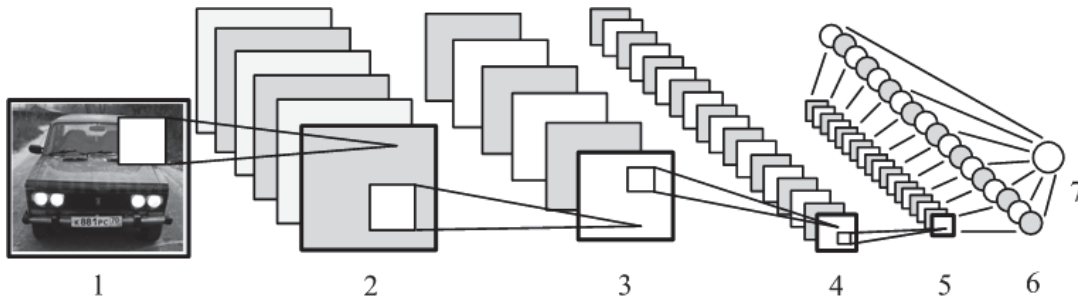


Рис. 1. Архитектура свёрточной нейронной сети для выделения области расположения символов: 1) вход; 2, 4) сверточные слои; 3, 5) подвыборочные слои; 6, 7) слои из обычных нейронов

$$h_c = h_u - K + 1, \quad (2)$$

где w_c , h_c – ширина и высота свёрточной плоскости соответственно; w_u , h_u – ширина и высота плоскости предыдущего слоя; K – ширина (высота) окна сканирования.

В качестве активационной функции был выбран гиперболический тангенс:

$$f(a) = A \tanh(Sa). \quad (3)$$

где $f(a)$ – искомое значение элемента, a – взвешенная сумма сигналов предыдущего слоя, A – амплитуда этой функции, S – определяет ее положение относительно начала отсчета.

Данная функция активации является нечётной, с горизонтальными асимптотами $+A$ и $-A$.

Эта функция имеет ряд преимуществ для решения задачи:

- симметричные активационные функции, типа гиперболического тангенса, обеспечивают более быструю сходимость, чем стандартная логистическая функция;
- функция имеет непрерывную первую производную;
- функция имеет простую производную, которая может быть вычислена через ее значение, что дает экономию вычислений.

Формула функционирования нейрона сверточного слоя:

$$y_k^{(i,j)} = b_k + \sum_{s=1}^K \sum_{t=1}^K w_{k,s,t} x^{((i-1)+s, (j-1)+t)}, \quad (4)$$

где $y_k^{(i,j)}$ – нейрон k -й плоскости свёрточного слоя; b_k – нейронной смещение k -й плоскости; K – размер рецептивной области нейрона; $w_{k,s,t}$ – матрица синаптических коэффициентов; x – выходы нейронов предыдущего слоя.

Формула функционирования нейрона подвыборочного слоя:

$$y_k^{(i,j)} = b_k + \frac{1}{4} w_k \sum_{s=1}^2 \sum_{t=1}^2 x^{((i-1)+s, (j-1)+t)}. \quad (5)$$

Используется стандартный для нейронных сетей алгоритм обратного распространения ошибки. Для измерения качества распознавания использовалась функция среднеквадратической ошибки:

$$E_p = \frac{1}{2} \sum_j (t_{pj} - o_{pj})^2, \quad (6)$$

где E_p – величина функции ошибки для образа p ; t_{pj} – желаемый выход нейрона j для образа p ; o_{pj} – действительный выход нейрона j для образа p .

Окончательная коррекция синаптических коэффициентов происходит по формуле:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_{pj} o_{pj}, \quad (7)$$

где η – коэффициент пропорциональности, влияющий на скорость обучения.

На каждой итерации алгоритм обратного распространения ошибки рассчитывается для всего обучающего набора данных, чтобы вычислить средний или истинный градиент.

Когда ненастроенной сети предъявляется входной образ, она выдает некоторый случайный выход. Функция ошибки представляет собой разность между текущим выходом сети и идеальным выходом, который необходимо получить. Для успешного обучения сети требуется приблизить выход сети к желаемому выходу, т. е. последовательно уменьшать величину функции ошибки. Это достигается настройкой межнейронных связей. Каждый нейрон в сети имеет свои веса, которые настраиваются, чтобы уменьшить величину функции ошибки.

Значения весовых коэффициентов были выбраны случайным образом из нормального распределения с нулевым средним и стандартным отклонением:

$$\delta_w = \sqrt{m},$$

где m – число связей, входящих в нейрон.

Для обучения сети была создана база данных из 1000 изображений автомобильных номерных знаков.

На рис. 2 представлены некоторые изображения автомобильных знаков из обучающей выборки, расположенные фронтально относительно регистрирующего устройства. Изображения получены при различных погодных условиях, времени суток и различном освещении и контрасте.

Для создания большего количества примеров и увеличения степени инвариантности нейронной сети к различным поворотам был создан набор изображений автомобильных номерных знаков, расположенных под различными углами по вертикали, горизонтали и на плоскости относительно регистрирующего устройства (рис. 3). Изображения получены при различных погодных условиях, времени суток, различном освещении и контрасте.



Рис. 2. Изображения автомобильных номерных знаков из обучающей выборки, расположенных фронтально относительно регистрирующего устройства



Рис. 3. Изображения автомобильных номерных знаков из обучающей выборки, расположенных под углами относительно регистрирующего устройства

Выделение отдельных символов с помощью гистограмм средней интенсивности

После выделения области расположения символов на изображении необходимо выделить отдельные символы, для дальнейшего их распознавания. Для этого предлагается использовать метод, основанный на построении гистограмм средней интенсивности.

Выделенная на предыдущем этапе область расположения символов сканируется попиксельно слева направо, сверху вниз, при этом вычисляется средняя интенсивность пикселей в каждом столбце. В тех местах, где нет символа, средняя интенсивность будет значительно отличаться от интенсивности тех мест, где символы присутствуют. Далее, выполняя ту же операцию по строкам, получается набор отдельных символов, которые уже можно анализировать [16, 17].

Для того чтобы отделить строку символов от всего изображения, для начала предлагается вычислить горизонтальные гистограммы. Так как на изображении самой яркой областью является фон номерной пластины, то два наибольших максимума будут соответствовать областям 1 и 2 (рис. 4).



Рис. 4. Построение горизонтальных гистограмм. Линиям 1 и 2 соответствуют два наибольших максимума. x – номер строки изображения, y – средняя интенсивность строки изображения

Далее строятся вертикальные гистограммы под углом взаимно перпендикулярным к n , и выделяется уже около 10 максимумов в промежутках между символами. Таким образом, выделяются области расположения отдельных символов на номерной пластине (рис. 5).

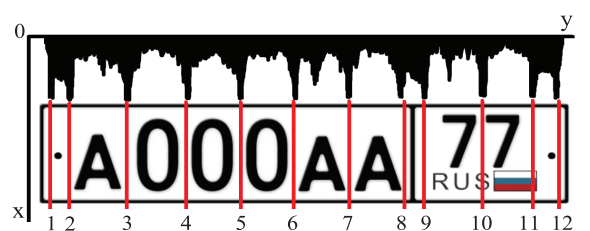


Рис. 5. Построение вертикальных гистограмм. Линиям от 1 до 12 соответствуют наибольшие максимумы. y – номер строки изображения, x – средняя интенсивность строки изображения

При регистрации изображение номерных пластин подвергается различным смещениям и искажениям, следовательно, линии, соответствующие областям 1 и 2, будут располагаться не по горизонтали, а под неизвестным углом. В связи с этим предлагается строить не одну, а n гистограмм средней интенсивности, каждая из которых строится не по горизонтали, а под заданным углом [18].

Необходимое количество гистограмм средней интенсивности определяется из технических условий регистрации изображений, по данным условиям угол разворота изображения не превышает 20° по горизонтали как в правую, так и в левую сторону, следовательно, $n=41$. Из n построенных гистограмм выбирается та, которая содержит наибольшее значение по y , т. к. наибольшее значение будет соответствовать области 1 или 2 (рис. 4) [19, 20].

Разработка сверточной нейронной сети для распознавания символов на изображениях

Для распознавания выделенных символов была разработана сверточная нейронная сеть с 4 скрытыми слоями (рис. 6).

Первый слой является входным и состоит из $28 \times 28 = 841$ нейрона.

Второй слой является свёрточным и состоит из шести плоскостей размером $24 \times 24 = 576$ нейронов.

Размер свёрточной плоскости определяется в соответствии с формулами (1, 2).

Третий слой является подвыборочным и также состоит из 5 плоскостей размером $12 \times 12 = 144$ нейронов.

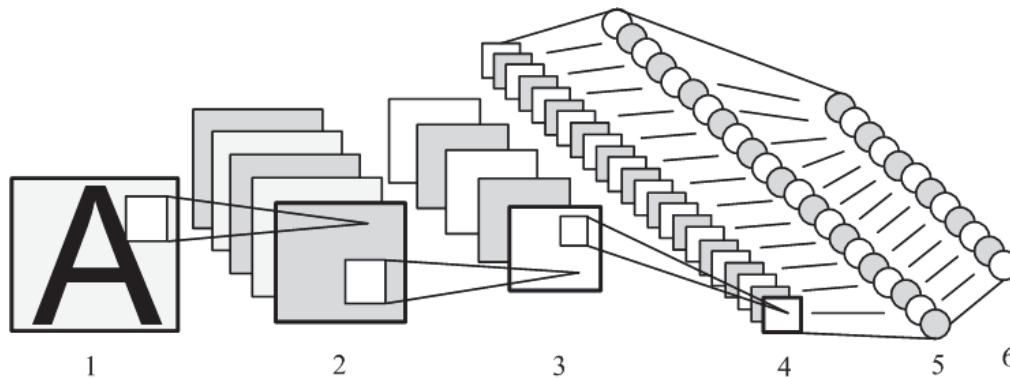


Рис. 6. Архитектура свёрточной нейронной сети для распознавания символов: 1) вход; 2, 4) свёрточные слои; 3) подвыборочный слой; 5, 6) слои из обычных нейронов

Четвёртый слой является свёрточным и состоит из 50 плоскостей размером $8 \times 8 = 64$ нейронов.

Пятый слой состоит из 126 простых сигмоидальных нейронов. Роль этого слоя – обеспечение классификации, после того, как выполнено извлечение особенностей и сокращение размерности входных данных.

Последний, шестой, слой является выходным слоем и состоит из 21 нейрона.

Согласно ГОСТ Р 50577–93 регистрационные знаки транспортных средств могут содержать следующие символы: А, В, Е, К, М, Н, О, Р, С, Т, Х, У и все цифры от 0 до 9. Поэтому выходной слой состоит из 21 нейрона, так как распознается 21 символ.

В качестве активационной функции был выбран гиперболический тангенс (3), формула функционирования нейрона свёрточного слоя (4), формула функционирования нейрона подвыборочного слоя (5). Для обучения сети используется алгоритм обратного распространения ошибки (6). Коррекция синаптических коэффициентов происходит по формуле (7).

Для обучения сети была использована база из 60000 изображений рукописных цифр (MNIST) и создана база из 20000 изображений букв. Размер тестирующей выборки – 10000 символов.

В табл. 1 отображены значения вероятности верного распознавания символов в зависимости от изменения углов регистрации.

На основе представленных алгоритмов разработана программная система, которая обеспечивает вероятность распознавания автомобильных номерных знаков на изображениях со сложным фоном не менее 98 % при следующих условиях регистрации:

- время обработки: 35 миллисекунд;
- высота символов не менее 12 пикселей;
- освещенность пластины номерного знака в диапазоне от 50 до 1000 лк;
- угол отклонения пластины номерного знака по горизонтали относительно регистрирующего устройства до $\pm 60^\circ$;
- угол отклонения пластины номерного знака по вертикали относительно регистрирующего устройства до $\pm 65^\circ$;

- угол поворота пластины номерного знака на плоскости до $\pm 20^\circ$.

Таблица 1. Вероятность верного распознавания символов в зависимости от изменения углов регистрации

Угол отклонения	Вероятность распознавания, %		
	по вертикали	по горизонтали	на плоскости
-75	87	85	–
-70	98	93	–
-65	100	99	–
-60	100	100	–
-55	100	100	–
-50	100	100	–
-45	100	100	–
-40	100	100	–
-35	100	100	–
-30	100	100	–
-25	100	100	–
-20	100	100	100
-15	100	100	100
-10	100	100	100
-5	100	100	100
0	100	100	100
5	100	100	100
10	100	100	100
15	100	100	100
20	100	100	100
25	100	100	–
30	100	100	–
35	100	100	–
40	100	100	–
45	100	100	–
50	100	100	–
55	100	100	–
60	100	100	–
65	100	98	–
70	99	92	–
75	88	86	–

Сравнение технических характеристик систем распознавания автомобильных номерных знаков

В табл. 2, 3 приведено сравнение технических характеристик разработанной программной системы с существующими на рынке системами распознавания автомобильных номеров.

Таблица 2. Сравнение технических характеристик систем распознавания автомобильных номеров

Название системы	Вероятность распознавания, %	Время распознавания	Освещенность, лк	Минимальная высота номера в кадре
«Авто-инспектор»	95	не указано	не менее 50	не указано
«Авто-интеллект»	90	не указано		не указано
«SL-Traffic»	90	не указано		25 пикселей
«Дигнум-авто»	90	не указано		не указано
«CarFlow II»	93...98	60 мс		не указано
Разработанное ПО	98	35 мс	от 50 до 1000	12 пикселей

Таблица 3. Возможности распознавания систем при различных углах отклонения пластины автомобильного номерного знака

Название системы	Угол отклонения, град.		
	Вертикальный	Горизонтальный	На плоскости
«Авто-инспектор»	±40	±30	±15
«Авто-интеллект»	±30	±20	±10
«SL-Traffic»	±60	±30	±10
«Дигнум-авто»	±40	±40	не указано
«CarFlow II»	±45	±45	±15
Разработанное ПО	±65	±60	±20

СПИСОК ЛИТЕРАТУРЫ

- Макаренко А.А., Калайда В.Т. Методика локализации изображения лица для систем видеоконтроля на основе нейронной сети // Известия Томского политехнического университета. – 2006. – Т. 309. – № 8. – С. 113–118.
- Елизаров А.И., Афонасенко А.В. Методика построения систем распознавания автомобильного номера // Известия Томского политехнического университета. – 2006. – Т. 309. – № 8. – С. 118–121.
- Le Cun Y., Bengio Y. Convolutional networks for images, speech and time series // The handbook of brain theory and neural networks. – 1998. – V. 7. – № 1. – P. 255–258.
- Болотова, Ю.А., Спицын В.Г. Сравнение способов обучения модели НТМ для задачи распознавания цифр // Молодежь и современные информационные технологии: сборник трудов IX Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых. – Томск: Изд-во СПБ Графикс, 2011. – Т. 1. – С. 252–253.
- Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.
- Болотова, Ю.А., Спицын В.Г., Кермани А.К. Распознавание символов на цветном фоне на основе иерархической временной модели с предобработкой фильтрами Габора // Электромагнитные волны и электронные системы. – 2012. – Т. 16. – № 1. – С. 14–19.
- Bundzel M., Hashimoto S. Object identification in dynamic images based on the memory-prediction theory of brain function // Journal of Intelligent Learning Systems and Applications. – 2010. – V. 2. – № 4. – P. 212–220.
- Буй Тхи Тху Чанг, Фан Нгок Хоанг, Спицын В.Г. Алгоритмическое и программное обеспечение для классификации цифровых изображений с помощью вейвлет-преобразования Хаара и нейронных сетей // Известия Томского политехнического университета. – 2011. – Т. 319. – № 5. – С. 103–106.

Как видно из табл. 2, разработанная программная система по всем параметрам не уступает существующим на рынке системам, а по некоторым параметрам превосходит.

Как видно из табл. 3, разработанная программная система способна распознавать автомобильные номерные знаки при самых больших углах отклонения в сравнении с другими системами.

Заключение

Для существенного повышения эффективности выделения и распознавания автомобильных номерных знаков на изображениях со сложным фоном предложено использовать сверточные нейронные сети, т. к. они являются устойчивыми к смещениям, поворотам, смене ракурса и прочим искажениям входных данных.

Разработана сверточная нейронная сеть, обеспечивающая нахождение и выделение области расположения символов на изображениях со сложным фоном. Для выделения отдельных символов предложено использовать алгоритм, основанный на построении гистограмм средней интенсивности пикселей. Для распознавания отдельных символов разработана сверточная нейронная сеть, работающая как классификатор символов. Проведена оптимизация структуры разработанных нейронных сетей, что позволило повысить эффективность их работы.

- Болотова Ю.А., Спицын В.Г., Фомин А.Э. Применение модели иерархической временной памяти в распознавании изображений // Известия Томского политехнического университета. – 2011. – Т. 318. – № 5. – С. 60–63.
- Кермани К.А., Спицын В.Г., Хамкер Ф. Нахождение параметров и удаление постоянной составляющей фильтра Габора для обработки изображений // Известия Томского политехнического университета. – 2011. – Т. 318. – № 5. – С. 57–59.
- Hansen D.W., Hansen J.P., Nielsen M. Eye typing using Markov and active appearance models // Applications of computer vision. – 2002. – V. 12. – P. 132–136.
- Rowley H.A., Baluja S., Kanade T. Neural network-based face detection // Pattern anal. mach. intell. – 1998. – V. 20. – P. 23–38.
- Le Cun Y., Huang F., Bottou L. Learning Methods for Generic Object Recognition with Invariance to Pose and Lighting // Proceedings of CVPR'04. – Washington, DC, USA: IEEE Computer Society, 2004. – P. 97–104.
- George D., Hawkins J. A hierarchical bayesian model of invariant pattern recognition in the visual cortex // Proceedings. 2005 IEEE International Joint Conference on Neural Networks. – Montreal, Canada: IEEE Computer Society, 2005. – V. 3. – P. 1812–1817.
- Chapelle O., Haffner P., Vapnik V. Support vector machines for histogram-based image classification // IEEE Trans. on Neural Networks & Learning Systems – 1999. – V. 10. – № 5. – P. 1055–1064.
- Feraud R., Bernier O., Viallet J., Collobert M. A fast and accurate face detector based on neural networks // IEEE Transactions on pattern analysis and machine intelligence. – 2001. – V. 23. – № 23. – P. 42–53.
- Karnowski T., Arel I., Rose D. Deep Spatiotemporal Feature Learning with Application to Image Classification // The 9th International conference on machine learning and applications (ICMLA'10). – 2010. – P. 883–888.

18. Кульчин Ю.Н., Ноткин Б.С., Седов В.А. Нейро-итерационный алгоритм томографической реконструкции распределенных физических полей в волоконно-оптических измерительных системах // Компьютерная оптика. – 2009. – Т. 33. – № 4. – С. 446–455.
19. Lai Z., Hongbin G., Ben N. Visual Hand Pose Estimation Based on Hierarchical Temporal Memory in Virtual Reality Cockpit Simulator // Information Technology Journal. – 2011. – V. 10. – № 9. – P. 1809–1816.
20. Le Cun Y., Kavukcuoglu K., Farabet C. Convolutional Networks and Applications in Vision // International Symposium on Circuits and Systems (ISCAS'10). – Paris, France: IEEE Computer Society, 2010. – P. 253–256.

Поступила 05.12.2013 г.

UDC 004.931

APPLICATION OF CONVOLUTIONAL NEURAL NETWORKS FOR EXTRACTION AND RECOGNITION OF CAR NUMBER PLATES ON IMAGES WITH COMPLEX BACKGROUND

Aleksey A. Druki,

Tomsk Polytechnic University, 30, Lenin avenue, Tomsk, 634050, Russia.

E-mail: druki2008@yandex.ru

The urgency of the discussed issue is caused by the need to develop methods, algorithms and programs to ensure efficiency of car number plate recognition on images with a complex background.

The main aim of the study: to increase the efficiency of character recognition on images with a complex background by developing methods, algorithms and programs invariant to affine and projective transformations of the input data.

The methods used in the study: To solve the task the author has used the methods of the artificial Intelligence, identification and pattern recognition in images, theory of artificial neural networks, convolutional neural networks, evolutionary algorithms, mathematical modeling, probability theory and mathematical statistics with the help of software Visual Studio and MatLab.

The results: The author developed the software allowing the recognition of car number plates on images with a complex background. The seven-layer convolutional neural network for character area selection on images is proposed. The algorithm based on the average pixel intensity histograms for individual characters selection is used. The six-layer convolutional neural network for character recognition on images is implemented. The represented software system can recognize license plates with deviation horizontally, vertically and in a plane angles. The system has high speed work.

Key words:

Image processing, artificial intelligence, character recognition, neural networks, histogram of average intensity.

REFERENCES

1. Makarenko A.A., Kalayda V.T. Metodika lokalizatsii izobrazhenia litsa dlya sistem videokontrolya na osnove neyronnoy seti [The locating face method for video surveillance systems based on neural network]. *Bulletin of Tomsk Polytechnic University*, 2006, vol. 309, no. 8, pp. 113–118.
2. Elizarov A.I., Afonashenko A.V. Metodika postroeniya sistem raspoznavaniya avtomobilnogo nomera [The constructing method of license plate recognition systems]. *Bulletin of Tomsk Polytechnic University*, 2006, vol. 309, no. 8, pp. 118–121.
3. Le Cun Y., Bengio Y. Convolutional networks for images, speech and time series. *The handbook of brain theory and neural networks*, 1998, vol. 7, no. 1, pp. 255–258.
4. Bolotova Yu.A., Spitzyn V.G. Sravnenie sposobov obucheniya modeli HTM dlya zadachi raspoznavaniya tsifr [Comparing ways of HTM model learning for digits recognition]. *Sbornik trudov IX Vserossiyskoy nauchno-prakticheskoy konferentsii studentov, aspirantov i molodykh uchenykh «Molodezh i sovremennye informatsionnye tekhnologii»* [Proceedings of the IX Russian scientific and practical conference of students, graduate students and young scientists. Youth and modern information technology]. Tomsk, Publ. SPB Grafiks, 2011, vol. 1, pp. 252–253.
5. Gonzalez R., Woods R. Digital image processing. Moscow, Technosphere, 2005. 1072 p.
6. Bolotova Yu.A., Spitzyn V.G., Kermani K.A. Raspoznavaniye simbolov na tsvetnom fone na osnove ierarkhicheskoy vremennoy modeli s predobrabotkoy filtrami Gabora [Character recognition on a colored background based on HTM model pretreatment with Gabor filters]. *Electromagnetic waves and electronic systems*, 2012, vol. 16, no. 1, pp. 14–19.
7. Bundzel M., Hashimoto S. Object identification in dynamic images based on the memory-prediction theory of brain function. *Journal of Intelligent Learning Systems and Applications*, 2010, vol. 2, no. 4, pp. 212–220.
8. Buy Thi Thu Chang, Fan Ngok Hoang, Spitzyn V.G. Algoritmicheskoe i programmnoe obespechenie dlya klassifikatsii tsifrovyykh izobrazheniy s pomoshchyu Veyvlet preobrazovaniya Haara i neyronnykh setey [Algorithms and software for digital images classification with the help the Haar wavelet transform and neural networks]. *Bulletin of Tomsk Polytechnic University*, 2011, vol. 319, no 5, pp. 103–106.
9. Bolotova Yu.A., Spitzyn V.G., Fomin A.E. Primeneniye modeli ierarkhicheskoy vremennoy pamyati dlya raspoznavaniya izobrazheniy [The hierarchical temporal memory model application for image recognition]. *Bulletin of Tomsk Polytechnic University*, 2011, vol. 318, no. 5, pp. 60–63.
10. Kermani K.A., Spitzyn V.G., Hamker F. Nakhozhdeniye parametrov i udalenie postoyannoy sostavlyayushchey filtra Gabora dlya obrabotki izobrazheniy [Finding parameters and removing the constant component of Gabor filter for image processing]. *Bulletin of Tomsk Polytechnic University*, 2011, vol. 318, no. 5, pp. 57–59.

11. Hansen D.W., Hansen J.P., Nielsen M. Eye typing using Markov and active appearance models. *Workshop on applications of computer vision*. Orlando, Florida, USA, IEEE Computer Society, 2002, vol. 12, pp. 132–136.
12. Rowley H.A., Baluja S., Kanade T. Neural network-based face detection. *Pattern anal. mach. intell.*, 2000, vol. 5, pp. 23–38.
13. LeCun Y., Huang F., Bottou L. Learning Methods for Generic Object Recognition with Invariance to Pose and Lighting. *Proceedings of CVPR'04*. Washington, DC, USA, IEEE Computer Society, 2004, pp. 97–104.
14. George D., Hawkins J. A hierarchical bayesian model of invariant pattern recognition in the visual cortex. *Proceedings. 2005 IEEE International Joint Conference on Neural Networks*. Montreal, Canada, IEEE Computer Society, 2005, vol. 3, pp. 1812–1817.
15. Chapelle O., Haffner P., Vapnik V. SVMs for histogram-based image classification. *IEEE Trans. Neural Networks*. Washington, DC, USA, IEEE Computer Society, 1999, vol. 10, no. 5, pp. 1055–1064.
16. Feraud R., Bernier O., Viallet J., Collobert M. A fast and accurate face detector based on neural networks. *Transactions on pattern analysis and machine intelligence*, 2002, vol. 3, no. 23, pp. 42–53.
17. Karnowski T., Arel I., Rose D. Deep Spatiotemporal Feature Learning with Application to Image Classification. *The 9th International conference on machine learning and applications (ICMLA'10)*. Washington, DC, USA, IEEE Computer Society, 2010, pp. 883–888.
18. Yu N., Notkin B.S., Sedov V.A. Neuro-iterative algorithm of tomographic reconstruction of the distributed physical fields in the fiberoptic measuring systems. *Computer optics*, 2009, vol. 33, no. 4, pp. 446–455.
19. Lai Z., Hongbin G., Ben N. Visual Hand Pose Estimation Based on Hierarchical Temporal Memory in Virtual Reality Cockpit Simulator. *Information Technology Journal*, 2011, vol. 10, no. 9, pp. 1809–1816.
20. LeCun Y., Kavukcuoglu K., Farabet C. Convolutional Networks and Applications in Vision. *International Symposium on Circuits and Systems (ISCAS'10)*. Paris, IEEE, 2010, pp. 253–256.

УДК 004.932

ПРИМЕНЕНИЕ МНОГОШАГОВОЙ СЕГМЕНТАЦИИ ДЛЯ РАСПОЗНАВАНИЯ НЕЧЕТКИХ ДУБЛИКАТОВ ИЗОБРАЖЕНИЙ

Немировский Виктор Борисович,

канд. физ.-мат. наук, доцент каф. информатики и проектирования систем
Института кибернетики Томского политехнического университета,
Россия, 634050, г. Томск, пр. Ленина, д. 30. E-mail: vbn@tpu.ru

Стоянов Александр Кириллович,

канд. техн. наук, доцент каф. информатики и проектирования систем
Института кибернетики Томского политехнического университета,
Россия, 634050, г. Томск, пр. Ленина, д. 30. E-mail: sto@tpu.ru

Актуальность работы обусловлена необходимостью распознавания нечётких дубликатов изображений в системах технического зрения, в работе с большими цифровыми архивами, а также при поиске изображений в сети Интернет.

Цель работы: исследование возможности применения многошаговой сегментации для распознавания нечётких дубликатов изображений.

Методы исследования: в выполненных исследованиях сегментация реализуется за счёт кластеризации яркостей пикселей изображений. Для кластеризации используется рекуррентная нейронная сеть, моделируемая одномерными точечными отображениями. Для оценки близости изображений применено косинусное расстояние между ранговыми распределениями мощностей кластеров яркости.

Результаты: Предложен поисковый образ изображения, основанный на ранговом распределении мощностей кластеров яркостей, выделенных на изображении. Приводятся экспериментальные результаты по распознаванию дубликатов изображений, основанному на применении предложенного образа. Показано, что использование многошаговой сегментации и рангового распределения мощности кластеров яркости позволяет надёжно определять нечёткие дубликаты оригинала изображения с большой степенью искажения на них, вплоть до радиуса гауссовых искажений, равного 8 пикселям. Применение такого подхода позволяет надёжно решать и обратную задачу обнаружения оригинала изображения даже по его пятикратно уменьшенной копии с радиусом гауссовых искажений на ней до 8 пикселей.

Ключевые слова:

Изображение, пиксель, точечные отображения, рекуррентная нейронная сеть, кластеризация, сегментация, распознавание изображений, ранговое распределение.

Введение

Задача распознавания изображений актуальна для систем технического зрения в робототехнике и иных технических системах; для поиска изобра-

жений по образцу в Интернете, цифровых архивах и библиотеках; при анализе содержания космических и аэрофотоснимков и т. д. [1, 2]. В основном это связано с отсутствием универсальных методов,

решающих задачу для всех видов изображений. Поэтому поиск новых алгоритмов, расширяющих возможности распознавания изображений, является одной из актуальных проблем, привлекающих внимание разработчиков.

Ответ на вопрос о схожести двух изображений можно было бы получить, поэлементно сравнивая их по ряду каких-либо параметров. Реализация такого прямого сравнения малоэффективна, что связано с отличиями от оригинала, которые приобретают распознаваемые изображения в процессе получения. Такие отличия приводят к появлению так называемых нечётких дубликатов. В качестве нечётких дубликатов могут пониматься изображения, отличающиеся разрешением или наличием шума, подвергшиеся небольшому фотометрическим преобразованиям; снимки одной и той же сцены, выполненные с небольшими изменениями ракурса камеры; видеокadres одной сцены. Другим случаем проявления нечётких дубликатов является изображение, подвергшееся искажениям. Кроме того, нечёткие дубликаты возникают в результате редактирования изображений и при создании коллажей. К понятию нечёткого дубликата приводит также поиск оригинала изображения по миниатюрной или увеличенной копии [3–11].

По причине проявления такого разнообразия отличий современные технологии распознавания изображения по образцу связаны с сопоставлением ему набора визуальных примитивов (характеристик яркости, цвета, формы, текстуры) и определением количественной оценки близости изображений по значениям примитивов. Визуальные примитивы – это характеристики изображения, которые вычисляются по исходным изображениям, позволяют эффективно индексировать их и использовать для распознавания и поиска. Поисковый образ изображения, сгенерированный из таких признаков, отражает визуальное его содержание, невелик по размеру в сравнении с самим изображением и удобен для организации поиска. Использование визуальных примитивов является на сегодняшний день достаточно эффективным и универсальным средством распознавания и поиска изображений в коллекциях оцифрованных изображений [12].

Очень часто при распознавании изображения в качестве поискового образа выбирают гистограммы – распределение пикселей по какому-то из выбранных визуальных примитивов [13]. Это могут быть гистограммы яркостей, цветов или выходов фильтров [1] и т. п. В [2], в частности, отмечено, что важную роль в анализе изображения играет форма объектов, присутствующих на нём. Гистограммы как раз и являются одной из простых характеристик формы.

Применение гистограмм привлекательно тем, что близость между ними определяется просто и её можно быстро вычислять, используя известные меры близости. Например, в качестве расстояния между гистограммами используется покомпонентная сумма модулей разности между ними. Несмо-

тря на предельную простоту подхода, он показывает довольно стабильные результаты [12]. Но всё же при сопоставлении изображений с помощью гистограмм бывает много ошибочных результатов.

Процесс получения цифрового изображения включает в себя дискретизацию и квантование яркости [14] исходного аналогового изображения. Гистограмма – это сжатое описание уже квантованного по яркости аналогового изображения, которое и используется в качестве поискового образа. Квантование можно рассматривать как простейшую кластеризацию, приводящую к получению «сегментированного» изображения, где под сегментацией понимается процесс выделения однородных областей на изображении. Особенность такой сегментации в том, что выполняется равномерное квантование всего диапазона яркостей. В этом случае сегментация никак не связана с формой объектов, присутствующих на изображении, а значит, не обеспечивает достаточно успешного распознавания. С другой стороны, применение специальных методов сегментации, основанных на кластеризации, способно дать больше информации об объектах на изображении. Действительно, естественно предположить, что если пиксели выделенных областей изображения близки друг к другу по своим параметрам (входят в один кластер), то они принадлежат одному объекту изображения.

Таким образом, можно предположить, что сегментированное изображение позволит решать задачу распознавания изображений более точно, не требуя при этом таких вычислительных затрат, как при использовании, например, корреляционных методов.

Целью настоящей работы является исследование возможности применения многошаговой сегментации на основе рекуррентной нейронной сети для распознавания нечётких дубликатов изображения.

Многошаговая сегментация изображения

Процедура многошаговой сегментации полутоновых изображений в серой шкале рассмотрена в [15, 16]. Она основана на кластеризации значений яркости пикселей изображения рекуррентной нейронной сетью, представленной на рис. 1.

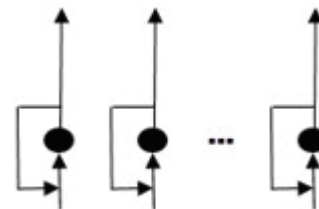


Рис. 1. Структура моделируемой нейронной сети

В [17] показано, что нейронная сеть с локальной обратной связью нейронов входного слоя может использоваться для кластеризации данных. Обратная связь моделируется с помощью одномерного отображения входных значений на активационной функции нейронов сети. Расчёт параме-

тров активационной функции нейронов, обеспечивающих кластеризацию (и, соответственно, сегментацию) подробно рассмотрен в [15]. Сама сегментация выполняется в несколько шагов, каждый из которых использует результаты предыдущего шага в качестве входных данных. При этом энтропия изображения, получаемого на каждом шаге, служит критерием завершения сегментации. Сегментация завершается, когда энтропия изображения, полученного на очередном шаге, перестаёт меняться. Описанная процедура была названа многошаговой сегментацией. С её помощью можно получить ряд сегментированных изображений с уменьшающейся детализацией сегментов и последовательно снижающейся энтропией.

Последнее изображение из такого ряда имеет энтропию, минимально возможную для данного распределения выделенных кластеров яркости. Следовательно, многошаговая сегментация обеспечивает сегментацию изображения на максимально упорядоченное по яркости распределение однородных областей пикселей. Это позволяет считать, что преобразованное изображение свободно от излишних подробностей и наиболее полно передаёт семантику сцены исходного изображения.

Распознавание сегментированных изображений

Чтобы разобраться в особенностях распознавания изображения, сегментированного описанной выше многошаговой процедурой, обратимся к диаграмме Ламерея (рис. 2). Она наглядно представляет процесс одномерного отображения значений яркости на активационной функции нейрона $f(x)$, лежащий в основе кластеризации. В качестве такой функции рассматривается сигмоида. В процессе отображения $x_{n+1} = f(x_n)$, (n – номер текущей итерации) любое значение x , попавшее в интервал между точками A и B , достигнет устойчивой точки x_3^* за одно и то же количество итераций. Именно это свойство и обеспечивает эффект кластеризации значений яркости и, как следствие, сегментацию изображения [15].

Покажем, что одномерное отображение, положенное в основу кластеризации (сегментации), позволяет отождествлять оригинальное изображение с его преобразованным дубликатом, несмотря на искажения яркости пикселей, а результаты сегментации могут использоваться для распознавания изображений.

Пусть в интервале AB располагается значение x , соответствующее яркости какого-то пикселя оригинального изображения. Допустим, в результате некоторых преобразований изображения яркость этого пикселя на полученном дубликате изменила своё значение на другое. При этом возможны два варианта.

Первый из них соответствует такому преобразованию, при котором вновь приобретённое значение хотя и изменится, но не выйдет за пределы интервала AB . Это означает, что в процессе отображения этого нового значения устойчивая точка x_3^* бу-

дет достигнута за то же количество итераций (т. е. попадет в тот же кластер), как и в случае оригинального изображения.

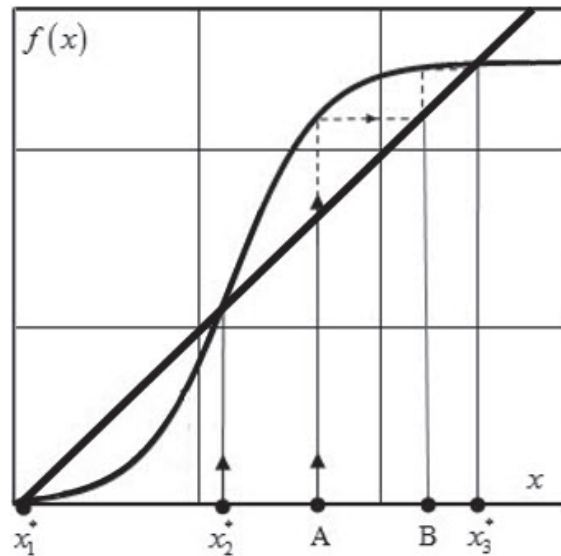


Рис. 2. Диаграмма Ламерея для одномерного отображения значений яркости x_1^* и x_3^* – устойчивые точки; x_2^* – неустойчивая точка отображения

Второй вариант соответствует преобразованию, которое выводит значение яркости пикселя оригинального изображения за пределы интервала AB . В этом случае устойчивая точка x_3^* для дубликата будет достигнута за иное количество итераций, то есть она попадет в другой кластер.

Таким образом, у сегментированного дубликата изображения для определенной части кластеров количество значений яркостей, попавшее в каждый из них, будет сохраняться, то есть распределение яркостей по кластерам останется неизменным по отношению к оригиналу. Будет также и иная ситуация – яркости пикселей в результате преобразований перейдут в соседние кластеры, меняя тем самым количество значений, входящих в него.

Очевидно, что любое сегментированное изображение можно характеризовать числом пикселей, попавших в выделенные в нём кластеры яркости. Назовём мощностью кластера число элементов, входящих в него. В этом случае величина мощности будет целым значением. Если же указывать долю элементов, попавших в кластер, от их общего числа, то такая величина будет дробной и меньше единицы. В обоих случаях распределение пикселей по мощности кластеров яркости может служить характеристикой любого изображения и использоваться для их распознавания. Чем больше различие в этих распределениях для двух изображений, тем более не похожи такие изображения.

Оценка близости изображений

Оценка близости распознаваемого изображения к оригиналу является важнейшей составляющей частью процедуры распознавания. При этой оценке учитывается тот факт, что различные формы пред-

ставления изображений в конечном счёте сводятся к представлению их в виде векторов в некотором n -мерном пространстве. Каждому изображению соответствует своя точка в этом пространстве.

В случае представления сегментированного изображения распределением мощностей кластеров для оценки близости этих распределений может быть принят такой же подход. Каждое распределение можно представить точкой и связанным с нею вектором в пространстве, размерность которого определяется числом выделенных в изображении кластеров яркости. Близость сравниваемых изображений будет измеряться расстоянием между точками, которые соответствуют полученным распределениям мощностей кластеров.

Процедура измерения расстояния d_{lp} между двумя точками l и p должна быть такой, чтобы в этом пространстве выполнялись следующие аксиомы метрического пространства:

- Расстояние между двумя точками l и p должно быть тождественным, то есть $d_{lp}=d_{pl}$.
- Для расстояний между тремя точками l , p и k должно соблюдаться правило треугольника, то есть $d_{lk}+d_{kp}>d_{lp}$.
- Расстояние между двумя точками l и p должно быть положительным, то есть $d_{lp}\geq 0$, причём $d_{lp}=0$, только если l совпадает с p .

Наиболее часто в метрическом пространстве используется евклидово расстояние

$$d_{lp} = \sqrt{\sum_i (x_{li} - x_{pi})^2},$$

где x_{li} , x_{pi} – i -е координаты точек l и p , соответственно.

Мера близости между двумя векторами (то есть изображениями) в n -мерном векторном пространстве может быть задана в виде угла. Если заданы вектор входного изображения $S=(s_1, s_2, \dots, s_n)$ и вектор изображения-оригинала $X=(x_1, x_2, \dots, x_n)$, то близость векторов определяется выражением

$$d_{sx} = \arccos \left(\frac{\sum_i s_i \cdot x_i}{|X| \cdot |S|} \right), \quad (1)$$

где $|S|$ и $|X|$ в знаменателе – соответственно, длины векторов S и X , а в числителе стоит их скалярное произведение. Такая мера называется косинусным расстоянием. Принадлежность входного изображения S к оригиналу X определяется с помощью решающего правила: $S \in X$, если $d_{sx} \leq \varepsilon$, где $\varepsilon \ll 1$ – наперёд заданная положительная величина.

Обе меры, евклидово расстояние и косинусное расстояние, основаны на вычислении скалярного произведения векторов. Это делает их использование во многих случаях равноправным. Хотя для случая выявления сходства двух векторов косинусное расстояние является более адекватной мерой по сравнению с евклидовым расстоянием.

При вычислении скалярного произведения для распределений мощностей кластеров необходимо учесть следующее обстоятельство. Пусть мощ-

ность кластера N является функцией яркости B , то есть $N=f(B)$. Для сопоставления двух распределений по скалярному произведению необходимо, чтобы сравниваемые значения мощности N брались для одного и того же значения яркости B . По отношению к яркостям кластеров это не всегда возможно. В многошаговой сегментации значения яркостей кластеров, выделенных на одном изображении, в общем случае не совпадают со значениями яркостей кластеров, выделенных на другом изображении. Это делает невозможным вычисление скалярного произведения для представления распределения мощностей кластеров, как функции яркости элементов кластера. Естественным выходом из указанного затруднения, по нашему мнению, является использование рангового распределения мощностей кластеров.

Ранговое распределение – это зависимость некоторой величины от ранга – порядкового номера в ряду её значений, упорядоченных по убыванию. Такое распределение широко используется при анализе текстов, технических и физических систем, и лежит в основе рангового анализа систем [18].

Для нашего исследования важно то, что можно сравнивать ранговые распределения с одинаковым числом рангов. Следовательно, возможно вычисление их скалярного произведения и мер, основанных на этом.

Экспериментальные результаты

С целью проверки возможности применения многошаговой сегментации для распознавания изображения и его нечётких дубликатов в серой шкале нами был проведён ряд экспериментов. При их проведении применялась программа сегментации изображений, реализующая описанный ранее алгоритм. В ней программная модель рекуррентной нейронной сети, представленная на рис. 1, использовалась в качестве модуля для кластеризации яркости. Параметры нейронной сети рассчитывались по методике, приведённой в [15, 16].

На каждом шаге обработки с помощью нейронной сети выделялись кластеры близких значений яркости пикселей. По завершению шага разные значения яркостей пикселей, оказавшихся в одном кластере, заменялись их средним по кластеру. Эти усреднённые значения использовались в качестве исходных данных для следующего шага сегментации.

Энтропия H , требуемая для оценки степени завершенности сегментации изображения, на каждом шаге оценивалась по формуле Шеннона

$$H = - \sum_i p_i \log_2 p_i.$$

Здесь $p_i = N_i/N$; N_i – число значений яркости, попавшей в i -й кластер; N – общее количество всех значений яркости. Суммирование велось по количеству кластеров, полученных при отображении входных значений яркостей пикселей на активационной функции нейрона.

С помощью описанной многошаговой процедуры получались сегментированные изображения оригинала изображения и его нечётких дубликатов. Создание дубликатов моделировалось гауссовым искажением яркостей пикселей эталонного изображения, выполненным с помощью растрового редактора Adobe Photoshop.

На рис. 3, 4 приведены некоторые из использованных для распознавания изображений. Соответствующие им ранговые распределения мощностей кластеров яркости приведены на рис. 5, а, б. На рисунке по ординате откладывались относительные значения мощности кластеров, вычисленные, как отношение числа пикселей, попавших в данный кластер, к общему числу пикселей изображения. По абсциссе откладывались номера рангов кластеров.

Графики на рис. 5, а позволяют выполнить качественное сравнение оригинала с изображениями других объектов, не являющихся его дубликатами. Нетрудно заметить, что распределения мощностей кластеров оригинала и изображений других объектов качественно отличаются друг от друга. Совпадение кривых наблюдается лишь в области больших значений рангов. Кластеры же с такими рангами содержат мало элементов и практически не отражают содержания изображения.

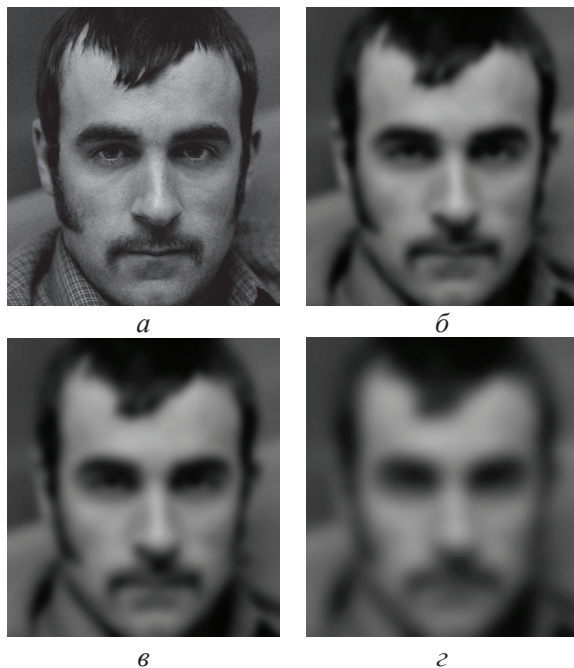


Рис. 3. Лицо мужчины: а) эталонное изображение, взятое из [19]; б) гауссово искажение изображения (радиус искажения 6 пикселей); в) гауссово искажение изображения (радиус искажения 8 пикселей); г) гауссово искажение изображения (радиус искажения 14 пикселей)

На рис. 5, б представлено ранговое распределение кластеров для оригинального изображения и его дубликатов. Рисунок демонстрирует очевидное качественное совпадение хода кривых.

Для целей распознавания важно иметь не

субъективную оценку близости сравниваемых изображений, а количественную её оценку. В эксперименте близость кривых рангового распределения мощностей кластеров яркости для дубликата и оригинала оценивалась по косинусному расстоянию $d_{\text{ск}}$, вычисленному по формуле (1).

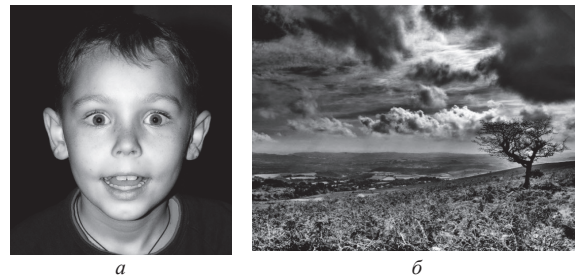
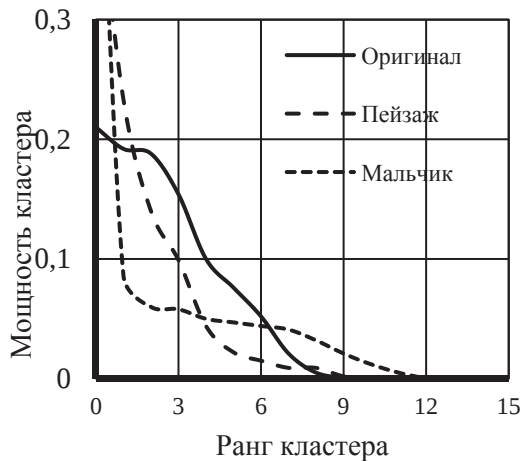


Рис. 4. Изображения объектов, полностью не совпадающих с эталонным: а) лицо мальчика [20]; б) пейзаж [21]

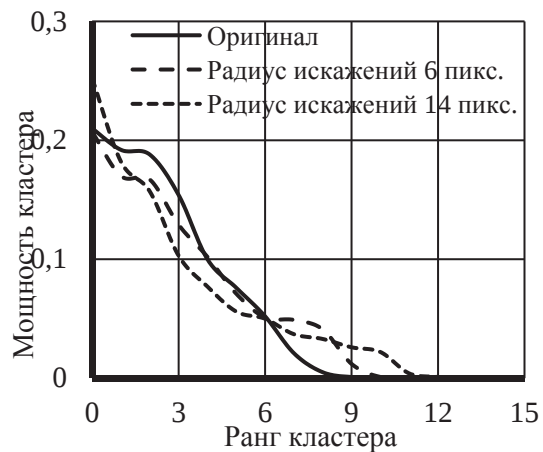
Чем более похожи сравниваемые изображения, тем меньше расстояние d между ними, определённое по (1). Полному совпадению изображений соответствует нулевое значение d . Пороговое расстояние, при котором изображение считается совпадающим с оригинальным, было выбрано равным $d=0,15$. Это соответствует примерно 10 % (в градусах) от угла в 90° , при котором n -мерные вектора сравниваемых изображений ортогональны, то есть изображения совершенно разные. Выбранный порог соответствует порогу, установленному на основании визуальной оценки совпадения изображений. Он обеспечивает количественную оценку близости, соответствующую субъективному человеческому восприятию схожих изображений. Действительно, согласно принятой нами оценке, $d=0,15$, при радиусе искажения, равном 8 пикселям, и сравниваемые изображения имеют достаточно высокую степень сходства. В этом несложно убедиться, обратившись к рис. 3. Изображения оригинала (рис. 3, а) и его дубликата (рис. 3, в) трудно считать относящимися к разным людям. Заметим, что предложенный порог различения довольно жёсткий. Человеческий глаз позволяет распознать дубликат и при более высоком уровне искажений (рис. 3, г).

Количественная оценка близости ранговых распределений, представленных на рис. 5, а, приведена в таблице. Расстояния вычислялись как для рангового распределения мощностей кластеров, так и для гистограмм. Определение этих расстояний выполнено из тех соображений, что, как отмечалось ранее, использование гистограмм для сравнения двух изображений до сих пор является популярным методом [22, 23]. Из таблицы видно, что вычисленные расстояния значительно превышают установленный порог близости $d=0,15$. При этом расстояние, вычисленное по ранговым распределениям, показывает большее различие изображений, чем при использовании гистограмм. Иными словами, качество распознавания изображения по мощностям кластеров выше.

Графики, отражающие количественную оцен-



а



б

Рис. 5. Ранговое распределение мощностей кластеров яркости: а) оригинал из [19]; б) оригинал и его искажённые дубликаты

ку сходства дубликатов с оригиналом, представлены на рис. 6. Здесь приведены расстояния между оригиналом и дубликатами, определённые по распределениям мощности кластеров, а также по гистограммам яркости.

Таблица. Косинусное расстояние между ранговыми распределениями мощностей кластеров яркости эталонного изображения и изображений других объектов

Изображение 1	Изображение 2	Вид искажения	Расстояние по мощности кластеров	Расстояние по гистограммам
Эталон – лицо мужчины	Лицо мальчика	Без искажения	0,764	0,560
	Пейзаж	Без искажения	0,482	0,459

Изучение результатов, представленных на рис. 6, показывает, что расстояние между оригиналом и его дубликатами, вычисленное по ранговому распределению мощности кластеров, в разы (примерно от двух до четырёх) меньше расстояния, определённого по гистограммам яркости. Порог распознавания $d=0,15$ для гистограмм достигается при значительно меньших искажениях, чем для распределения мощности кластеров.

Дубликаты оригинала могут быть искажёнными за счёт изменений яркости, а также иметь другой масштаб. Многошаговая сегментация позволяет распознавать изображение по его уменьшенному и искажённому дубликату. На рис. 7 приведен график зависимости косинусного расстояния от радиуса гауссового размытия при сравнении неискажённого оригинала изображения с его уменьшенными в 5 раз и искажёнными дубликатами. Дубликаты с разной степенью искажения были сделаны для оригинала изображения с рис. 3, а. Для наглядности на рис. 7 показан и график рас-

стояния между оригиналом и искажённым, но не уменьшенным дубликатом. Видно, что расхождение между кривыми значительно лишь при малых радиусах искажения (примерно до трёх пикселей). При больших радиусах кривые почти совпадают, что свидетельствует о высоком совпадении результатов сравнения в обоих случаях.

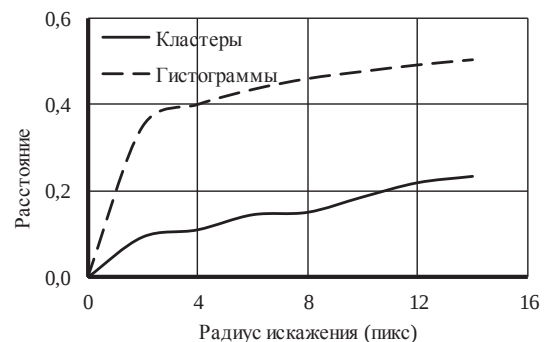


Рис. 6. Зависимость косинусного расстояния между эталоном и его дубликатами от радиуса гауссового искажения

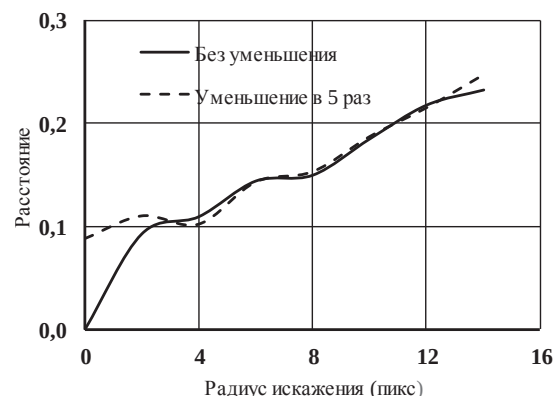


Рис. 7. Зависимость косинусного расстояния между оригиналом и его уменьшенными дубликатами от радиуса гауссового искажения

Выводы

1. Многошаговая сегментация, выполненная с помощью рекуррентной нейронной сети, позволяет создать образ изображения, основанный на ранговом распределении мощностей кластеров яркостей, выделенных на изображении.
2. Ранговое распределение мощности кластеров яркости на сегментированном изображении является характеристикой, достаточной для распознавания нечётких дубликатов оригинала изображения. Его использование дает более точные результаты распознавания по сравнению с методом, основанным на гистограммах.

3. При выбранном пороге различения $d=0,15$ использование многошаговой сегментации и рангового распределения мощности кластеров яркости позволяет надёжно определять нечёткие дубликаты оригинала изображения вплоть до радиуса гауссовых искажений на них, равного 8 пикселям.
4. Использование многошаговой сегментации и рангового распределения мощности кластеров позволяет надёжно решать задачу обнаружения оригинала изображения даже по его пятикратно уменьшенной копии с радиусом гауссовых искажений на ней до 8 пикселей.

Работа выполнена в рамках государственного задания «Наука».

СПИСОК ЛИТЕРАТУРЫ

1. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход. – М.: Изд-во «Вильямс», 2004. – 928 с.
2. Шапиро Л., Стокман Дж. Компьютерное зрение / под ред. С. Соколова, пер. с англ. А. Богуславского. – М.: Бином. Лаборатория знаний, 2006. – 752 с.
3. An Evaluation of Popular Copy-Move Forgery Detection Approaches / V. Christlein, J. Jordan, C. Riess, E. Angelopoulou // IEEE Transactions on information forensics and security. – 2012. – V. 7. – Iss. 6. – P. 1841–1854.
4. Farid H. Image Forgery Detection // IEEE Signal processing magazine. – 2009. – V. 26. – № 2. – P. 16–25.
5. Fridrich J., Soukal D. Detection of copy-move forgery in digital images // Proceedings of Digital Forensic Research Workshop. – Cleveland, 2003. – P. 55–61.
6. Sridevi M., Mala C., Sanyam S. Comparative Study of Image forgery and Copy-move Techniques // Proceedings of II International Conference on Computer Science, Engineering and Applications (ICCSEA 2012). – New Delhi, India, 2012. – P. 715–723.
7. Глузов Н.И., Кузнецов А.В., Мясников В.В. Поиск дубликатов на цифровых изображениях // Компьютерная Оптика. – 2013. – Т. 37. – № 3. – С. 360–368.
8. Кузнецов А.В., Мясников В.В. Алгоритм обнаружения дубликатов на цифровых изображениях с использованием эффективных линейных локальных признаков // Компьютерная Оптика. – 2013. – Т. 37. – № 4. – С. 489–496.
9. Пименов В.Ю. Метод поиска нечётких дубликатов изображений на основе выявления точечных особенностей // Труды РОМИП 2007–2008. – СПб.: НУ ЦСИ, 2008. – С. 145–158.
10. Мельниченко А., Гончаров А. Методы поиска изображений по визуальному подобию и детекции нечётких дубликатов изображений // Труды РОМИП 2009. – СПб.: НУ ЦСИ, 2009. – С. 108–121.
11. Добров Г.Б., Пятков Е.А. Алгоритм поиска нечётких дубликатов на основе простых признаков // Труды РОМИП 2009. – СПб.: НУ ЦСИ, 2009. – С. 78–82.
12. Некоторые подходы к организации содержательного поиска изображений и видеоинформации / Н.С. Байгарова, Ю.А. Бухштаб, Н.Н. Евтеева, Д.А. Корягин // Препринт ИПМ им. М.В. Келдыша РАН. – 2002. – № 78. URL: http://www.keldysh.ru/papers/2002/prep78/prep2002_78.html (дата обращения: 12.03.2014).
13. Котов В.В. Использование гистограммных оценок в задачах распознавания // Успехи современного естествознания. – 2004. – № 4. – С. 40–42.
14. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.
15. Немировский В.Б., Стоянов А.К. Сегментация изображений с помощью рекуррентной нейронной сети // Известия Томского политехнического университета. – 2012. – Т. 321. – № 5. – С. 205–210.
16. Nemirovsky V.B., Stoyanov A.K. Multi-Step Segmentation of Images by Means of a Recurrent Neural Network // Proceedings of VII International Forum on Strategic Technology (IFOST-2012). – Tomsk, September 18–21, 2012. – Tomsk: TPU Press, 2012. – V. 1. – P. 557–560.
17. Стоянов А.К. Применение рекуррентной нейронной сети для решения задачи кластеризации // Известия Томского политехнического университета. – 2009. – Т. 315. – № 5. – С. 144–149.
18. Кудрин Б.И. Математика ценозов: видовое, ранговидовое, ранговое по параметру гиперболические Н-распределения и законы Лотки, Ципфа, Парето, Мандельброта // Техногенная самоорганизация. Материалы к конференциям 2004 г. Ценологические исследования, вып. 25. – М.: Центр системных исследований, 2004. – 248 с.
19. Фотоблог о моде и стиле 2012–2014. URL: <http://fashion2012.ru/muzhskie-borodyi-foto.html> (дата обращения: 12.03.2014).
20. Детские портреты в ч/б. URL: <http://www.liveinternet.ru/users/katiava/post285466584> (дата обращения: 26.02.2014).
21. Черно-белые обои для рабочего стола. URL: <http://oboi-pny.ru/chno-belye-obo-i-dlya-rabochego-stola> (дата обращения: 26.02.2014).
22. Ioannidis Y. The History of Histograms (abridged) // Proceedings of XXIX International Conference on Very Large Data Bases. – Berlin, Germany, September 9–12, 2003. – P. 19–30.
23. Dalal N., Triggs W. Histograms of Oriented Gradients for Human Detection // IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR05. – San Diego, 2005. – V. 1 (3). – P. 886–893.

Поступила 18.03.2014 г.

APPLICATION OF MULTI-STEP IMAGE SEGMENTATION FOR NEAR-DUPLICATE IMAGE RECOGNITION

Victor B. Nemirovskiy,

Cand. Sc., Tomsk Polytechnic University, 30, Lenin Avenue, Tomsk, 634050,
Russia. E-mail: vbn@tpu.ru

Alexander K. Stoyanov,

Cand. Sc., Tomsk Polytechnic University, 30, Lenin Avenue, Tomsk, 634050,
Russia. E-mail: sto@tpu.ru

The urgency of the paper is caused by the need to detect image near-duplicate in computer vision systems, as well as when image searching on Internet or in large digital archives.

The main aim of the study: usage of multi-step segmentation for near-duplicate image recognition.

The methods used in the study: clustering of image pixels brightness is used for segmentation. The recurrent neural network is used for clustering. To estimate images similarity the authors have applied the cosine distance between rank distributions of clusters cardinality.

The results: The authors suggested the search patterns based on the rank distributions of brightness clusters cardinality. The paper introduces the experimental results on the near-duplicate image recognition based on application of the suggested search patterns. It is shown that the use of a multi-step segmentation and rank distribution of the brightness clusters cardinality allows determining reliably the near-duplicate of the original image with a high degree of distortion on them, up to the radius of the Gaussian distortion equal 8 pixels. Such an approach also allows solving reliably the inverse problem of detecting the original image even in its fivefold reduced copy with radius Gaussian distortion on it to 8 pixels.

Key words:

Image, pixel, point mapping, recurrent neural network, clustering, segmentation, recognition of images, ranking distribution.

REFERENCES

1. Forsyth D.A., Ponce J. *Computer Vision, a Modern Approach*. Moscow, Prentice Hall, 2003. 693 p.
2. Shapiro L.G., Stockman G.C. *Computer vision*. Moscow, Prentice Hall, 2001. 580 p.
3. Christlein V., Jordan J., Riess C., Angelopoulou E. An Evaluation of Popular Copy-Move Forgery Detection Approaches. *IEEE Transactions on information forensics and security*, 2012, vol. 7, Iss. 6, pp. 1841–1854.
4. Farid H. Image Forgery Detection. *IEEE Signal processing magazine*, 2009, vol. 26, no. 2, pp. 16–25.
5. Fridrich J., Soukal D. Detection of copy-move forgery in digital images. *Proceedings of Digital Forensic Research Workshop*. Cleveland, 2003. pp. 55–61.
6. Sridevi M., Mala C., Sanyam S. Comparative Study of Image forgery and Copy-move Techniques. *Proceedings of II International Conference on Computer Science, Engineering and Applications (ICCSEA 2012)*. New Delhi, India, 2012. pp. 715–723.
7. Glumov N.I., Kuznetsov A.V., Myasnikov V.V. Poisk dublikatov na tsifrovyykh izobrazheniyakh [The Algorithm for Copy-move Detection on Digital Images]. *Kompyuternaya optika – Computer optics*, 2013, vol. 37, no. 3, pp. 360–368.
8. Kuznetsov A.V., Myasnikov V.V. Algoritm obnaruzheniya dublikatov na tsifrovyykh izobrazheniyakh s ispolzovaniem effektivnykh lineynykh lokalnykh priznakov [Efficient Linear Local Features Based Copy-Move Detection Algorithm]. *Kompyuternaya optika – Computer optics*, 2013, vol. 37, no. 4, pp. 489–496.
9. Pimenov V.Yu. Metod poiska nechetkikh dublikatov izobrazheniy na osnove vyyavleniya tochechnykh osobennostey [Near-Duplicate Image Detection with Local Interest Point Extraction]. *Trudy ROMIP 2007–2008* [Proc. ROMIP 2007–2008]. Saint Petersburg, NU TCSI, 2008. pp. 145–158.
10. Melnichenko A., Goncharov A. Metody poiska izobrazheniy po vizualnomu podobiyu i detektsii nechetkikh dublikatov izobrazheniy [Image retrieval methods by the visual similarity and the detection of near-duplicate image]. *Trudy ROMIP 2009* [Proc. ROMIP 2009]. Saint Petersburg, NU TCSI, 2009. pp. 108–121.
11. Dobrov G.B., Pyatkov E.A. Algoritm poiska nechetkikh dublikatov na osnove prostykh priznakov [Near duplicates detection algorithm based on simple features]. *Trudy ROMIP 2009* [Proc. ROMIP 2009]. Saint Petersburg, NU TCSI, 2009. pp. 78–82.
12. Baygarova N.S., Bukhshtab Yu.A., Evteeva N.N., Koryagin D.A. Nekotorye podkhody k organizatsii soderzhatelnogo poiska izobrazheniy i videoinformatsii. Institut prikladnoy matematiki im. M. V. Keldysha RAN [Various Questions Connected with Content-Based Search of Visual Information and Videoinformation Preprint, Inst. Appl. Math., the Russian Academy of Science, no. 78]. Available at: http://www.keldysh.ru/papers/2002/prep78/prep2002_78.html (accessed 12 March 2014).
13. Kotov V.V. Ispolzovanie gistogramnykh otsenok v zadachakh raspoznavaniya [Use of histogram estimates in recognition tasks]. *Uspekhi sovremennogo estestvoznaniya*, 2004, no. 4, pp. 40–42. Available at: http://www.rae.ru/use/?section=content&op=show_article&article_id=7780895 (accessed 12 March 2014).
14. Gonzalez R., Woods R. *Digital Image Processing*. Moscow, Prentice Hall, 2002. 1072 p.
15. Nemirovsky V.B., Stoyanov A.K. Segmentatsiya izobrazheniy s pomoshchyu rekurrentnoy neyronnoy seti [Image Segmentation by Recurrent Neural Network]. *Bulletin of the Tomsk Polytechnic University*, 2012, vol. 321, no. 5, pp. 205–210.
16. Nemirovsky V.B., Stoyanov A.K. Multi-Step Segmentation of Images by Means of a Recurrent Neural Network. *Proc. VII International Forum on Strategic Technology (IFOST-2012)*. Tomsk, September 18–21, 2012. Tomsk, TPU Press, 2012. Vol. 1, pp. 557–560.
17. Stoyanov A.K. Primenenie rekurrentnoy neyronnoy seti dlya resheniya zadachi klasterizatsii [Application of recurrent neural network for clustering problem solution]. *Bulletin of the Tomsk Polytechnic University*, 2009, vol. 315, no. 5, pp. 144–149.

18. Kudrin B.I. Matematika tsenozov: vidovoe, rangovidovoe, rangovoe po parametru giperbolicheskie N-raspredeleniia i zakony Lotki, Tsipfa, Pareto, Mandelbrota. Tekhnogennaia samoorganizatsiya [Mathematics of cenoses: species, the rank of species, ranking in the parameter hyperbolic H-distribution and the laws of Lotka, Zipf, Pareto, Mandelbrot]. *Tsenologicheskie issledovaniya*. Moscow, Tsentr sistemnykh issledovaniy, Iss. 25, 2004. 248 p.
19. *Fotoblog o mode i stile 2012–2014* [Photoblog about fashion and style 2012–2014]. Available at: <http://fashion2012.ru/muzhskie-borodyi-foto.html> (12 March 2014).
20. *Detskie portrety v ch/b* [Children's Portraits in Black and White]. Available at: <http://www.liveinternet.ru/users/katiava/post285466584> (accessed 26 February 2014).
21. *Cherno-belye oboi dlya rabocheho stola* [Black and white wallpaper]. Available at: <http://oboiny.ru/cherno-belye-oboi-dlya-rabocheho-stola> (accessed 26 February 2014).
22. Ioannidis Y. The History of Histograms (abridged). *Proc. XXIX International Conference on Very Large Data Bases*. Berlin, Germany, September 9–12, 2003. pp. 19–30.
23. Dalal N., Triggs W. Histograms of Oriented Gradients for Human Detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR05*. San Diego, 2005. Vol. 1 (3), pp. 886–893.

Алгоритмическое и программное обеспечение

УДК 517.4:519.652

К ВОПРОСУ ВОССТАНОВЛЕНИЯ УЧЕТНЫХ ДАННЫХ НА ХИМИЧЕСКИХ ПРЕДПРИЯТИЯХ

Волошко Анатолий Васильевич,

канд. техн. наук, доцент кафедры электроснабжения Института энергосбережения и энергоменеджмента Национального технического университета Украины «Киевский политехнический институт», Украина, 03056, г. Киев, ул. Борщаговская, 115. E-mail: a-voloshko@yandex.ru

Бедерак Ярослав Семенович,

инженер ПАО «АЗОТ», Украина, 18016, г. Черкассы, ул. Первомайская, 72.
E-mail: yarbak@yandex.ua

Лутчин Татьяна Николаевна,

аспирант Института энергосбережения и энергоменеджмента Национального технического университета Украины «Киевский политехнический институт», Украина, 03056, г. Киев, ул. Борщаговская, 115.
E-mail: t.lutchyn@gmail.com

Кудрицкий Максим Юриевич,

магистрант Института энергосбережения и энергоменеджмента Национального технического университета Украины «Киевский политехнический институт», Украина, 03056, г. Киев, ул. Борщаговская, 115.
E-mail: bugs_bunni@ukr.net

Актуальность работы обусловлена наличием пропущенных данных в показаниях приборов учета энергии.

Цель работы: обоснование выбора метода восстановления пропущенных данных об энергопотреблении на промышленных предприятиях.

Методы исследования: модели рассчитываются с помощью приложения Curve Fitting Toolbox программного комплекса «Matlab 7.0». В состав библиотеки графических моделей Curve Fitting Toolbox входит приложение cftool, которое позволяет определить параметрическую модель (например, функции экспоненциальную Exr, полиномиальную Polynomial, рациональную RAT, а также сумму синусоидальных функций SumSin), выполнить подбор параметров, анализ пригодности приближения, отобразить результат графически. В библиотеке графических моделей Curve Fitting Toolbox определяются методом перебора модели из более чем 50 различных математических функций.

Результаты: Рассмотрены особенности простых и сложных методов восстановления данных с дальнейшим оцениванием их ошибок (погрешностей). Указаны способы повышения точности n -факторных моделей. Исследованы прямые и обратные зависимости восстановления утерянных учетных данных на примере химического предприятия. Обоснованы оптимальные диапазоны исследования исходных выборок данных. Также предусмотрены варианты определения наиболее рациональных методов восстановления значений в единичных случаях их отсутствия.

Ключевые слова:

Восстановление данных, n -факторные модели, ошибка модели, энергопотребление, утерянные данные.

Введение

Отсутствие данных технического учета об энергопотреблении на промышленных предприятиях

приводит к недоучету энергоресурсов, отсутствию возможности контроля энергоэффективности производственных объектов. Для решения данных во-

просов принято использовать восстановление данных. Восстановление данных необходимо начинать с проверки их выборок на наличие случайных значений.

Знание механизма, приводящего к отсутствию значений, является ключевым при выборе методов анализа и интерпретации результатов [1].

Причинами потери информации об энергопотреблении являются, как правило, следующие [2]: аппаратные и системные отказы, человеческий фактор, программные ошибки, вирусы, кражи и хищения, стихийные бедствия (пожары, наводнения, землетрясения, удар молнии).

Утерянные данные по виду их пропусков принято подразделять на виды [3]:

- 1) полностью случайные пропуски (data are missing completely at random – MCAR), если условная вероятность не зависит ни от самого пропущенного значения переменной, ни от значений прочих переменных (эта вероятность постоянна для всех наблюдений);
- 2) случайные пропуски данных (missing at random – MAR), если их вероятность не зависит от самого пропущенного значения переменной, но может зависеть от значений других переменных (в этих случаях механизм пропусков несущественен и к данным применимо большинство методов восстановления пропусков);
- 3) существенные пропуски данных, если их вероятность зависит от самого пропущенного значения переменной (механизм пропусков является существенным, и для корректного анализа данных необходимо знать этот механизм).

Краткое описание алгоритмов восстановления данных

Можно выделить следующие группы методов заполнения пропусков: простые и сложные [4]. К *простым (неитеративным) алгоритмам* на основе простых арифметических операций относятся: заполнение пропусков средним арифметическим, метод ближайшего соседа, подбор в группе и регрессионное моделирование пропусков.

Самым простым методом является заполнение средним арифметическим значением по учетным данным. Он не требует применения специального программного обеспечения. Средние значения, вычисленные на исходном и преобразованном массивах, совпадают. Однако такого рода преобразование «усредняет» данные, уменьшая дисперсию признака, и, следовательно, показатели корреляции, что приводит к занижению оценки.

Метод подбора в группе предполагает, что пропуски будут заполнены значениями, полученными в результате оценивания распределения данных по группам. Недостатком данного метода является то, что он требует значительных вычислительных затрат.

При использовании метода парной или множественной регрессии строится модель линейной зави-

симости переменной, в которой необходимо заполнить пропуски, от ряда других имеющихся признаков. Регрессионные коэффициенты для каждого из предикторов находятся методом наименьших квадратов в массиве с полными данными. Подставляя значения предикторов в регрессионное уравнение, получают прогноз пропущенного показателя.

Хорошее качество восстановления данных обеспечивает метод сплайн-интерполяции, особенно для одиночных пропусков и небольших выборок. В случае восстановления группы последовательных пропусков результат аппроксимации сплайном данной группы не всегда дает оценки, приближающиеся с достаточной точностью к значениям, которые могли бы быть на месте пропусков [3]. На практике чаще всего используют кубические сплайны и сплайны, не изменяющие форму кривой (сплайны shape-preserving).

Метод экспоненциального сглаживания также применим для восстановления одиночных данных на выборках небольшого объема (например, временной ряд почасовых значений за сутки).

Сложные (итеративные) алгоритмы предполагают оптимизацию некоторого функционала, отражающего точность расчета подставляемых на место пропуска значений. Их делят на глобальные и локальные.

Особенностью локальных алгоритмов является оценивание (предсказание) каждого пропущенного значения с использованием полного наблюдения, которые находятся в некоторой окрестности предсказываемого объекта.

Глобальные алгоритмы для оценивания каждого пропущенного значения оперируют всеми объектами рассматриваемой выборки. К ним относятся [4]:

- Метод Бартлетта, который представляет собой алгоритм, включающий три итерации. На первой итерации пропуски заполняются некоторым начальным значением (например, средним арифметическим по имеющимся данным). На второй итерации для преобразованной переменной строится регрессионная модель. На заключительном этапе на основе полученного регрессионного уравнения предсказываются новые значения для пропусков.
- Алгоритм Resampling (метод попарного сравнения): выборки данных, содержащие пропущенные данные, заменяют случайно подобранными строками из матрицы полных наблюдений. Затем строится регрессионное уравнение для предсказания отсутствующего значения. Процедура построения регрессионного моделирования повторяется несколько раз. После определенного количества повторений значения полученных регрессионных коэффициентов усредняют и получают окончательное решение с максимальной точностью прогноза пропущенного значения [5].

Особенности n-факторных моделей

Рассмотрим особенности, которые необходимо учитывать при построении моделей восстановления. Чем больше объем исследуемой выборки, тем лучше будут учтены в математической модели особенности ведения технологического процесса. С другой стороны, чем меньше объем выборки, тем меньше влияние сезонных составляющих [6].

Пропуски как зависимых, так и независимых переменных ставят задачу поиска определенного вида математической модели, которую можно использовать для восстановления данных. Для однофакторной модели $y=f(x)$ это может достигаться путем построения моделей вида $y=f(x)$ или $x=f(y)$ с помощью парной регрессии.

Таблица 1. Однофакторные и многофакторные математические модели для восстановления данных

Тип математической модели	Метод построения модели	Способы повышения точности построения модели
Однофакторная	Регрессионный метод с использованием пакета CurveFitting-Toolbox и SplineToolbox программы Matlab [7]	1. Метод перебора всех видов математических моделей (экспоненциальных, степенных, полиномиальных, отношения полиномов, суммы синусов и т. п., в сумме их более 50) [7–11]. 2. Критерии качества моделей: средняя абсолютная процентная ошибка $MAPE$, F-критерий Фишера, критерий Акаике AIC [12]. 3. Способ определения выборки данных наименьшего объема, что обеспечивает минимальную ошибку модели [12]
Многофакторная	Метод группового учета аргументов [13]	1. Включение в проверочную последовательность характерных точек (например, точки с нулевым энергопотреблением энергоресурсов с координатами (0, 0, 0) при объеме выборки в 150 единиц) [14]. 2. Учет как текущих, так и предыдущих значений независимых переменных [15]. 3. Критерии качества моделей: регулярности $\Delta^2(B)$, минимума смещения, точности кратковременного прогноза $\Delta^2(C)$ и коэффициента простоты $K_{пр}$ [13]

Для двухфакторной зависимости при парном коэффициенте корреляции между зависимой и независимыми переменными более 0,75 для выборок данных за длительный период (300 значений и более) целесообразно строить три однофакторные модели при помощи парной регрессии вместо множественной регрессии [7]. Это проще и эффективнее, чем построение множественных линейных регрессий, которые требуют специального программного обеспечения, подобно методу группового учета аргументов. Результаты исследований в области по-

вышения точности построения однофакторных и многофакторных математических моделей для восстановления данных энергопотребления сведены в табл. 1.

Оценивание ошибок простых и сложных методов восстановления данных

Для построения и сравнения многофакторных моделей рассмотрим данные энергопотребления, расхода ресурсов и выработки продукции на химическом предприятии. Так, если электропотребление E цеха по производству аммиака зависит от объема выпуска аммиака A и от потребления природного газа G , то при наличии пропусков данных и в зависимых, и в независимых переменных и при тесной связи между этими переменными необходимо для восстановления данных строить 3 модели: $E=f(A)$, $A=f(G)$, $G=f(E)$.

Для наглядности фрагмент исходных данных приведен в табл. 2 в виде среднечасовых значений за 01.05.2012.

Таблица 2. Исходные данные об энергопотреблении

Время	Выработка аммиака A , т	Расход электроэнергии E , МВт·ч	Расход природного газа G , тыс. м ³
0:00:00	39,699	31,988	45,211
1:00:00	39,292	32,005	45,182
2:00:00	39,644	31,932	45,357
3:00:00	39,929	31,923	45,122
4:00:00	39,684	32,105	45,481
5:00:00	*	32,056	45,782
6:00:00	39,422	32,063	*
7:00:00	43,174	32,098	45,602
8:00:00	42,055	31,971	45,608
9:00:00	40,449	31,953	45,023
10:00:00	41,385	31,860	44,973
11:00:00	38,386	31,759	45,042
12:00:00	39,183	*	45,100
13:00:00	40,331	31,640	44,743
14:00:00	*	31,806	44,836
...	...	...	...

* – пропуски данных.

Для данных химического производства математические модели определялись тремя простыми методами (замены пропуска средним арифметическим значением, подбора в группе и регрессионным методом) и двумя сложными методами (Барлетта и Resampling), а также методами сплайн-интерполяции кубическим сплайном и одним из методов экстраполяции – методом экспоненциального сглаживания. При восстановлении данных методом Resampling модель строилась без повторений.

Методом парной регрессии рассчитывались ошибки прямых и обратных моделей. Для определения параметрической модели целесообразно использовать отрезки рядов Фурье Fourier, сумму синусоидальных функций SumSin, экспоненциальные Exр, степенные Power, полиномиальные Polynomial, рациональные RAT и другие функции. Далее выполнялся подбор параметров, проводился

анализ пригодности приближения с графическим отображением результата [16, 17]. Затем выбирались лучшие параметрические модели каждого вида функции.

Результаты выбора метода, обеспечивающего наилучшее качество восстановления данных (наименьшую среднюю абсолютную процентную ошибку (MAPE)), определяли согласно [18]. Результаты расчетов простых и сложных методов для трехфакторной модели с разными интервалами определения сведены в табл. 3.

Таблица 3. Результаты расчетов простых и сложных методов для трехфакторной модели с разными интервалами определения

Количество временных интервалов в выборке	Вид зависимости	Ошибка простых методов, %			Ошибка метода экспоненциального сглаживания, %		Ошибка метода сплайн-интерполяции, %		Ошибка сложных методов, %	
		Метод среднего арифметического	Метод подбора в группе	Регрессионный метод	Ошибка метода экспоненциального сглаживания, %	Ошибка метода сплайн-интерполяции, %	Metod Resampling	Metod Bartlett		
24	$A=f(G)$	0,08	2,42	1,25	2,2	2,42	2,60	1,02		
	$E=f(A)$	0,01	0,75	0,26	0,08	0,75	0,10	0,09		
	$G=f(E)$	0,01	0,54	0,37	0,11	0,54	0,61	0,24		
168	$A=f(G)$	0,91	0,01	2,1	–	–	3,43	9,45		
	$E=f(A)$	0,36	0,03	0,53	–	–	0,08	3,62		
	$G=f(E)$	0,64	0,03	1,03	–	–	0,06	3,78		
700	$A=f(G)$	1,04	0,10	2,07	–	–	2,16	6,53		
	$E=f(A)$	0,52	0,01	1,41	–	–	0,91	4,01		
	$G=f(E)$	0,62	0,01	1,53	–	–	0,74	2,44		
1050	$A=f(G)$	0,94	0,04	1,97	–	–	1,38	2,28		
	$E=f(A)$	1,93	0,02	1,26	–	–	0,63	3,91		
	$G=f(E)$	0,47	0,02	1,61	–	–	2,56	0,96		

После выбора моделей необходимо проверить их результаты на адекватность. Для этого необходимо выбрать степень значимости (например, 0,05) и рассчитать значение F -критерия Фишера, а также F -критическое значение $F_{кр}$. Если $F > F_{кр}$ при данной степени значимости, то модель адекватна согласно работе [18]. Другие критерии, которыми можно оценивать результаты методов восстановления, приведены в работах [19–21].

На основании полученных результатов можно сделать вывод, что лучшее качество восстановления данных энергопотребления в цехе химического производства обеспечивает простой метод подбора в группе. Оптимальным объемом выборки данных, обеспечивающим минимальную ошибку, является минимальная по объему суточная выборка данных (24 значения независимой и 24 значения зависимой переменных). Поэтому исследуемый процесс электропотребления на данном химическом производстве можно отнести к необратимым процессам [22], то есть увеличение числа наблюде-

ний только ухудшит прогнозные и аналитические свойства модели [23].

Как следует из условия стационарности, для наиболее полного анализа стационарных процессов следует собрать как можно больше статистических данных о них. В этом случае удастся тем более точно определить и спрогнозировать характеристики процесса, чем более полной будет выборка наблюдений за ними [22]. Для нестационарных процессов такое правило неприменимо.

Под необратимыми понимаются неоднородные во времени процессы, характеристики которых необратимо меняются с течением времени t и являются вариантными относительно временных сдвигов:

$$t \rightarrow t + T, Y(t) \rightarrow Y(t + T) + \Delta Y(T)$$

при любом фиксированном T (действительном или целочисленном), где приращение $\Delta Y(T)$ однозначно не вытекает из характеристик процессов в момент времени t . В случае, когда приращения $\Delta Y(T)$ не имеют какой-либо достаточно гладкой тенденции во времени и их изменения непредсказуемы (например, на первом же наблюдении $\Delta Y(T)$ может быть достаточно велико по сравнению с самим показателем $Y(T)$), то такие необратимые процессы хаотические [22].

Хорошее качество восстановления данных для выборки объемом 24 значения показывают методы сплайн-интерполяции и экспоненциального сглаживания. Установлено, что метод экспоненциального сглаживания обеспечивает высокую точность восстановления данных при коэффициенте вариации значений временного ряда до 2 %.

Необходимо указать, что эти методы обеспечивают хорошее качество восстановления одиночных пропущенных данных энергопотребления. При отсутствии нескольких данных подряд лучше использовать регрессионный метод.

Ошибки прямых и обратных зависимостей математических моделей восстановления учетных данных химического предприятия

Рассмотрим способы восстановления данных табл. 2 при помощи различных видов математических моделей (табл. 4). Для расчета утраченных данных использовался регрессионный метод, определяющий ошибки моделей для прямых и обратных зависимостей $E=f(A)$ и $E=f(G)$.

Результаты вычислений табл. 4 указывают на то, что при одинаковых значениях коэффициента парной корреляции значительно отличаются значения ошибок прямых и обратных моделей $E=f(A)$ и $E=f(G)$. Таким образом, один и тот же метод восстановления данных не может обеспечить высокую точность на всем интервале изменения физической величины во времени (на временном ряде). Поэтому для практического применения рекомендуется использовать один из методов, определяющих ошибку модели, используя только временной ряд, и один из методов, учитывающий взаимосвязь между физическими величинами.

Таблица 4. Способы восстановления данных при помощи различных видов математических моделей для прямых и обратных зависимостей энергопотребления

Вид математической модели	Объем выборки, ч	$E=f(A)$	$A=f(E)$	Среднее значение	$E=f(G)$	$G=f(E)$	Среднее значение
Exp	24	0,26	2,16	1,21	0,21	0,38	0,29
Fourier		0,27	2,15	1,21	0,30	0,44	0,37
Polynomial		0,28	2,21	1,25	0,21	0,38	0,29
Power		0,28	2,15	1,21	0,21	0,38	0,29
RAT		0,28	2,16	1,22	0,21	0,37	0,29
SumSin		0,27	2,18	1,23	0,31	1,34	0,82
Exp	48	1,13	2,64	1,88	0,95	3,71	2,33
Fourier		3,13	2,69	2,91	1,08	2,99	2,03
Polynomial		1,12	2,79	1,96	1,91	3,73	2,82
Power		1,39	2,69	2,04	2,02	3,74	2,88
RAT		1,11	2,57	1,84	0,97	2,48	1,73
SumSin		0,87	2,57	1,72	0,65	2,06	1,36
Exp	168	0,97	2,26	1,61	0,99	1,55	1,27
Fourier		0,75	6,16	3,45	1,20	1,39	1,29
Polynomial		0,83	1,98	1,41	0,99	1,42	1,21
Power		1,21	2,07	1,64	0,96	1,46	1,21
RAT		0,91	1,91	1,41	0,95	1,22	1,09
SumSin		0,53	1,68	1,11	0,58	1,03	0,80
Exp	350	1,18	1,94	1,56	1,03	1,48	1,26
Fourier		1,14	4,07	2,61	1,10	1,04	1,07
Polynomial		1,14	1,94	1,54	1,03	1,53	1,28
Power		1,31	1,94	1,63	1,03	1,74	1,39
RAT		1,11	1,94	1,52	1,03	1,19	1,11
SumSin		1,24	1,69	1,46	0,85	1,04	0,94
Exp	700	1,66	2,00	1,83	1,37	1,82	1,60
Fourier		1,50	2,08	1,79	1,35	1,70	1,52
Polynomial		1,51	2,13	1,82	1,35	1,71	1,53
Power		1,62	2,09	1,86	1,34	2,14	1,74
RAT		1,47	1,90	1,69	1,32	1,75	1,54
SumSin		1,41	1,85	1,63	1,35	1,53	1,44
Exp	1050	1,46	1,96	1,71	1,71	1,84	1,78
Fourier		1,33	1,95	1,64	1,53	1,67	1,60
Polynomial		1,32	2,12	1,72	1,56	1,71	1,64
Power		1,42	2,08	1,75	1,64	2,07	1,85
RAT		1,36	1,82	1,59	1,61	1,65	1,63
SumSin		1,26	1,83	1,55	1,37	1,61	1,49

СПИСОК ЛИТЕРАТУРЫ

1. Литтл Р.Дж., Рубин Д.Б.А. Статистический анализ данных с пропусками. – М.: Финансы и статистика, 1990. – 336 с.
2. Зинкевич В.С., Штатов Д.К. Информационные риски: анализ и количественная оценка. – М.: Бухгалтерия и бланки. – 2007. – № 1. – С. 50–55.
3. Круглов В.В., Абраменкова И.В. Методы восстановления пропусков в массивах данных // Программные продукты и системы. – 2005. – № 2. URL: <http://www.swsys.ru/index.php?page=article&id=528> (дата обращения: 20.01.2014).
4. Бых А.И., Высоцкая Е.В., Рак Л.И. и др. Выбор метода восстановления пропущенных данных для оценки сердечно-сосудистой деятельности подростков // Восточно-Европейский журнал передовых технологий. – 2010. – № 3. – С. 4–7.
5. Злоба Е.А., Яцкив И.Р. Статистические методы восстановления пропущенных данных // Computer Modelling & New Technologies. – 2004. – V. 6. – С. 51–61.

Выводы

В результате исследований:

- 1) упорядочены способы повышения точности построения однофакторных и многофакторных математических моделей;
- 2) доказано, что лучшее качество восстановления данных энергопотребления обеспечивает метод подбора в группе;
- 3) сведено восстановление данных сложными глобальными методами Барлетта и Resampling к методу парной регрессии, которая проста в вычислении;
- 4) предложено для коротких выборок при восстановлении утерянных одиночных данных энергопотребления использовать сплайн-интерполяцию и метод экспоненциального сглаживания, обеспечивающие ошибку моделей до 1...2 %;
- 5) установлено, что метод экспоненциального сглаживания применим для восстановления данных при коэффициенте вариации значений временного ряда до 2 %;
- 6) доказано, что процесс энергопотребления химического производства является необратимым процессом;
- 7) установлено, что ошибки моделей прямой и обратных парных зависимостей не зависят от коэффициента парной корреляции;
- 8) предложено использовать на практике одновременно один из методов, определяющих ошибку модели, используя только временной ряд, и один из методов, учитывающий взаимосвязь между физическими величинами, для обеспечения высокой точности восстановления данных.

6. Бедерак Я.С., Лутчин Т.Н., Кудрицкий М.Ю. Влияние объема выборки данных энергопотребления на ошибку математической модели // Международный научно-исследовательский журнал. – 2013. – № 12 (Ч. 1). – С. 37–40.
7. Волошко А.В., Лутчин Т.Н., Бедерак Я.С. Восстановление учетных данных энергопотребления на промышленных предприятиях // Материалы VII МНПК. – Москва, 2012. – С. 179–188.
8. Pentland A., Pentland S. Honest Signals: How They Shape Our World. – Cambridge: MIT Press, 2008. – 208 p.
9. Mayer P. Data Recovery: Choosing the Right Technologies. Data-link, 2003.
10. Holden J.M., Bhagwat S.A., Pat K.Y. Development of a multivariate data quality evaluation system // J. Food Compos. Anal. – 2002. – № 15 (4). – С. 339–348.
11. Schafer J., Graham J. Missing data: our view of the state of the art // Psychological Methods. – 2002. – № 7 (2). – С. 147–177.

12. Волошко А.В., Бедерак Я.С., Лутчин Т.М. Проблеми вибору оптимальної математичної моделі енергоспоживання на промислових підприємствах // Восточно-европейский журнал передовых технологий. – 2013. – Вип. 5/8 (65). – С. 19–23.
13. Ивахненко А.Г. Долгосрочное прогнозирование и управление сложными системами. – Киев: Техника, 1975. – 312 с.
14. Стеценко І.В., Бедерак Я.С. Побудова багатofакторних математичних моделей енергоспоживання на хімічному виробництві // Енергосбережение, энергетика, энергоаудит. – 2013. – № 7. – С. 41–48.
15. Находов В.Ф., Стеценко І.В., Бедерак Я.С. Застосування методів самоорганізації математичних моделей енергоспоживання для встановлення «стандартів» в системах оперативного контролю енергоефективності // Енергосбережение, энергетика, энергоаудит. – 2010. – № 5. – С. 23–33.
16. Дьяконов В., Круглов В. Математические пакеты расширения Matlab. Специальный справочник. – СПб.: Питер, 2001. – 480 с.
17. Відновлення втрачених облікових / А.В. Волошко, Я.С. Бедерак, Т.М. Лутчин, Д.К. Міщенко // Вісник КНУ ім. М. Остроградського. – 2012. – Т. 2 (73). – С. 426–428.
18. Лук'яненко І.Г., Краснікова Л.І. Економетрика: Підручник. – К.: Знання, 1998. – 494 с.
19. Ивахненко А.Г., Мюллер Й.А.К. Самоорганизация прогнозирующих моделей. – К.: Наукова думка, 1985. – 219 с.
20. Горбунов В.М. Теория принятия решений. – Томск: Изд-во ТПУ, 2010. – 67 с.
21. Zwillinger D. CRC Standard Mathematical Tables and Formulae. – Boca Raton: CRC Press, 2003. – 857 с.
22. Светульников С.Г., Светульников И.С. Методы социально-экономического прогнозирования. – СПб.: Изд-во СПбГУЭФ, 2009. – Т. I. – 147 с.
23. Новицкий П.В., Зograф И.А. Оценка погрешностей результатов измерений. – Л.: Энергоатомиздат, 1985. – 302 с.

Поступила 30.03.2014 г.

UDC 517.4:519.652

THE PROBLEM OF ACCOUNTING DATA RECOVERY ON CHEMICAL ENTERPRISE

Anatoliy V. Voloshko,

Cand. Sc., Institute for Energy Saving and Energy Supply within National Technical University of Ukraine «Kiev Polytechnic Institute»,
115, Borshchagivska Street, Kiev, 03056, Ukraine. E-mail: a-voloshko@yandex.ru

Yaroslav S. Bederak,

AZOT, 72, Pervomayska Str., Cherkassy, 18016, Ukraine.
E-mail: yarbak@yandex.ua

Tetiana M. Lutchyn,

Institute for Energy Saving and Energy Supply within National Technical University of Ukraine «Kiev Polytechnic Institute»,
115, Borshchagivska Street, Kiev, 03056, Ukraine. E-mail: t.lutchyn@gmail.com

Maxim Yu. Kudritskiy,

Institute for Energy Saving and Energy Supply within National Technical University of Ukraine «Kiev Polytechnic Institute»,
115, Borshchagivska Street, Kiev, 03056, Ukraine. E-mail: bugs_bunni@ukr.net

Relevance of the work is caused by the presence of missing data in the readings of energy meters.

The main aim of the research is to study the choice of method for recovering missing data on energy consumption in industry.

The methods used in the study: the models are calculated using the application Curve Fitting Toolbox of the software complex «Matlab 7.0». The library of graphical models Curve Fitting Toolbox includes an application cftool, which allows defining a parametric model (such as, exponential function Exp, polynomial Polynomial, rational RAT, as well as the sum of sinusoidal functions SumSin), selecting parameters, analyzing approach suitability, displaying the result graphically. In the library of graphical models Curve Fitting Toolbox the models from more than 50 different mathematical functions are determined by search method.

The results: The paper describes the features of simple and complex data recovery methods with further estimation of their errors and indicates the ways to improve the accuracy of n-factor models. The authors have studied direct and inverse dependences of recovering lost accounting data for a chemical enterprise. The optimal limits of initial research data samples are proved. The paper also provides options for defining the best methods for value recovery in cases of their absence.

Key words:

Data recovery, n-factor models, model error, energy consumption, lost data.

REFERENCES

1. Littl R.J., Rubin D.B.A. *Statisticheskii analiz dannykh s propuskami* [Statistical analysis of data with gaps]. Moscow, Finansy i statistika Publ., 1990. 336 p.
2. Zinkevich V.S., Shtatov D.K. Informatsionnye riski: analiz i kolichestvennaya otsenka [Information Risk: analysis and quantification]. Moscow, Bukhgalteriya i blanki Publ., 2007, no. 1, pp. 50–55.
3. Kruglov V.V., Abramenkova I.V. Metody vosstanovleniya propuskov v massivakh dannykh [Recovery Methods omissions in the data]. *Programmnyye produkty i sistemy*, 2005, no. 2. Available at: <http://www.swsys.ru/index.php?page=article&id=528> (accessed 20.01.2014).
4. Bykh A.I., Vysotskaya E.V., Rak L.I. Vybory metoda vosstanovleniya propushchennykh dannykh dlya otsenki serdechno-sosudistoy deyatel'nosti podrostkov [Selecting a method of reconstructing the missing data for evaluating cardiovascular activity of teenagers]. *Vostochno-Evropeyskiy zhurnal peredovykh tekhnologiy*, 2010, no. 3, pp. 4–7.
5. Zloba E.A., Yatskiv I.R. Statisticheskiye metody vosstanovleniya propushchennykh dannykh [Statistical methods for recovering missing data]. *Computer Modelling & New Technologies*, 2004, vol. 6, pp. 51–61.
6. Bederak Ya.S., Lutchyn T.M., Kudritskiy M.Yu. Vliyaniye obyema vyborke dannykh energopotrebleniya na oshibku matematicheskoy modeli [Influence of data sampling of energy consumption on mathematical model error]. *Mezhdunarodny nauchno-isledovatel'skiy zhurnal*, 2013, no. 12 (P. 1), pp. 37–40.
7. Voloshko A.V., Lutchyn T.M., Bederak Ya.S. Vosstanovleniye uchetykh dannykh energopotrebleniya na promyshlennyykh predpriyatiyakh [Energy recovery measurements in industrial enterprises]. *Materialy VII MNPk* [Materials of VII MNPk]. Moscow, 2012, pp. 179–188.
8. Pentland A., Pentland S. *Honest Signals: How They Shape Our World*. Cambridge, MIT Press, 2008. 208 p.
9. Mayer P. Data Recovery: Choosing the Right Technologies. Data-link, 2003.
10. Holden J.M., Bhagwat S.A., Pat K.Y. Development of a multinutrient data quality evaluation system. *J. Food Compos. Anal.*, 2002, no. 15 (4), pp. 339–348.
11. Schafer J., Graham J. Missing data: our view of the state of the art. *Psychological Methods*, 2002, no. 7 (2), pp. 147–177.
12. Voloshko A.V., Bederak Ya.S., Lutchyn T.M. Problemy vboru optimal'noy matematicheskoy modeli energospozhivaniya na promyshlennykh podpriemstvakh [Problems of choosing the optimal mathematical model of energy consumption in industrial enterprises]. *Vostochno-yevropeyskiy zhurnal peredovykh tekhnologiy*, 2013, vol. 5/8 (65), pp.19–23.
13. Ivakhnenko A.G. *Dolgosrochnoye prognozirovaniye i upravleniye slozhnyimi sistemami* [Long-term forecasting and management of complex systems]. Kiev, Tekhnika Publ., 1975. 312 p.
14. Stetsenko I.V., Bederak Ya.S. Pobudova bagatofaktornikh matematicheskikh modeley energospozhivaniya na khimichnomu virobnitstvi [Construction of multivariate mathematical models of power consumption by the chemical industry]. *Energoberezhniye, energetika, energoaudit*, 2013, no. 7, pp. 41–48.
15. Nakhodov V.F., Stetsenko I.V., Bederak Ya.S. Zastosuvannya metodiv samoorganizatsii matematicheskikh modeley energospozhivaniya dlya vstanovlennya «standartiv» v sistemakh operativnogo kontrolyu energoefektivnosti [Application of self-organizing power of mathematical models for establishing «standards» in the system of operational control efficiency]. *Energoberezhniye, energetika, energoaudit*, 2010, no. 5, pp. 23–33.
16. Dyakonov V., Kruglov V. *Matematicheskiye pakety rasshireniya Matlab. Spetsialny spravochnik* [Mathematical packets Expansion Matlab. Special Directory]. Saint Petersburg, Piter Publ., 2001. 480 p.
17. Voloshko A.V., Bederak Ya.S., Mishchenko D.K., Lutchyn T.M. Vidnovlennya vtrachenykh oblikovykh [Recovery of lost accounting data]. *Visnyk KNU by M. Ostrogradskogo*, 2012, vol. 2 (73), pp. 426–428.
18. Lukyanenko I.G., Krasnikova L.I. *Ekonometrika: Pidruchnik* [Econometrics: Textbook]. Kiev, «Znannya», 1998. 494 p.
19. Ivakhnenko A.G., Myuller Y.A.K. *Samoorganizatsiya prognozirovushchikh modeley* [Self-organization of predictive models]. Kiev, Naukova dumka, 1985. 219 p.
20. Gorbunov V.M. *Teoriya prinyatiya resheniy* [Decision theory]. Tomsk, TPU, 2010. 67 p.
21. Zwillinger D. *CRC Standard Mathematical Tables and Formulae*. Boca Raton, CRC press, 2003. 857 p.
22. Svetunkov S.G., Svetunkov I.S. *Metody sotsialno-ekonomicheskogo prognozirovaniya* [Methods of social and economic forecasting]. Saint Petersburg SPbGUEF Press, 2009, Vol. I, 147 p.
23. Novitskiy P.V., Zograf I.A. *Otsenka pogreshnostey rezultatov izmereniy* [Evaluation of errors in measurement results]. Leningrad, Energoatomizdat Publ., 1985. 302 p.

УДК 389.14:621.317

ОРГАНИЗАЦИЯ ДИСТАНЦИОННОЙ КАЛИБРОВКИ СРЕДСТВ ИЗМЕРЕНИЙ ЭЛЕКТРИЧЕСКИХ ВЕЛИЧИН

Величко Олег Николаевич,

д-р техн. наук, профессор, заслуженный деятель науки и техники Украины,
директор Научно-производственного института метрологического
обеспечения электромагнитных величин Государственного предприятия
«Всеукраинский государственный научно-производственный
центр стандартизации, метрологии, сертификации и защиты прав
потребителей» (ГП «Укрметрестандарт»), Украина, 03680, г. Киев,
ул. Метрологическая, д. 4. E-mail: velychko@ukrcsm.kiev.ua

Гурин Роман Витальевич,

инженер Научно-производственного института метрологического
обеспечения электромагнитных величин ГП «Укрметрестандарт»,
Украина, 03680, г. Киев, ул. Метрологическая, д. 4. E-mail: gurin@ukr.net

Баранов Павел Федорович,

канд. техн. наук, инженер кафедры компьютерных измерительных систем
и метрологии Института кибернетики Томского политехнического
университета, Россия, 634050, г. Томск, пр. Ленина, д. 30.
E-mail: bpf@tpu.ru

Актуальность работы обусловлена необходимостью развития и совершенствования процедур калибровки средств измерений. По сравнению с обычной калибровкой дистанционная калибровка обеспечивает экономический выигрыш, т. к. все необходимое для калибровки оборудования находится в лаборатории заказчика, и нет необходимости транспортировать его к месту проведения калибровки.

Цель работы: выявить особенности построения системы дистанционной калибровки через Интернет средств измерения электрических величин.

Методы исследования: теория криптография.

Результаты: предложен способ защиты передаваемых данных измерений при проведении дистанционной калибровки. Представлены результаты экспериментального исследования разработанной системы дистанционной калибровки с использованием локальной сети предприятия, а также с использованием глобальной сети Интернет с размещением элементов системы в разных странах. Экспериментальные исследования разработанной системы дистанционной калибровки с использованием алгоритма шифрования данных Blowfish показали, что общее время проведения калибровки при локальном и глобальном размещении отличается незначительно, от 8,6 до 24,4 % в зависимости от калибруемого параметра, что подтверждает возможность использования разработанной системы для дистанционной калибровки средств измерений электрических величин.

Ключевые слова:

Дистанционная калибровка, данные измерений, средства измерений, программное обеспечение, защита данных калибровки, Интернет, электрические величины.

Введение

Развитие средств автоматизации и измерений, персональных компьютеров и сетевых технологий привело к широкому распространению распределенных измерительных информационных систем, построенных на основе интегрирования средств измерений (СИ) и персональных компьютеров (ПК) [1–3]. Такие системы используются как для автоматизации технологических процессов [4, 5], так и при построении интерактивных образовательных сред [6, 7].

Любое СИ требует проведения периодической поверки или калибровки. Калибровку СИ проводят аккредитованные на право проведения калибровочных работ метрологические службы. Использование современных информационных и те-

лекоммуникационных технологий позволяет организовать дистанционное сличения метрологических характеристик СИ [8–12] непосредственно на местах их эксплуатации.

Впервые было заявлено о возможности проведения дистанционной калибровки через Интернет в 1997 г. [13], а в 2000 г. Национальный институт стандартов и технологий США (NIST) провел первую дистанционную калибровку СИ с использованием многофункционального мобильного рабочего эталона (РЭ) для лаборатории фирмы «Sandia» [14].

Основным преимуществом дистанционной калибровки СИ является минимизация временных и экономических затрат на транспортировку СИ из лаборатории заказчика в калибровочную лабора-

торию, поскольку все РЭ и СИ находятся в лаборатории заказчика.

При проведении дистанционной калибровки СИ возникают следующие проблемы [15–18]:

- требуется специализированное программное обеспечение (ПО) для дистанционного управления СИ в реальном масштабе времени и передачи результатов измерений из лаборатории заказчика в калибровочную лабораторию;
- результаты измерения должны быть надежно защищены с использованием специализированного ПО, как в процессе самой калибровки, так и при передаче результатов измерений;
- необходимо осуществлять дистанционный контроль климатических условий в лаборатории заказчика;
- необходимо обеспечивать выполнение требований безопасности в лаборатории заказчика при проведении дистанционной калибровки СИ и осуществлять соответствующий визуальный контроль.

Из перечисленных проблем наибольшую актуальность представляет защита результатов измерений при их передаче через сеть Интернет. Основной угрозой является несанкционированное подключение к каналу передачи данных и модификация результатов измерений. Поэтому для защиты данных, передаваемых через сеть Интернет, необходимо применять соответствующие средства криптографической защиты [19, 20].

В статье описана разработанная система для дистанционной калибровки СИ электрических величин и представлены результаты ее экспериментального тестирования с использованием средств криптографической защиты данных.

Особенности разработанной системы дистанционной калибровки

Обобщенная структурная схема разработанной системы для дистанционной калибровки СИ приведена на рис. 1. Функциональная схема разработанной системы дистанционной калибровки для СИ электрических величин представлена на рис. 2.

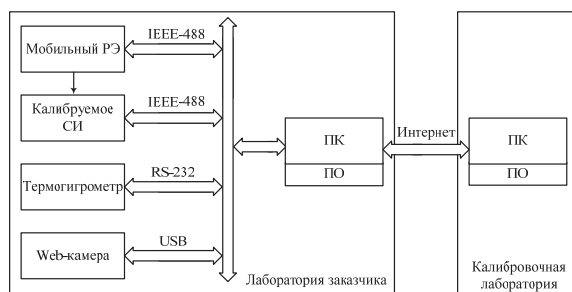


Рис. 1. Обобщенная структурная схема системы для дистанционной калибровки СИ

Основными компонентами предложенной системы являются:

- специализированное ПО, устанавливаемое на ПК в лаборатории заказчика и на ПК в калибровочной лаборатории;

- многофункциональный мобильный РЭ, размещаемый в лаборатории заказчика;
- калибруемое СИ электрических величин, размещенное в лаборатории заказчика;
- измеритель параметров воздуха «Атмосфера-1» (термогигрометр) и Web-камера, установленные в лаборатории заказчика;
- канал связи сети Интернет.

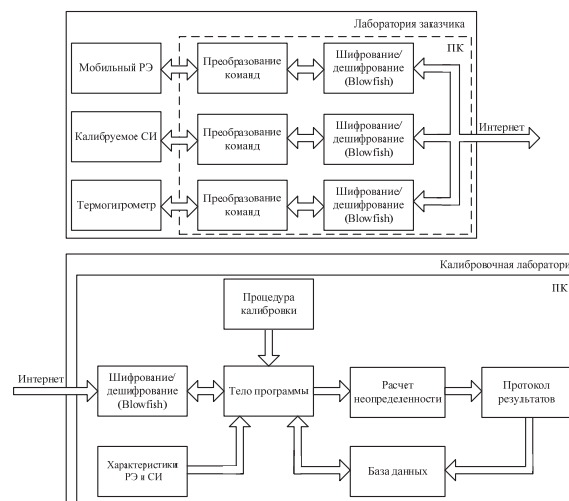


Рис. 2. Функциональная схема разработанной системы дистанционной калибровки для СИ электрических величин

Специализированное ПО разработано на языке графического программирования LabVIEW и состоит из двух основных частей: основной программы (программа-клиент) и программы управления устройствами (программа-сервер). Программа-клиент размещена на ПК в калибровочной лаборатории, а программа-сервер на ПК в лаборатории заказчика.

Процедура дистанционной калибровки СИ

В калибровочной лаборатории и лаборатории заказчика могут быть установлены сетевые экраны, поэтому для взаимодействия ПО создается виртуальная сеть с помощью программы TeamViewer с поддержкой VPN драйвера. TeamViewer используется в режиме полного доступа с генерацией случайных паролей при его запуске. Процедура создания виртуальной сети осуществляется в следующем порядке.

1. На ПК в калибровочной лаборатории с помощью программы TeamViewer создается видео и чат конференция.
2. В лабораторию заказчика передается идентификационный номер (ID) конференции (например, по e-mail), пароль (например, по чату конференции), и осуществляется вход в систему дистанционной калибровки.
3. Запускается процесс тестирования связи между ПК в калибровочной лаборатории и ПК в лаборатории заказчика. Если связь установлена, выдается разрешение на осуществление работ по дистанционной калибровке СИ.

На обоих ПК должен быть установлен бесплатный распространяемый плагин LabVIEW Runtime Engine для выполнения законченных приложений, разработанных в среде LabVIEW. При запуске программы-клиента для дистанционной калибровки СИ производится авторизация и загружается конфигурационный файл процедуры калибровки. На рис. 3 представлен внешний вид программы-клиента для дистанционной калибровки.

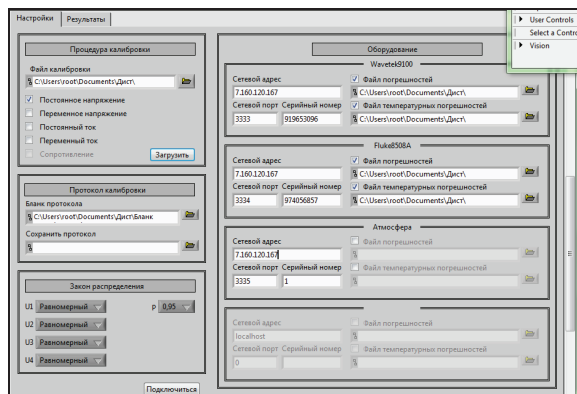


Рис. 3. Внешний вид разработанной программы-клиента для дистанционной калибровки

До начала калибровки отмечаются величины, для которых осуществляется калибровка СИ (постоянное напряжение и ток, переменные напряжение и ток и т. д.). Для расчета комбинированной стандартной неопределенности измерений при калибровке СИ предварительно вводятся значения таких составляющих неопределенности как: U_1 – неопределенность, вызванная погрешностью эталона; U_2 – неопределенность, вызванная температурным дрейфом эталона; U_3 – неопределенность, вызванная конечной разрешающей способностью эталона (измерителя); U_4 – суммарный закон распределения выходной величины, P – уровень доверительной вероятности.

Затем вводятся сетевые адреса и порты, серийные номера используемого эталонного СИ, дополнительного (для контроля параметров окружающей среды) оборудования и калибруемого СИ. Загружаются специально созданные файлы с составляющими неопределенности и температурных зависимостей для используемого эталонного, дополнительного оборудования и калибруемого СИ. Процесс калибровки СИ начинается после того, как сотрудник в лаборатории заказчика получит сообщение с предупреждением о начале процесса калибровки и подтвердит свою готовность.

Программа-сервер считывает серийный номер РЭ, дополнительного оборудования и калибруемого СИ, сравнивает их с серийными номерами, полученными от программы-клиента из калибровочной лаборатории. Если эти номера не совпадают, программа-сервер прекращает процедуру дистанционной калибровки СИ с выводом специального сообщения на оба ПК. Если идентификация обо-

дования и СИ прошла успешно, программа-сервер начинает последовательно получать команды от программы-клиента, устанавливая для используемого оборудования те или иные режимы и параметры (установка частоты, амплитуды, выполнение измерения и т. д.).

По окончании калибровки СИ по одному из параметров на стороне калибровочной лаборатории будет предупреждение о начале калибровки СИ по следующему параметру, после чего уполномоченный персонал в лаборатории заказчика должен перекоммутировать схему измерения. Если было выбрано несколько параметров для калибровки СИ, то таблица с результатами калибровки СИ будет обновляться после получения необходимых результатов измерений (рис. 4).

Рис. 4. Внешний вид окна результатов калибровки СИ

Следует отметить, что конфигурационный файл процедуры калибровки находится вне тела основной программы и может легко изменяться без изменения самой программы, что позволяет проводить калибровку большой номенклатуры СИ.

Защита передаваемых данных измерений

Для защиты передаваемых данных измерений из лаборатории заказчика в калибровочную лабораторию криптопреобразования могут осуществляться аппаратным или программным способом. В первом случае между интерфейсом СИ (IEEE-488, RS-232, USB) и интерфейсом ПК подключается специальный модуль, который производит шифрование и дешифрование данных [19]. Достоинством данного способа реализации криптографической защиты является то, что данные проходят максимальный путь, являясь зашифрованными (от ПК в лаборатории заказчика до ПК в калибровочной лаборатории и наоборот), но его реализация влечет за собой дополнительные затраты на разработку аппаратной и программной частей криптографического модуля.

В случае, когда необходимо обеспечить защиту информации на сетевом уровне, рационально использовать криптографические алгоритмы, реализуемые программным путем. Существует множе-

ство криптографических алгоритмов отличающихся между собой сложностью реализации, скоростью шифрования (дешифрования) и криптостойкостью.

При реализации системы дистанционной калибровки СИ для защиты данных, передаваемых по сети, был выбран алгоритм шифрования данных Blowfish, достоинствами которого являются: простота реализации, высокая скорость шифрования, криптоустойчивость и свободное распространение.

Основные параметры алгоритма Blowfish: размер ключа – от 32 до 448 бит; размер блока – 64 бита; число раундов – 16; тип – сеть Фейстеля.

Алгоритм Blowfish является симметричным шифром, который подразумевает использование одного ключа, как на стороне сервера, так и на стороне клиента (рис. 5). Недостатком выбранного алгоритма и аналогичных является необходимость передачи ключа от одной стороны к другой с использованием защищенных каналов связи или непосредственно на информационных носителях.

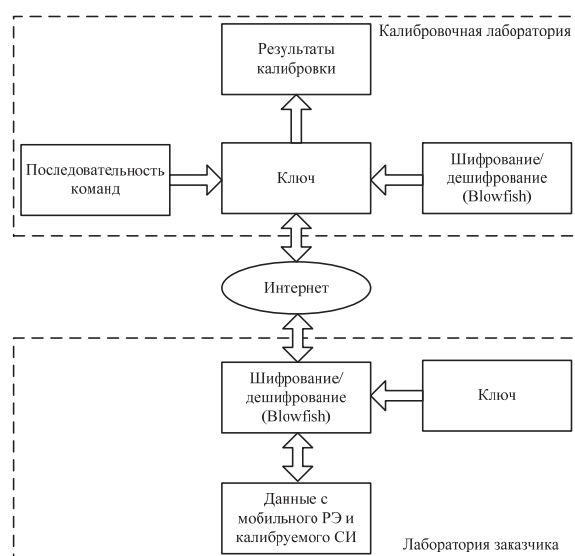


Рис. 5. Обобщенная схема передачи данных измерений

Таблица. Результаты калибровки при применении разработанной системы

Диапазон	Точки калибровки		Время калибровки, мин:с					
			Локальная сеть			Глобальная сеть		
			1	2	3	1	2	3
Постоянное напряжение								
320 мВ	0; ±300 мВ		3:20	3:21	3:22	4:04	4:07	4:00
3,2 В	±0,45; ±1,5; ±3,0 В							
32 В	±4,5; ±30 В							
320 В	±45; ±300 В							
1050 В	±330; ±1000 В							
Средняя величина:			3:21 (100 %)			4:04 (121,4 %)		
Постоянный ток								
320 мкА	0; ±320 мкА		3:21	3:20	3:21	4:11	4:08	4:11
3,2 мА	±0,45; ±3,0 мА							
32 мА	±4,5; ±30 мА							
320 мА	±45; ±300 мА							
3,2 А	±0,45; ±3,0 А							
10,5 А	±3,3; ±10,0 А							
Средняя величина:			3:21 (100 %)			4:10 (124,4 %)		
Переменное напряжение								
320 мВ	0,3 В	0.01; 0.04; 1; 3; 10; 30; 50; 100 кГц	17:15	17:14	17:21	18:40	18:42	18:55
3,2 В	1,5 В	100 кГц						
	3,0 В	0,01; 0,04; 1; 3; 10; 30; 50; 100 кГц						
32 В	30 В							
105 В	100 В							
320 В	300 В							
800 В	330 В	30 кГц						
	500 В	20 кГц						
	750 В	0,04; 0,1; 0,2; 3; 10 кГц						
Средняя величина:			17:17 (100 %)			18:46 (108,6 %)		
Переменный ток								
320 мкА	300 мкА	0,01;0,04; 0,1; 3; 10 кГц	8:02	8:04	8:02	9:00	8:58	9:05
3,2 мА	3,0 мА							
32 мА	30 мА							
320 мА	300 мА							
3,2 А	3,0 А							
10,5 А	10 А							
Средняя величина:			8:03 (100 %)			9:01 (112,0 %)		
Всего:			31:59	31:58	32:06	35:56	35:55	36:11
Общая средняя величина:			32:01 (100 %)			36:01 (112,5 %)		

Экспериментальные исследования разработанной системы дистанционной калибровки

Экспериментальные исследования разработанной системы дистанционной калибровки СИ электрических величин проводились с использованием локальной сети предприятия (в разных лабораториях одной организации), а также с использованием глобальной сети Интернет с размещением элементов системы в разных странах (лаборатория заказчика в г. Киев, Украина, калибровочная лаборатория в г. Томск, Россия).

В качестве эталонного СИ был использован универсальный калибратор Wavetek 9100, в качестве калибруемого СИ был использован вольтметр Fluke 8508A. Для каждого из размещений было проведено по три сеанса дистанционной калибровки СИ. При этом измерялось время проведения калибровки СИ для каждого калибруемого параметра в каждой точке калибровки для всех диапазонов (начало синхронизировалось со стороны калибровочной лаборатории с лабораторией заказчика). Рассчитаны средние величины (время калибровки) для каждого калибруемого параметра СИ и общая средняя величина (время калибровки) при двух размещении системы дистанционной калибровки. При этом общая средняя величина получена путем суммирования всех средних величин без учета времени перекоммутации схемы измерений

для разных параметров калибровки в лаборатории заказчика уполномоченным персоналом.

Результаты калибровки при применении разработанной системы приведены в таблице.

Выводы

1. При реализации системы дистанционной калибровки для защиты данных, передаваемых по сети, предложено использовать алгоритм шифрования данных Blowfish, достоинствами которого являются: простота реализации, высокая скорость шифрования, криптоустойчивость и свободное распространение.
2. Результаты экспериментального исследования разработанной системы дистанционной калибровки с применением алгоритма шифрования данных Blowfish с использованием локальной сети предприятия, а также с использованием глобальной сети Интернет с размещением элементов системы в разных странах показали, что общее время проведения калибровки при локальном и глобальном размещении отличается незначительно, от 8,6 до 24,4 % в зависимости от калибруемого параметра, что подтверждает возможность использования разработанной системы для дистанционной калибровки средств измерений электрических величин.

СПИСОК ЛИТЕРАТУРЫ

1. Benetazzo L. et al. A web-based distributed virtual laboratory // IEEE Transactions on Instrumentation and Measurement. – 2000. – V. 49. – № 2. – P. 349–356.
2. A client-server architecture for distributed measurement systems / M. Bertocco, F. Ferraris, C. Offelli, M. Parvis // IEEE Transactions on Instrumentation and Measurement. – 1998. – V. 47. – № 5. – P. 1143–1148.
3. De Capua C., Liccario A., Morello R. On the web service-based remote didactical laboratory: further developments and improvements // Proceedings of Instrumentation and Measurement Technology Conference IMTC-2005. – Canada, Ottawa, May 16–19, 2005. – P. 1692–1696.
4. Wu Xu-Li, Zhao Yu-Xiao. Exploit of remote monitoring and control system for bell prover device based on virtual instrument // Proceedings of International Symposium on Computational Intelligence and Design ISCID-2012. – Hangzhou, October 28–29, 2012. – P. 178–181.
5. Cristaldi L., Ferrero A., Salicone S. A distributed system for electric power quality measurement // IEEE Transactions on Instrumentation and Measurement. – 2002. – V. 51. – № 4. – P. 776–781.
6. Cristaldi L., Ferrero A., Piuri V. Programmable instruments, virtual instruments, and distributed measurement systems: What is really useful, innovative, and technically sound // IEEE Transactions on Instrumentation and Measurement. – 1999. – V. 42. – № 3. – P. 20–27.
7. Spoedler H.J.W. Virtual instruments and virtual environments // IEEE Instrumentation & Measurement Magazine. – 1999. – V. 2. – № 3. – P. 14–19.
8. Remote calibration using mobile multiagent technology / M.M. Albu, A. Ferrero, F. Mihai, S. Salicone // IEEE Transactions on Instrumentation and Measurement. – 2005. – V. 54. – № 1. – P. 24–30.
9. Dimitrijevic B.R., Simic M.M. Remote wireless calibration of equipment in the distributed measurement systems // Proceedings of Telecommunications in Modern Satellite, Cable and Broadcasting Services TELSIKS-2007. – Serbia, September 26–28, 2007. – P. 479–482.
10. Dudley R.A., Ridler N.M. Internet calibration direct to national measurement standards for automatic network analyzers // Proceedings of Instrumentation and Measurement Technology Conference IMTC-2001. – Budapest, May 21–23, 2001. – P. 255–258.
11. A remote calibration system for industrial thermometers / Le Chen, Yiyi Shao, Yaqiong Fu, Min Xie // Proceedings of Industrial Electronics and Applications ICIEA-2013. – Melbourne, June 19–21, 2013. – P. 1574–1577.
12. Remote time and frequency calibration system for telecommunication synchronization applications / Jia-Lun Wang, Chang P.C., Shinn-Yan Lin, Huang-Tien Lin, Chia-Shu Liao // Proceedings of Network Operations and Management Symposium APNOMS-2011. – Taipei, September 21–23, 2011. – P. 1–6.
13. Remote characterization of optoelectronic devices over the internet / R. O'Dowd, D. Maxwell, T. Farrell, J. Dunne // Proceedings of 4th Optical Fibre Measurement Conference. – Teddington, UK, October 24–27, 1997. – P. 155–158.
14. Internet-based calibration of a multifunction calibrator / L.B. Baca, L. Duda, R. Walker, N. Oldham, M. Parker // Proceedings of National Conference of Standards Laboratories. – Toronto, Canada, April 12–14, 2000. – P. 10–12.
15. Величко О.М. Калибровка засобів вимірювальної техніки через Інтернет: стан і перспективи впровадження // Український метрологічний журнал. – 2006. – № 1. – С. 45–49.
16. Величко О.Н. Особенности внедрения современных информационных технологий в аккредитованных лабораториях // Метрология и метрологическое обеспечение: Междунар. научно-практ. конф. – Минск, 2007. – С. 30–34.

17. Величко О., Гурін Р. Калібрування цифрових мультиметрів і калібраторів електричних сигналів із застосуванням мережі Інтернет // Метрологія та прилади. – 2011. – № 5. – С. 51–55.
18. Величко О.М., Анохін Ю.Л. Виконання вимог міжнародного стандарту ISO/IEC 17025 при дистанційному калібруванні засобів виміральної техніки // Метрологія та прилади. – 2013. – № 2 (40). – С. 37–41.
19. Jurčević M., Hegeduš H., Golub M. Generic system for remote testing and calibration of measuring instruments: security architecture // Measurement science review. – 2010. – V. 10. – № 2. – P. 50–55.
20. Carullo A., Parvis M., Vallan A. Security issues for internet-based calibration activities // Proceedings of Instrumentation and Measurement Technology Conference IMTC-2002. – Budapest, May 21–23, 2002. – P. 817–822.

Поступила 18.02.2014 з.

UDC 389.14:621.317

ORGANIZATION OF REMOTE CALIBRATION FOR MEASURING INSTRUMENTS OF ELECTRICAL QUANTITIES

Oleg N. Velichko,

Dr. Sc., State Enterprise «All-Ukrainian State Research and Production Center for Standardization, Metrology, Certification and Consumer's Rights Protection» (SE «Ukrmetrteststandard»), 4, Metrologicheskaya st., Kiev, 03680, Ukraine. E-mail: velychko@ukrcsm.kiev.ua.

Roman V. Gurin,

State Enterprise «All-Ukrainian State Research and Production Center for Standardization, Metrology, Certification and Consumer's Rights Protection» (SE «Ukrmetrteststandard»), 4, Metrologicheskaya st., Kiev, 03680, Ukraine. E-mail: gurin@ukr.net

Pavel F. Baranov,

Cand. Sc., Tomsk Polytechnic University, 30, Lenin Avenue, Tomsk, 634050, Russia. E-mail: bpf@tpu.ru

The importance of the work is conditioned by the need to develop and to improve calibration procedures of measuring instruments. In contrast to conventional calibration, a remote calibration provides economic benefits due to all necessary equipment is in the customer's laboratory; there is no need to transport it to the calibration location.

The main aim of the study: to identify the features of building the remote calibration system for measuring instruments of electrical quantities via the Internet.

The methods used in the study: cryptographic security.

The results: The authors have proposed the method for secure transmission of measurement data during remote calibration. The paper introduces the experimental results for the developed remote calibration system using the local network, as well as using the Internet global network with the deployment of the system instruments in different countries. The experimental investigations of the developed remote calibration system using Blowfish encryption algorithm shown that the total calibration time under local and global deployments differs insignificantly, from 8,6 to 24,4 % in dependence on quantities to be calibrated. This confirms the possibility to use the developed system for remote calibration of electrical measuring instruments.

Key words:

Remote calibration, measurement data, measuring instrument, software, calibration data protection, Internet, electrical quantities.

REFERENCES

1. Benetazzo L. A web-based distributed virtual laboratory. *IEEE Transactions on Instrumentation and Measurement*, 2000, vol. 49, no. 2, pp. 349–356.
2. Bertocco M., Ferraris F., Offelli C., Parvis M. A client-server architecture for distributed measurement systems. *IEEE Transactions on Instrumentation and Measurement*, 1998, vol. 47, no. 5, pp. 1143–1148.
3. De Capua C., Liccario A., Morello R. On the web service-based remote didactical laboratory: further developments and improvements. *Proceedings of Instrumentation and Measurement Technology Conference IMTC-2005*. Canada, Ottawa, May 16–19, 2005. pp. 1692–1696.
4. Wu Xu-Li, Zhao Yu-Xiao Exploit of Remote Monitoring and Control System for Bell Prover Device Based on Virtual Instrument. *Proceedings of International Symposium on Computational Intelligence and Design ISCID-2012*. Hangzhou, October 28–29, 2012. pp. 178–181.
5. Cristaldi L., Ferrero A., Salicone S. A distributed system for electric power quality measurement. *IEEE Transactions on Instrumentation and Measurement*, 2002, vol. 51, no. 4, pp. 776–781.
6. Cristaldi L., Ferrero A., Piuri V. Programmable instruments, virtual instruments, and distributed measurement systems: What is

- really useful, innovative, and technically sound. *IEEE Transactions on Instrumentation and Measurement*, 1999, vol. 42, no. 3, pp. 20–27.
7. Spiedler H.J.W. Virtual instruments and virtual environments. *IEEE Instrumentation & Measurement Magazine*, 1999, vol. 2, no. 3, pp. 14–19.
8. Albu M.M., Ferrero A., Mihai F., Salicone S. Remote calibration using mobile multiagent technology. *IEEE Transactions on Instrumentation and Measurement*, 2005, vol. 54, no. 1, pp. 24–30.
9. Dimitrijevic B.R., Simic M.M. Remote wireless calibration of equipment in the distributed measurement systems. *Proceedings of Telecommunications in Modern Satellite, Cable and Broadcasting Services TELSIS-2007*. Serbia, September 26–28, 2007. pp. 479–482.
10. Dudley R.A., Ridler N.M. Internet calibration direct to national measurement standards for automatic network analyzers. *Proceedings of Instrumentation and Measurement Technology Conference IMTC-2001*. Budapest, May 21–23, 2001. pp. 255–258.
11. Le Chen, Yiyi Shao, Yaqiong Fu, Min Xie. A remote calibration system for industrial thermometers. *Proceedings of Industrial Electronics and Applications ICIEA-2013*. Melbourne, June 19–21, 2013. pp. 1574–1577.
12. Jia-Lun Wang, Chang P.C., Shinn-Yan Lin, Huang-Tien Lin, Chia-Shu Liao. Remote time and frequency calibration system for telecommunication synchronization applications. *Proceedings of Network Operations and Management Symposium APNOMS-2011*. Taipei, September 21–23, 2011. pp. 1–6.
13. O'Dowd R., Maxwell D., Farrell T., Dunne J. Remote characterization of optoelectronic devices over the internet. *Proceedings of 4th Optical Fibre Measurement Conference*. Teddington, UK, October 24–27, 1997. pp. 155–158.
14. Baca L. B., Duda L., Walker R., Oldham N., Parker M. Internet-Based Calibration of a Multifunction Calibrator. *Proceedings of National Conference of Standards Laboratories*. Toronto, Canada, April 12–14, 2000. pp. 10–12.
15. Velichko O.N. Kalibrovka zasobiv vimiryuvalnoy tekhniki cherez Internet: stan i perspektivi vprovadzhennya [Calibration of measuring instruments over the Internet: state and prospects of implementation]. *Ukrainskiy metrologicheskii zhurnal – Ukrainian Journal of Metrology*, 2006, no. 1, pp. 45–49.
16. Velichko O.N. Osobennosti vnedreniya sovremennykh informatsionnykh tekhnologiy v akkreditovannykh laboratoriyakh [Features of introducing modern information technologies in accredited laboratories]. *Metrologiya i metrologicheskoe obespechenie. Mezhdunarodnaya nauchno-prakticheskaya konferentsiya* [International scientific-practical conference. Metrology and measurement assurance]. Minsk, 2007, pp. 30–34.
17. Velichko O., Gurin R. Kalibruvannya tsifrovikh multimetriv i kalibratoriv elektrichnikh signaliv iz zastosuvannyam merezhi Internet [Calibration of digital multimeters and calibrators electrical signals using the Internet]. *Metrologiya ta priladi – Metrology and Instruments*, 2011, no. 5, pp. 51–55.
18. Velichko O.N., Anokhin Yu.L. Vikonannya vimog mizhnarodnogo standartu ISO/IEC 17025 pri distantsiynomu kalibruvanni zasobiv vimiryuvalnoy tekhniki [Compliance with the requirements of the international standard ISO/IEC 17025 for remote calibration of measuring instruments]. *Metrologiya ta priladi – Metrology and Instruments*, 2013, no. 2 (40), pp. 37–41.
19. Jurčević M., Hegeduš H., Golub M. Generic system for remote testing and calibration of measuring instruments: security architecture. *Measurement science review*, 2010, vol. 10, no. 2, pp. 50–55.
20. Carullo A., Parvis M., Vallan A. Security issues for internet-based calibration activities. *Proceedings of Instrumentation and Measurement Technology Conference IMTC-2002*. Budapest, May 21–23, 2002. pp. 817–822.

ПРИМЕНЕНИЕ ТЕХНОЛОГИИ NVIDIA CUDA ДЛЯ ОБУЧЕНИЯ И ДЕКОДИРОВАНИЯ СКРЫТЫХ МАРКОВСКИХ МОДЕЛЕЙ

Зацепин Павел Михайлович,

канд. физ.-мат. наук, доцент, заведующий кафедрой вычислительной техники и электроники Физико-технического факультета Алтайского Государственного университета, Россия, 656049, г. Барнаул, пр. Ленина, 61.
E-mail: zpm@phys.asu.ru

Гефке Денис Алексеевич,

аспирант кафедры вычислительной техники и электроники Физико-технического факультета Алтайского Государственного университета, Россия, 656049, г. Барнаул, пр. Ленина, 61. E-mail: bspugda@mail.ru

Актуальность работы обусловлена необходимостью оптимизации алгоритмов обработки больших речевых баз данных при разработке качественных систем автоматического распознавания речи. Развитие современных многоядерных процессоров, в частности графических процессоров GPU, позволяет получить существенный прирост производительности при реализации сложных ресурсоемких алгоритмов цифровой обработки сигналов и значительно сократить время обработки данных.

Цель работы: оптимизация алгоритмов обучения (Baum-Welch re-estimation) и декодирования (Витерби) Скрытых Марковских Моделей с помощью технологии параллельного программирования NVIDIA CUDA и оценка прироста производительности относительно центрального процессора.

Методы исследования: определение участков алгоритмов обучения и декодирования Скрытых Марковских Моделей, подходящих для эффективной параллельной реализации с учетом особенности программной модели CUDA, с последующей реализацией.

Результаты: Получена практическая параллельная реализация алгоритмов обучения и декодирования Скрытых Марковских Моделей с помощью графического процессора GPU. Произведена оценка прироста производительности относительно центрального процессора для различных параметров модели (количества состояний и размерности параметрического вектора). Результаты данной работы могут быть полезны как инженерам, работающим над созданием и улучшением систем автоматического распознавания речи, так и исследователям, работающим в области обработки сигналов и искусственного интеллекта.

Ключевые слова:

Распознавание речи, параллельные вычисления, Скрытые Марковские Модели, NVIDIA CUDA, алгоритм Витерби, алгоритм Баума-Велча.

Введение

Математический аппарат Скрытых Марковских Моделей (СММ) представляет собой универсальный инструмент моделирования стохастических процессов, для описания которых не существует точных математических моделей, а их свойства меняются с течением времени в соответствии с некоторыми статистическими законами. Наиболее широкое применение СММ нашли при решении таких задач, как распознавание речи, анализа последовательностей ДНК и ряда других [1].

Современные системы распознавания речи предполагают наличие нескольких сотен, а то и тысяч Скрытых Марковских Моделей и их сочетаний, вследствие чего работа со СММ связана со значительными вычислительными затратами, как на этапе обучения – при обработке огромного массива речевых данных, так и при последующем декодировании – в зависимости от сложности языковой модели. Например, обучение хорошей помехозащищенной дикторо-независимой системы распознавания слитной речи может занять несколько недель, а то и месяцев. Поэтому задача оптимизации алгоритмов обработки СММ остается актуальной в настоящее время [2, 3].

Применение современных технологий параллельного программирования, в частности графиче-

ских мультипроцессоров, позволяет получить значительный прирост производительности и перейти на качественно более высокий уровень в решении задач распознавания речи.

Целью данной работы является оптимизация алгоритмов обучения Скрытых Марковских Моделей (Baum-Welch re-estimation и forward-backward) и алгоритма декодирования (обобщенный алгоритм Витерби) с помощью графического процессора (CUDA) и оценка прироста производительности относительно центрального процессора (CPU).

Структура Скрытой Марковской Модели

В основе Скрытой Марковской Модели лежит конечный автомат, состоящий из N состояний. Переходы между состояниями в каждый дискретный момент времени t не являются детерминированными, а происходят в соответствии с некоторым вероятностным законом и описываются матрицей вероятностей переходов A_{NN} . Схематическое изображение диаграммы переходов между состояниями в СММ приведено на рис. 1 [4].

При каждом переходе в новое состояние i в момент времени t происходит генерация выходного значения x_t в соответствии с функцией распределения f_i . Результатом работы СММ является последо-

вательность параметрических векторов $\{x_1, x_2, \dots, x_T\}$ длиной T [5].

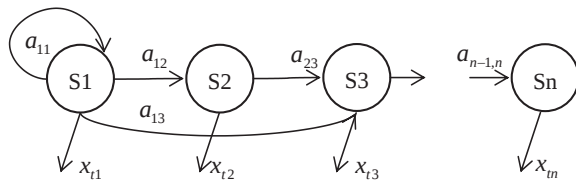


Рис. 1. Структурная схема переходов в СММ

На практике, как правило, решается обратная задача: при известной структуре Марковской Модели требуется определить, какова вероятность, что наблюдаемая последовательность $\{x_1, x_2, \dots, x_T\}$ может быть сгенерирована данной СММ [6].

Таким образом, основными параметрами СММ являются:

- 1) N – количество состояний;
- 2) матрица вероятностей переходов A_{NN} между состояниями;
- 3) N функций плотности вероятности $f_i(x)$.

Функция плотности вероятности $f_i(x)$ описывается, как правило, взвешенной Гауссовой смесью:

$$f(x) = \sum_{i=1}^M w_i p_i(x),$$

где M – количество компонент смеси; w_i – вес компонента смеси, а $p_i(x)$ – нормальное распределение вероятностей для D -мерного случая:

$$p_i(x) = \frac{1}{(2\pi)^{D/2} |\sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2(x - \mu_i)^T \sigma_i^{-1} (x - \mu_i)} \right\},$$

где μ_i – вектор математического ожидания; σ_i – матрица ковариации.

Работа со Скрытыми Марковскими Моделями, как и с любой другой адаптивной системой, осуществляется в 2 этапа: обучение – алгоритм Баума–Велча (Baum–Welch re-estimation), декодирование – алгоритм максимума правдоподобия (Витерби) [7].

В качестве параметрического вектора чаще всего используют мел-частотные кепстральные коэффициенты [8] (MFCC – mel-frequency cepstral coefficients), либо коэффициенты линейного предсказания [9] (LPC – linear prediction coefficients), а также их первые и вторые производные. Данные преобразования, в силу их гармонической природы, позволяют с высокой достоверностью локализовать вокализованные составляющие сигнала (гласные звуки) [10]. В тоже время, речевой сигнал имеет более сложную природу и помимо вокализованных звуков содержит ударные, шипящие и прочие составляющие. Поэтому для обработки речевого сигнала представляется перспективным применение других операторов, например Вейвлет-преобразований [11], рассмотрение которых выходит за рамки данной работы, т. к. реализованные алгоритмы являются универсальными и не зависят от природы параметрического вектора.

Сравнение Скрытой Марковской Модели и искусственных нейронных сетей при решении задачи автоматического распознавания речи

В настоящее время Скрытые Марковские Модели являются основой большинства успешных систем автоматического распознавания речи. Это связано с наличием ряда важных свойств, которыми обладают СММ по сравнению с основной альтернативной моделью классификации – искусственными нейронными сетями [12, 13]:

- 1) возможность моделирования длительных временных последовательностей (слов или целых высказываний), в то время как искусственные нейронные сети (ИНС) хорошо подходят для классификации кратковременных акустико-фонетических единиц, а их эффективность сильно снижается, когда на входе появляется некоторая динамика, т. е. образы подвержены, например, нелинейным изменениям во времени;
- 2) обучение ИНС при построении систем автоматического распознавания речи происходит, как правило, с учителем и требует точной временной разметки на фонемы обучающей выборки, что является чрезвычайно трудоемкой задачей для современных речевых баз данных. В тоже время обучение СММ может производиться без точной временной разметки;
- 3) возможность объединения, совместного обучения и декодирования набора СММ, представляющих отдельные фонемы, в соответствии с языковой моделью. Говоря простым языком, СММ хорошо адаптированы для построения составных моделей из исходных «примитивных» языковых элементов – фонем.

В то же время СММ обладают существенным недостатком по сравнению с ИНС – слабой дискриминантной мощностью, т. е. возможностью разделять классы образов. Особенно это проявляется при обучении с использованием критерия максимума правдоподобия. В связи с этим в последнее время исследования в области построения систем автоматического распознавания речи направлены на поиск гибридных моделей, объединяющих достоинства СММ и ИНС [14].

Рассмотрение гибридных моделей СММ/ИНС выходит за рамки данной статьи. Однако в их основе по-прежнему лежат СММ, поэтому результаты данной работы (оптимизация алгоритмов обучения и декодирования СММ с помощью GPU) могут быть успешно применены и при построении гибридных моделей.

Краткое описание алгоритма обучения Скрытой Марковской Модели

Процесс обучения Скрытой Марковской Модели заключается в определении на основе набора обучающих образцов следующих параметров:

- 1) матрицы вероятностей переходов между состояниями A_{NN} ;

2) параметров Гауссовых смесей (математическое ожидание, матрица ковариации и весового коэффициента) для каждого состояния.

Для решения этих задач совместно применяются два итерационных алгоритма: forward-backward и Baum-Welch re-estimation [2].

В алгоритме forward-backward вводятся две функции: прямого распространения вероятности $\alpha_j(t)$ и обратного $\beta_j(t)$.

Значение величины $\alpha_j(t)$ представляет собой вероятность наблюдения последовательности векторов $\{x_1, x_2, \dots, x_T\}$ и нахождения СММ в состоянии j в момент времени t :

$$\alpha_j(t) = P(x_1, x_2, \dots, x_t | state_t = j). \quad (1)$$

Величины $\alpha_j(t)$ и $\alpha_j(t-1)$ связаны итерационным выражением:

$$\alpha_j(t) = \left[\sum_{i=2}^{N-1} \alpha_i(t-1) A_{ij} \right] f_j(x_t). \quad (2)$$

Обратная функция $\beta_j(t)$ представляет собой вероятность нахождения СММ в состоянии j в момент времени t с последующим наблюдением последовательности $\{x_{t+1}, x_{t+2}, \dots, x_T\}$:

$$\beta_j(t) = P(x_{t+1}, x_{t+2}, \dots, x_T | state_t = j). \quad (3)$$

Величины $\beta_j(t)$ и $\beta_j(t+1)$ связаны аналогичным тождеством:

$$\beta_j(t) = \sum_{i=2}^{N-1} A_{ji} f_i(x_{t+1}) \beta_i(t+1). \quad (4)$$

Величины $\alpha_j(t)$ и $\beta_j(t)$ позволяют определить вероятность нахождения СММ в состоянии j в момент времени t при наблюдении последовательности $\{x_{t+1}, x_{t+2}, \dots, x_T\}$:

$$L_j(t) = \frac{1}{P} \alpha_j(t) \beta_j(t),$$

где $P = \alpha_N(t)$ – общая вероятность наблюдения последовательности $\{x_{t+1}, x_{t+2}, \dots, x_T\}$ данной СММ.

Алгоритм Баума-Велча (Baum-Welch re-estimation) на очередном шаге обучения позволяет, используя вышеприведенные выражения, сделать переоценку параметров модели [2].

Пусть имеется R обучающих образцов, тогда вероятность перехода из состояния i в состояние j определяется как:

$$A_{ij} = \frac{\sum_{r=1}^R \frac{1}{P_r} \sum_{t=1}^{T_r-1} \alpha_i^r(t) A_{ij} f_j^r(x_{t+1}^r) \beta_j^r(t+1)}{\sum_{r=1}^R \sum_{t=1}^{T_r} L_j^r(t)}. \quad (5)$$

Для каждого состояния j и для каждой компоненты Гауссовой смеси m математическое ожидание, матрица ковариации и вес определяются следующими выражениями

$$\mu_{jm} = \frac{\sum_{r=1}^R \sum_{t=1}^{T_r} L_{jm}^r(t) x_t^r}{\sum_{r=1}^R \sum_{t=1}^{T_r} L_{jm}^r(t)}, \quad (6)$$

$$\sigma_{jm} = \frac{\sum_{r=1}^R \sum_{t=1}^{T_r} L_{jm}^r(t) (x_t^r - \mu_{jm})(x_t^r - \mu_{jm})^T}{\sum_{r=1}^R \sum_{t=1}^{T_r} L_{jm}^r(t)}, \quad (7)$$

$$w_{jm} = \frac{\sum_{r=1}^R \sum_{t=1}^{T_r} L_{jm}^r(t)}{\sum_{r=1}^R \sum_{t=1}^{T_r} L_j^r(t)}. \quad (8)$$

Для качественного обучения Скрытой Марковской Модели требуется множество образцов сигнала: от нескольких десятков до нескольких сотен и тысяч экземпляров. Также необходимо соблюдать условие линейной независимости обучающих образцов, в противном случае в процессе обучения происходит вырождение матрицы ковариации, следствием чего является полная неработоспособность модели [2].

Алгоритм максимума правдоподобия (Витерби)

Суть алгоритма декодирования СММ заключается в поиске последовательности состояний, при прохождении которой наблюдаемая входная последовательность имела бы максимальную вероятность. Схема переходов и выбора цепочки состояний в момент времени t изображена на рис. 2.

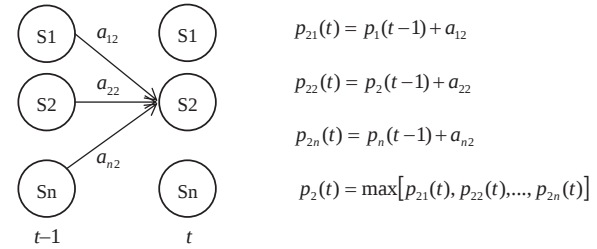


Рис. 2. Алгоритм декодирования Витерби

В момент времени t осуществляется переход в состояние i из *всех* предыдущих состояний, после чего выбирается последовательность, имеющая максимальную суммарную вероятность в моменты времени $t-1$ и t . Время выполнения алгоритма пропорционально длине последовательности L и квадрату количества состояний N . Алгоритм имеет сложность $O(N^2L)$.

Декодирование последовательности одной моделью не является ресурсоемкой задачей. Однако при наличии сложной языковой модели СММ объединяются вместе в виде графа, в зависимости от последовательностей следования фонем, и процесс декодирования осуществляется одновременно для множества моделей. В связи с чем алгоритмическая сложность задачи резко возрастает [3].

Особенности параллельного программирования на CUDA

Применение современных графических процессоров позволяет получить многократный прирост производительности при решении ряда научных и технических задач. Однако для достижения наилучших результатов необходимо учитывать ряд особенностей [15–18]:

- 1) графический процессор (GPU) состоит из нескольких мультипроцессоров, которые, в свою очередь, состоят из ядер. Каждое ядро одновременно выполняет 32 потока (warp). Например, NVidia GeForce GTX 480 состоит из $15 \times 32 = 480$ ядер и параллельно может выполнять до 15360 легковесных потоков. Потоки объединяются в блоки и сетки блоков. Каждый поток имеет идентифицирующие его координаты;
- 2) максимальной производительности удается достичь при выполнении *однотипных* действий над большим числом *обрабатываемых единиц* данных;
- 3) архитектура памяти имеет сложную организацию: глобальная память (объемная, но медленная), локальная память, разделяемая память (быстрая), память констант и т. д.;
- 4) особенности доступа к памяти: для получения максимальной пропускной способности все запросы к памяти должны быть выровнены.

Подробное описание и особенности применения графических процессоров можно найти в соответствующих руководствах [17].

Реализация декодера СММ на CUDA

Пусть $N=(8,16,32)$ – количество состояний модели; $L=50$ – длина последовательности; $C=32$ – число одновременно декодируемых последовательностей одной СММ; N_{NN} – матрица вероятностей переходов между состояниями; B_{LN} – матрица вероятности наблюдения вектора последовательности X_L в состоянии N (C матриц рассчитываются заранее).

Для декодирования одной последовательности используются N параллельных потоков, каждый из которых декодирует свое состояние $1...N$. Данная схема изображена на рис. 3. Итоговая вероятность выбирается как максимум вероятностей, полученных каждым потоком. Сложность алгоритма для одного потока – $O(N \times L)$ [19].

Один вычислительный блок CUDA одновременно декодирует C последовательностей (для одной СММ). Таким образом, для декодирования СММ из 32 состояний задействовано $N \times C = 1024$ потока, что соответствует максимальному количеству потоков на один мультипроцессор для CUDA версии 2.0.

Каждый мультипроцессор используется для декодирования отдельной СММ. В итоге на GPU параллельно происходит декодирование $32 \times 15 = 480$ последовательностей для видеопроцессора NVidia GeForce GTX 480.

Разделяемая (быстрая) память используется для хранения матрицы вероятностей A и промежуточных вероятностей для N состояний. Объем необходимой памяти [19] – $O(N^2 + N \times C)$.

Реализация алгоритма обучения СММ с помощью CUDA

Пусть R – количество обучающих образцов для одной СММ (32); N – количество состояний модели (8, 16, 32); M – количество компонент Гауссовой смеси (16); D – размерность параметрического вектора (16, 32).

Предварительный анализ показал, что в процедуре forward-backward основную долю времени (порядка 95 %), при вычислении вероятностей $\alpha_i(t)$ и $\beta_i(t)$, занимает расчет вероятности наблюдения вектора x_i в состоянии j для модели $r - f_j(x_i)$. В тоже время эта часть вычислений является *менее* сложной с алгоритмической точки зрения [2].

Размерность параметрического вектора D выбиралась равной warp или половине warp (количеству потоков, выполняемых ядром CUDA одновременно). Это же количество потоков (D) использовалось для реализации матричных операций при вычислении функции плотности вероятности

$$p_i(x) = \frac{1}{(2\pi)^{D/2} |\sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2(x - \mu_i)^T \sigma_i^{-1} (x - \mu_i)} \right\},$$

где вектора x , μ имеют размерность D , и матрица σ_i^{-1} имеет размерность $D \times D$.

Для видеопроцессора GeForce GTX 480 каждый мультипроцессор содержит 32 ядра CUDA. Таким образом, на одном мультипроцессоре производится одновременный расчет значений функции вероятности для 32 последовательностей в одном из состояний. Вычисления производятся по всей длине последовательности – T . В свое время, каждый мультипроцессор обрабатывает одно из N состояний. Итоговая схема расчета приведена на рис. 4.

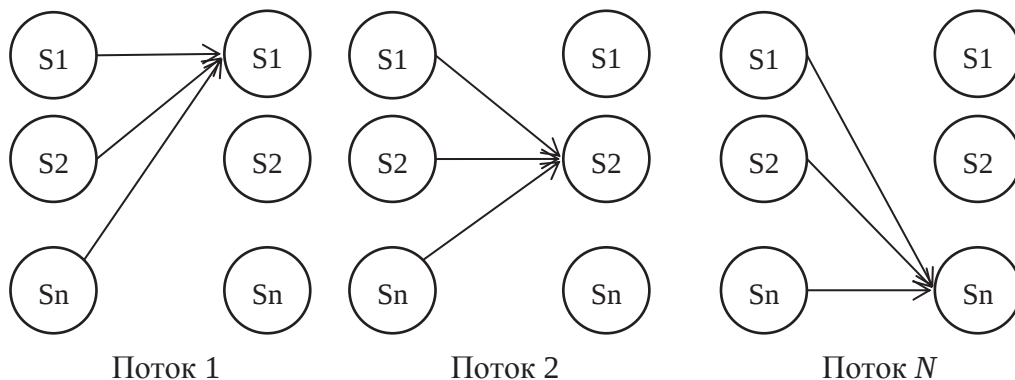


Рис. 3. Параллельная схема декодирования



Рис. 4. Параллельная схема вычисления значений вероятности модели во всех состояниях для всех обучающих образцов

В результате работы алгоритма получается массив вероятностей $f(x_n)$, который затем используется для вычисления окончательных $\alpha(t)$ и $\beta(t)$ на CPU в соответствии с итеративными формулами (1–4). В процессе вычисления одновременно задействовано $32 \times 32 \times 16 = 16384$ потока.

В алгоритме Баума–Велча (формулы 5–8) схема распределения вычислений по ядрам аналогична. Векторные и матричные операции для каждого обучающего образца используют по 32 потока (размерность параметрического вектора). При этом каждый мультипроцессор обрабатывает одновременно 32 обучающих образца. После чего осуществляется редуцирование (свертка) для окончательного расчета величин математического ожидания μ_{jm} и матрицы ковариации σ_{jm} . Каждый мультипроцессор независимо обрабатывает одно из N -состояний модели.

Расчет вероятностей переходов между состояниями A_{ji} и весов компонент Гауссовых смесей w_{im} осуществляется с помощью CPU, т. к. все входящие в формулы (5) и (8) величины были предварительно рассчитаны, и для получения окончательных значений не требуются существенные вычислительные затраты [20].

Тестирование

Для тестирования алгоритма использовалась вычислительная машина следующей конфигурации:

1. Центральный процессор – Intel Core i7 930 2,8 ГГц.
2. Операционная система – Windows 7 64-Bit.
3. Оперативная память – Kingston 12 Gb RAM DDR3.
4. Видеокарта – Nvidia GeForce GTX 580 (1536 Mb, 512 ядер CUDA).
5. Версия CUDA – 4.0.
6. Компилятор Microsoft Visual Studio 2008 (векторизация SSE2).
7. Тип данных – double (64 разряда).

Результаты тестирования алгоритмов обучения [19] и декодирования [20] приведены в табл. 1, 2 соответственно.

Таблица 1. Результаты тестирования алгоритмов forward-backward и Baum-Welch re-estimation

N	D	CPU FB, с.	GPU FB, с.	CPU/GPU FB	CPU BW, с.	GPU BW, с.	CPU/GPU BW
8	16	8,688	1,494	5,815	21,926	5,118	4,284
16	16	17,286	1,965	8,797	44,753	5,890	7,598
32	16	33,881	3,778	8,968	95,077	11,574	8,215
8	32	27,578	2,594	10,631	72,714	8,276	8,786
16	32	55,241	3,153	17,520	146,278	9,537	15,338
32	32	109,676	6,006	18,261	302,857	18,473	16,395

Таблица 2. Результаты тестирования алгоритма декодирования CMM

Число состоя- ния (N)	Про- цес- сор	Коли- че- ство CMM	Количество декоди- руемых последова- тельств	Время, мс	CPU/GPU
8	CPU	1	32	1,54	50,7
		64		1,22	
1		5,70		41,9	
64		3,73			
32		1		2,120	58,2
		64		1,930	

Примечание к табл. 1, 2: FB – Forward-Backward. BW – Baum-Welch. CPU/GPU – отношение времени выполнения CPU к GPU. Время пересылки в память GPU включено в общее время. Наибольший прирост производительности достигается при размере параметрического вектора $D=32$ и $N=32$ состояниях модели, т. к. в этом случае задействованы все ресурсы GPU.

Выводы

Применение современных графических процессоров для реализации математического аппарата Скрытых Марковских Моделей позволяет получить существенный прирост производительности относительно центрального процессора, как на этапе обучения, так и при декодировании CMM, особенно при построении моделей с большим количеством состояний.

Однако с практической точки зрения целесообразно оптимизировать лишь те участки алгоритма, на которых выполняются ресурсоемкие *матричные и векторные операции*, в частности вычисление значений функции Гауссовой смеси для обрабатываемых последовательностей и расчет матрицы ковариации. Итеративная часть алгоритма for-

ward-backward (формулы 2 и 4) занимает малую часть от общего времени вычисления и сложна в реализации на GPU, поэтому ее оптимизация не представляется целесообразной. Такой же вывод можно сделать относительно расчета матрицы вероятностей переходов между состояниями и весов компонент Гауссовой смеси (формулы 5 и 8).

Повышение итоговой производительности в несколько десятков раз позволит применять более сложные Марковские Модели, а также увеличить объем материала, используемого при обучении, что представляется перспективным с точки зрения реализации более качественных систем распознавания речи.

СПИСОК ЛИТЕРАТУРЫ

1. Jurafsky D., Martin J.H. Speech and Language processing. 2nd ed. – Englewood Cliffs, New Jersey: Prentice Hall Inc., 2008. – 302 p.
2. Hidden Markov Model Toolkit Book. – Cambridge: Cambridge University Engineering Department, 2001–2009. – 399 p.
3. Rabiner L.R. A tutorial on Hidden Markov Models and selected applications in speech recognition // Proceedings of The IEEE. – 1989. – V. 77. – № 2. – P. 257–286.
4. Rabiner L.R., Juang B.H. Fundamentals of Speech Recognition. – Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1993. – 553 p.
5. Rabiner L.R., Levinson S.E. Isolated and Connected word Recognition – Theory and Selected Applications // IEEE Transactions on communications. – 1981. – V. 29. – № 5. – P. 301–315.
6. Vaseghi S.V. Advanced digital signal processing and noise reduction. 3rd ed. – Chichester, England: John Wiley & Sons, 2006. – 453 p.
7. Huang Xuedong. Spoken language processing: a guide to theory, algorithm, and system development. – Englewood Cliffs, New Jersey: Prentice Hall Inc., 2001. – 480 p.
8. Айфичер Э.С., Джервис Б.У. Цифровая обработка сигналов: практический подход, 2-е изд. / пер. с англ. – М.: Изд. дом «Вильямс», 2004. – 992 с.
9. Vaidyanathan P.P. The Theory of Linear Prediction. – California: Morgan & Claypool, 2008. – 198 p.
10. Сорокин В.Н., Цыплихин А.И. Сегментация и распознавание гласных // Информационные процессы. – 2004. – Т. 4. – № 2. – С. 202–220.
11. Гейфе Д.А., Зацепин П.М. Применение нейронных сетей для классификации сигналов звукового диапазона // Нейроинформатика, ее приложения и анализ данных: XVII Всеросс. семинар. – Красноярск, 2009. – С. 37–40.
12. Хайкин С. Нейронные сети: полный курс, 2-е изд. / пер. с англ. – М.: Изд. дом «Вильямс», 2006. – 1104 с.
13. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы / пер. с польск. И.Д. Рудинского. – М.: Телеком, 2006. – 452 с.
14. Маковкин К.А. Гибридные модели – Скрытые марковские модели/Многослойный персептрон – и их применение в системах распознавания речи // Речевые технологии. – 2012. – № 3. – С. 58–83.
15. Сандерс Дж., Кэндрот Э. Технология CUDA в примерах. Введение в программирование графических процессоров. – М.: ДМК Пресс, 2011. – 256 с.
16. Боресков А.В., Харламов А.А. Основы работы с технологией CUDA. – М.: ДМК Пресс, 2011. – 232 с.
17. NVIDIA CUDA SDK 4.0 // NVIDIA Corporation. URL: http://www.nvidia.com/object/cuda_sdks.html (дата обращения: 10.07.2013).
18. Гейфе Д.А. Применение технологии CUDA для частотного разделение каналов широкополосного тракта // Многоядерные процессоры и параллельное программирование: Регион. научно-практ. конф. – Барнаул, 2011. – С. 12–15.
19. Гейфе Д.А., Зацепин П.М. Применение технологии CUDA для декодирования Скрытых Марковских Моделей // Многоядерные процессоры, параллельное программирование, ПЛИС, системы обработки сигналов: сборник статей II регион. научно-практ. конф. – Барнаул, 2012. – С. 45–51.
20. Гейфе Д.А., Зацепин П.М. Применение технологии CUDA для обучения Скрытых Марковских Моделей // Многоядерные процессоры, параллельное программирование, ПЛИС, системы обработки сигналов: сборник статей III Всеросс. научно-практ. конф. – Барнаул, 2013. – С. 30–39.

Поступила 10.07.2013 г.

NVIDIA CUDA APPLICATION TO TRAIN AND DECODE THE HIDDEN MARKOV MODELS

Pavel M. Zatsepin,

Cand. Sc., Altai State University, 61, Lenin Avenue, 656049, Barnaul, Russia.

E-mail: zpm@phys.asu.ru

Denis A. Gefke,

Altai State University, 61, Lenin Avenue, 656049, Barnaul, Russia.

E-mail: bspugda@mail.ru

The urgency of the discussed issue is caused by the need of optimization of huge speech corpus's processing algorithms required for developing robust automatic speech recognition systems. The evolution of modern multicore processors, specifically graphical processor units GPU, allows improving sufficiently the performance of difficult and resource-intensive digital signal processing algorithms and reducing sufficiently a data processing time.

The main aim of the study is to optimize education (Baum–Welch re-estimation) and decoding (Viterbi) algorithms of Hidden Markov Models by parallel programming technology NVIDIA CUDA and to estimate performance increase in comparison within the CPU.

The methods used in the study: the search of education and decoding algorithm's parts suitable for effective parallel realization by NVIDIA CUDA and its implementation.

The results: The authors have developed parallel realization of education and decoding Hidden Markov Models algorithms by GPU and have estimated the performance increase in comparison within the CPU for different model's parameters (the number of model state and dimension of a feature vector). The results of the paper can be used both by engineers developing and improving the automatic speech recognition systems and by explorers working on a digital signal processing and artificial intelligence systems.

Key words:

Speech recognition, parallel computing, Hidden Markov Models, NVIDIA CUDA, Viterbi algorithm, Baum–Welch re-estimation algorithm.

REFERENCES

1. Jurafsky D., Martin J.H. *Speech and Language processing*. 2nd ed. Englewood Cliffs, New Jersey: Prentice Hall Inc., 2008. 302 p.
2. *Hidden Markov Model Toolkit Book*. Cambridge, Cambridge University Engineering Department, 2001–2009. 399 p.
3. Rabiner L.R. A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE*. 1989, vol. 77, no. 2, pp. 257–286.
4. Rabiner L.R., Juang B.H. *Fundamentals of Speech Recognition*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1993. 553 p.
5. Rabiner L.R., Levinson S.E. Isolated and Connected word Recognition – Theory and Selected Applications. *IEEE Transactions on communications*, 1981. vol. 29, no. 5, pp. 301–315.
6. Vaseghi S.V. *Advanced digital signal processing and noise reduction*. 3rd ed. Chichester, England, John Wiley & Sons, 2006. 453 p.
7. Huang Xuedong. *Spoken language processing: a guide to theory, algorithm, and system development*. Englewood Cliffs, New Jersey, Prentice Hall Inc., 2001. 480 p.
8. Iteachor E.C., Jervis B.W. *Digital Signal Processing. A practical approach*. 2nd ed. Englewood Cliffs, New Jersey, Prentice Hall Inc., 2004. 992 p.
9. Vaidyanathan P.P. *The Theory of Linear Prediction*. California, Morgan & Claypool, 2008. 198 p.
10. Sorokin V.N., Tsyplikhin A.I. Segmentatsiya i raspoznavanie glasnykh [Segmentation and vocal recognition]. *Informatsionnye protsessy – Informational processes*, 2004, vol. 4, no. 2, pp. 202–220.
11. Gefke D.A., Zatssepin P.M. Primenenie neyronnykh setey dlya klassifikatsii signalov zvukovogo diapazona [The application of neural networks for voice signal classification]. *Neuroinformatika, ee primeneniye i analiz dannykh. XVII Vserossiyskiy seminar* [Neuroinformatics, its application and data analysis. XVII Russian conference]. Krasnoyarsk, 2009. pp. 37–40.
12. Haykin S. *Neural Networks – A Comprehensive Foundation*. 2nd ed. New Jersey, Pearson Education Inc., 1999. 823 p.
13. Rutkovskaya D., Pilinskiy M., Rutkovskiy L. *Neironnye sistemy, geneticheskie algoritmy i nechetkie sistemy* [Neural networks, genetic algorithms and fuzzy systems]. Moscow, Telekom, 2006. 452 p.
14. Makovkin K.A. Gibridnye modeli – Skrytye Markovskie Modely/ Mnogosloynny pertseptron – i ikh primeneniye v sistemakh avtomaticheskogo raspoznavaniya rechi [The Hybrid models – Hidden Markov Models/Multi-Layer Perceptron – and its application in automatic speech recognition systems]. *Rechevye tekhnologii – Speech Technologies*, 2012, no. 3, pp. 58–83.
15. Sanders J., Kandrot E. *CUDA by Example: An Introduction to General-Purpose GPU Programming Code*. Boston, Addison-Wesley Professional, 2010. 312 p.
16. Borekov A.V., Harlamov A.A. *Osnovy raboty s CUDA* [Basic usage of CUDA]. Moscow, DMK Publ., 2011. 232 p.
17. NVIDIA CUDA SDK 4.0. *NVIDIA Corporation*. Available at: http://www.nvidia.com/object/cuda_sdks.html (accessed 10 July 2013).
18. Gefke D.A., Zatssepin P.M. Primeneniye tekhnology CUDA dlya dekodirovaniya Skrytykh Markovskikh Modeley [The application of CUDA technology for decoding Hidden Markov Models]. *Mnogoyadernye protsessory, parallelnoe programmirovaniye, PLIS, sistemy obrabotki signalov. Sbornik statey III Regionalnoy konferentsii* [Multicore processors, parallel programming, PLD, digital signal processing systems. Paper collection of the III Regional science-practical conference]. Barnaul, 2012. pp. 45–51.
19. Gefke D.A., Zatssepin P.M. Primeneniye tekhnology CUDA dlya obucheniya Skrytykh Markovskikh Modeley [The application of CUDA technology for education of Hidden Markov Models]. *Mnogoyadernye protsessory, parallelnoe programmirovaniye, PLIS, sistemy obrabotki signalov. Sbornik statey III Vserossiyskoy konferentsii* [Multicore processors, parallel programming, PLD, digital signal processing systems. Paper collection of the III Russian science-practical conference]. Barnaul, 2013. pp. 30–39.
20. Gefke D.A. Primeneniye tekhnologii CUDA dlya chastotnogo razdeleniya kanalov shirokopolosnogo trakta [The application of CUDA technology for frequency-separation of wideband signal]. *Mnogoyadernye protsessory i parallelnoe programmirovaniye. Regionalnaya nauchno-prakticheskaya konferentsiya* [Multicore processors and parallel programming. Regional science-practical conference]. Barnaul, 2011. pp. 12–15.

УДК 004.22

ПРИМЕНЕНИЕ СРЕДСТВ OPEN SOURCE ДЛЯ ВИЗУАЛИЗАЦИИ СТАНДАРТИЗИРОВАННЫХ МЕДИЦИНСКИХ ДАННЫХ

Копаница Георгий Дмитриевич,

канд. техн. наук, доцент кафедры оптимизации систем управления
Института кибернетики Томского политехнического университета, Россия,
634050, г. Томск, пр. Ленина, д. 30; доцент Томского Государственного
архитектурно-строительного университета, Россия, 634003, г. Томск,
пл. Соляная, д. 2. E-mail: georgy.kopanitsa@gmail.com

Тараник Максим Алексеевич,

аспирант кафедры оптимизации систем управления Института кибернетики
Томского политехнического университета,
Россия, 634050, г. Томск, пр. Ленина, д. 30. E-mail: taranik@tpu.ru

Актуальность работы обусловлена потребностью использования гибких и тривиальных программных средств для визуализации стандартизированных медицинских данных.

Цель работы: *Используя средства open source реализовать представление стандартизированных медицинских данных, структура которых регламентирована стандартом ISO 13606.*

Методы: *Для структурирования медицинских данных был использован стандарт клинической информации ISO 13606, который описывает структуру данных для передачи их между различными медицинскими информационными системами. Для преобразования данных в стандартный формат был применен канонический подход, при котором данные из различных систем преобразовывались к архетипной модели ISO 13606. JavaScript библиотеки с открытым программным кодом (open source). Среда разработки MS Visual Studio.*

Результаты: *Для работы с имеющимся набором медицинских архетипов была разработана информационная модель, описанная на языке XML. Используя open source библиотеки JavaScript, были построены диаграмма показателей инсулина и уровня сахара в крови пациента, а также динамическая диаграмма, определяющая уровень сахара в крови пациента.*

Выводы: *Применяя средства open source (библиотек JavaScript), разработчик получает многообразие готовых решений в виде шаблонов, используя которые реализуются гибкие и тривиальные способы графического представления медицинских данных.*

Ключевые слова:

Open source, визуализация, архетипы, JavaScript, ISO 13606.

Введение

Одним из наиболее важных аспектов при работе с персональными медицинскими записями (ПМЗ) является пользовательский интерфейс. Разработка простых и удобных средств отображения клинического статуса пациента особенно актуальна при наблюдении за пациентами в престарелом возрасте. Такие средства реализуются в рамках подхода ALL (Ambient Assisted Living). В основе данного подхода лежит концепция, предполагающая разработку интеллектуальных систем для обработки данных, полученных при помощи различных сенсорных устройств, объединенных в сеть и расположенных по всему периметру дома наблюдаемого пациента [1–5]. Также подход AAL направлен на разработку специальных средств, таких как социальные сети (social networking) [6] и веб-сервисы, позволяющие пациенту удаленно взаимодействовать с медицинским персоналом [7]. Для формирования полного и развернутого представления о клиническом статусе пациента необходимы данные из различных источников. Таким образом, данный подход требует стандартизации хранимых данных.

Стандарты хранения клинических данных не дают ответа на вопрос о представлении медицин-

ской информации пользователю [8–12]. Эта проблема стала следствием ранее проведенных исследований, направленных на разработку методов визуализации стандартизированных медицинских данных [13]. Одним из таких стандартов является стандарт ISO 13606, позволяющий описывать медицинские документы любой сложности. Однако сами медицинские концепции, описываемые архетипом, являются семантически несвязными. Таким образом, если пользователь хочет определить референтный интервал для поля с физической величиной (Physical Quantity, PQ), это будет обработано системой только в том случае, если она заранее знает, что референтный интервал должен быть связан с определенным полем архетипа. Решением данной проблемы может быть разработка модели верхнего уровня, которая будет дополнять собой модель данных стандарта ISO 13606 [14].

При использовании стандартизированных медицинских данных разработка графического пользовательского интерфейса требует дополнительных существенных временных и финансовых затрат для его реализации. Таким образом, пользовательский интерфейс разрабатывается под конкретные EHR (Electronic Healthcare Records) системы [15]. Одним из способов представления стандартизированной

ванных медицинских данных является применение средств open source, использующих библиотеки JavaScript [16, 17]. В качестве хранилища информации для работы с данными библиотеками обычно выступают файлы с расширением JBOS или csv. Однако стандартизированные медицинские данные представлены в формате XML [8, 18–20].

Целью настоящей статьи является исследование средств open source на предмет их использования для визуализации стандартизированных медицинских данных. В соответствие с целью был обозначен следующий ряд задач:

- Определить существующие средства open source для решения задачи представления медицинских данных.
- Разработать инструментарий для визуализации клинических архетипов, основанных на стандарте ISO 13606.
- Осуществить проверку разработанного инструмента визуализации стандартизированных медицинских данных на клинических результатах наблюдения за пациентом с сахарным диабетом.

Методы и материалы

Для проведения исследования был определен набор медицинских данных пациента с сахарным диабетом, полученных из [21–23]. Данный набор был сформирован медицинскими работниками посредством автоматизированного фиксирования соответствующих клинических показателей в различные периоды времени наблюдения за пациентом. Для настоящего исследования полученный набор данных был представлен в виде совокупности архетипов, структура которых определена стандартом ISO 13606 [18]. Данная структура представлена в виде XML файла, хранящего данные из разнородных источников. Также такой формат необходим для сохранения семантики при

передаче данных между различными EHR системами. Для преобразования данных в требуемый формат используется канонический подход, предусматривающий преобразование данных из модели информационной системы в каноническую модель для обмена клинической информацией. Последняя модель является архетипной моделью стандарта ISO 13606 [10]. Для определения характеристик визуализации медицинских данных была разработана информационная модель (ИМ) [22], представляющая собой XML схему. Используемые в проекте библиотеки JQuery [16, 17] применялись для построения статических и динамических диаграмм с целью определения корреляции между уровнем сахара в крови и инсулином в различные периоды наблюдения: ежедневный, еженедельный и ежемесячный. В качестве среды разработки интерфейса использовалась MS Visual Studio, полученные результаты визуализации были отображены с применением браузера Mozilla Firefox.

Результаты

Используемая совокупность медицинских данных была определена как соответствующая совокупность архетипов, которая позволяет совершать обмен данной медицинской информацией между пользователями. Определенные архетипы представляют собой совокупность XML файлов. Настоящие медицинские данные, необходимые для представления пользователю в графическом виде, были получены из различных EHR систем. Различные данные, представленные различными архетипами, были объединены в группы. Характеристики визуализации медицинских данных, описанные в ИМ, определяют тип диаграмм, а также другие свойства конечного графического представления пользователю. Применяя соответствующий шаблон для отображения данных, результатом визуализации становится сгенерированная HTML стра-

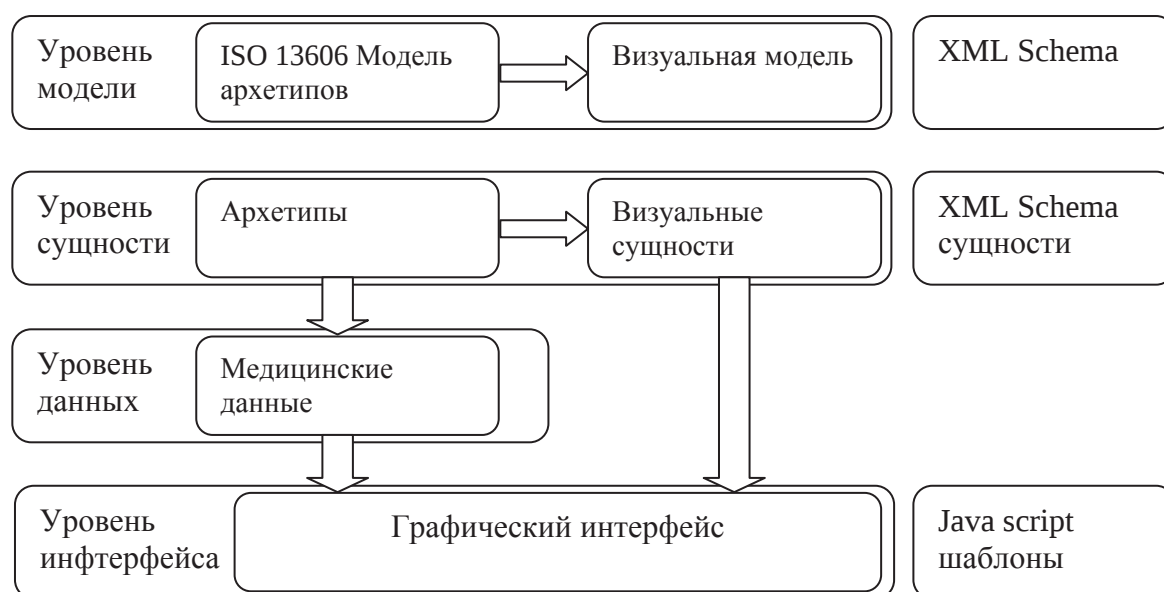


Рис. 1. Процесс визуализации медицинских данных

ница с диаграммой, за построение которой отвечают библиотеки JavaScript. На рис. 1 представлен процесс модельного преобразования, обеспечивающий графическое отображение медицинских данных пользователю.

Визуальная модель (Visual model) основана на архетипной модели стандарта ISO 13606. Уровень сущности (Instance layer) содержит архетипы, а также информационную модель представления данных (Visual Medical Concept). На уровне данных (Data layer) происходит связь файлов ИМ (Visual Medical Concept) с файлами XML, содержащими медицинские данные. На следующем уровне представления медицинских данных применяются JavaScript шаблоны для графического отображения диаграмм. Результаты настоящей работы представлены на рис. 2–5. Данные диаграммы отображают динамику изменения различных зафиксированных клинических параметров пациента в различное время. Диаграмма на рис. 2 демонстрирует совокупность различных параметров, опреде-

ленных различными архетипами, но содержащихся в одном медицинском документе. Данные архетипы описывают различные показания инсулина пациента: регулярный инсулин (regular insulin), инсулин НПХ (NPH insulin), а также содержания сахара в крови (Glucose).

На шаблоне, изображенном на рис. 3, также представлены данные, характеризующие динамику изменения уровня сахара в крови пациента, однако он позволяет масштабировать необходимые части диаграммы для более детального просмотра.

Гибкость применяемых шаблонов позволяет отображать необходимые медицинские данные в любой форме. Таким образом, модифицировав один из представленных выше шаблонов, были получены графики, изображенные на рис. 4, 5. На рис. 4 показана динамика изменения уровня сахара в крови пациента (Glucose) а также пульса (Pulse). Более того, на данном графике отображен нормальный интервал первого показателя (Glucose norm).

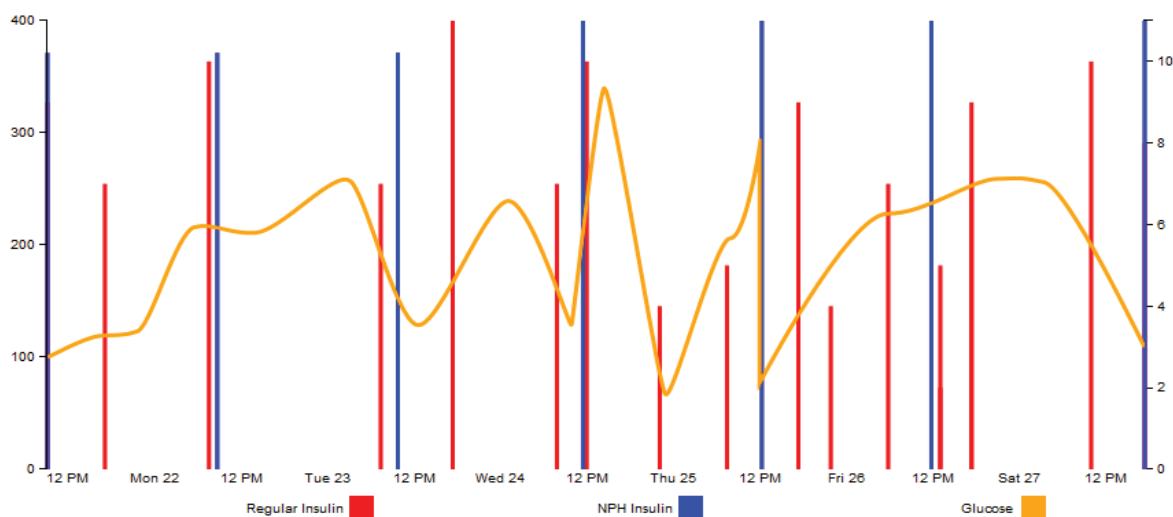


Рис. 2. Диаграмма показателей инсулина и уровня сахара в крови пациента

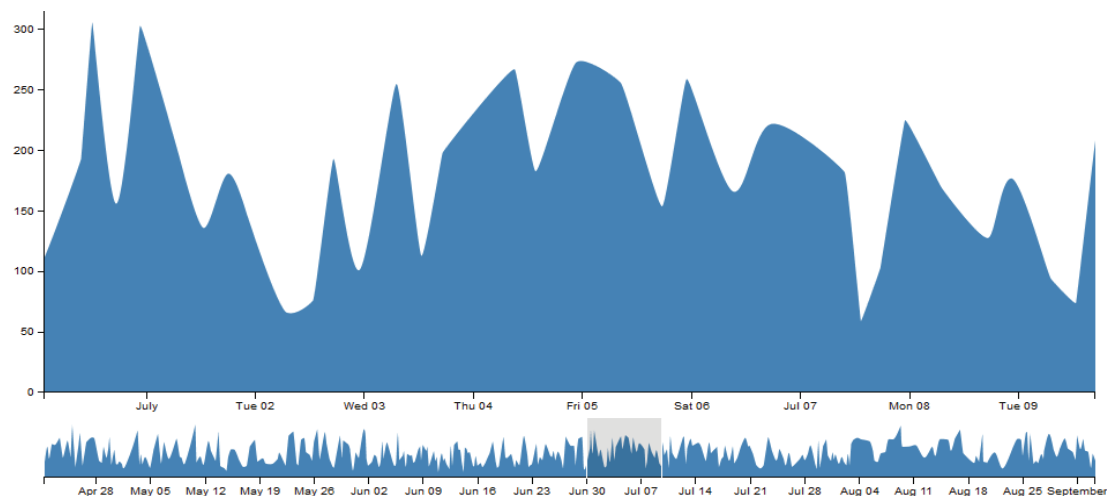


Рис. 3. Динамическая диаграмма уровня сахара в крови пациента

На заключительной диаграмме (рис. 5) представлены такие показатели, как уровень сахара в крови (Glucose) с нормальным интервалом, артериальное давление (Pressure) и пульс (Pulse). Крайняя точка линии артериального давления, наиболее близкая к оси абсцисс, характеризует нижнюю планку показателя. Планка верхней границы давления отображена точкой максимально отдаленной от оси абсцисс.

Практическая реализация представленных выше клинических диаграмм была выполнена с применением открытой графической библиотеки D3JS [16]. Одним из наиболее важных достоинств использования данной библиотеки является ассоциативный набор тегов, отвечающих за соответствующие блоки представленной диаграммы. Так, например, диаграмма, изображенная на рис. 1, имеет такие блоки, как графическая область, блок

графиков, шкала глюкозы, шкала инсулина, шкала дат и легенда (рис. 6).

Основной элемент – графическая область – является контейнером и содержит весь набор остальных блоков, представленных как графические элементы, описанные на языке масштабируемой векторной графики (SVG). Данный элемент прикрепляется в body секции HTML документа с определенными параметрами (рис. 7).

Представленные на рис. 6 шкалы определяют значения для соответствующих показателей глюкозы и инсулина, представленных осью ординат, а также времени, значения которого расположены на оси абсцисс. Для добавления осей на диаграмму была использована следующая кодовая конструкция (рис. 8, 9):

Для построения соответствующих графиков на диаграмме применяются три различных метода

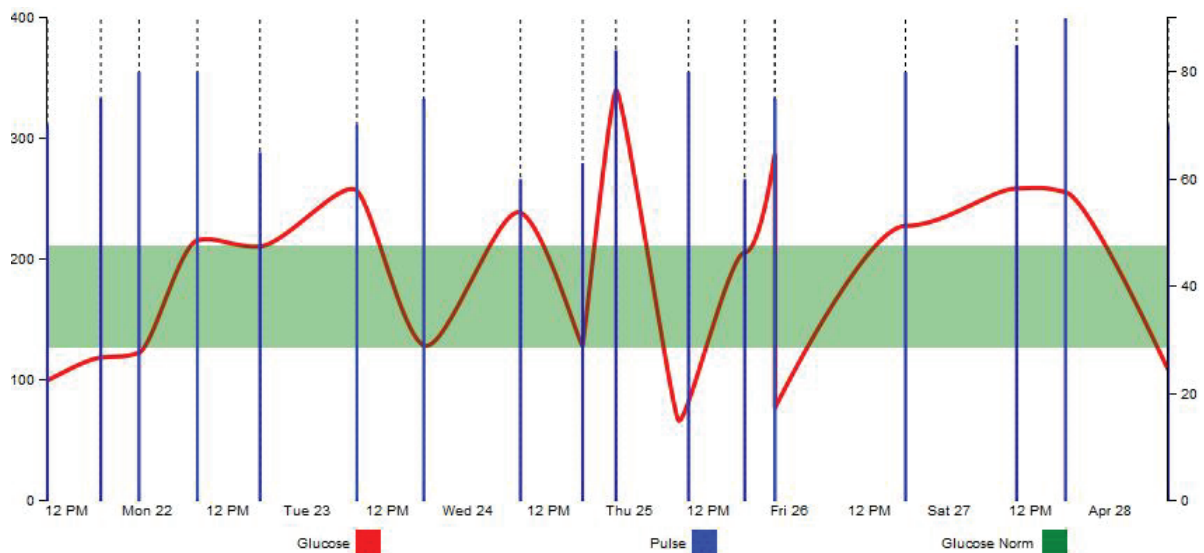


Рис. 4. Диаграмма показателей пульса и уровня сахара в крови пациента

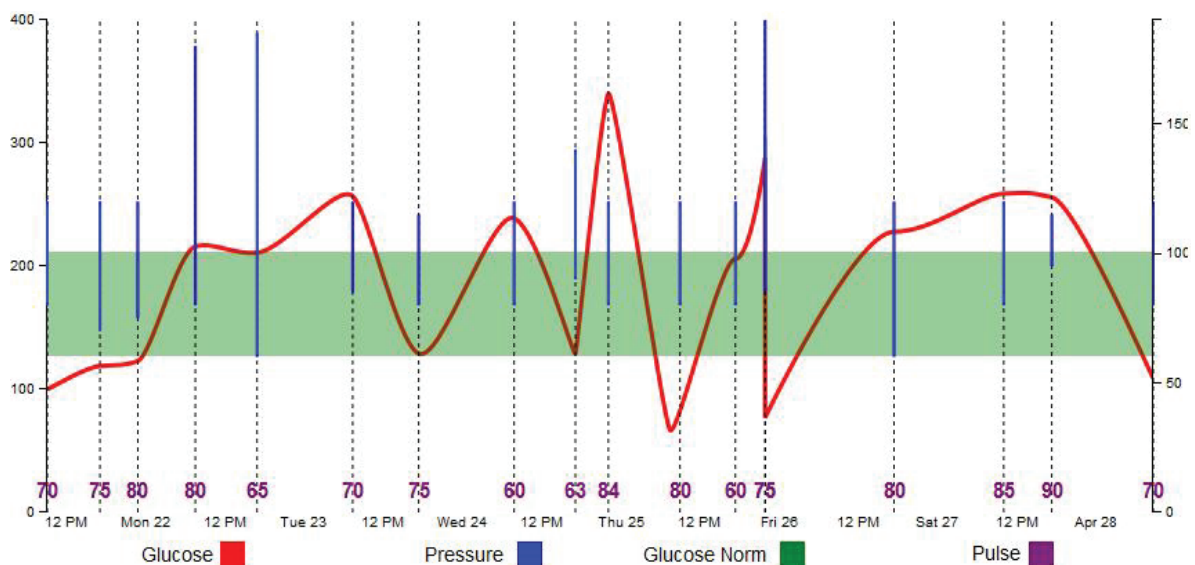


Рис. 5. Диаграмма показателей пульса, артериального давления и уровня сахара в крови пациента

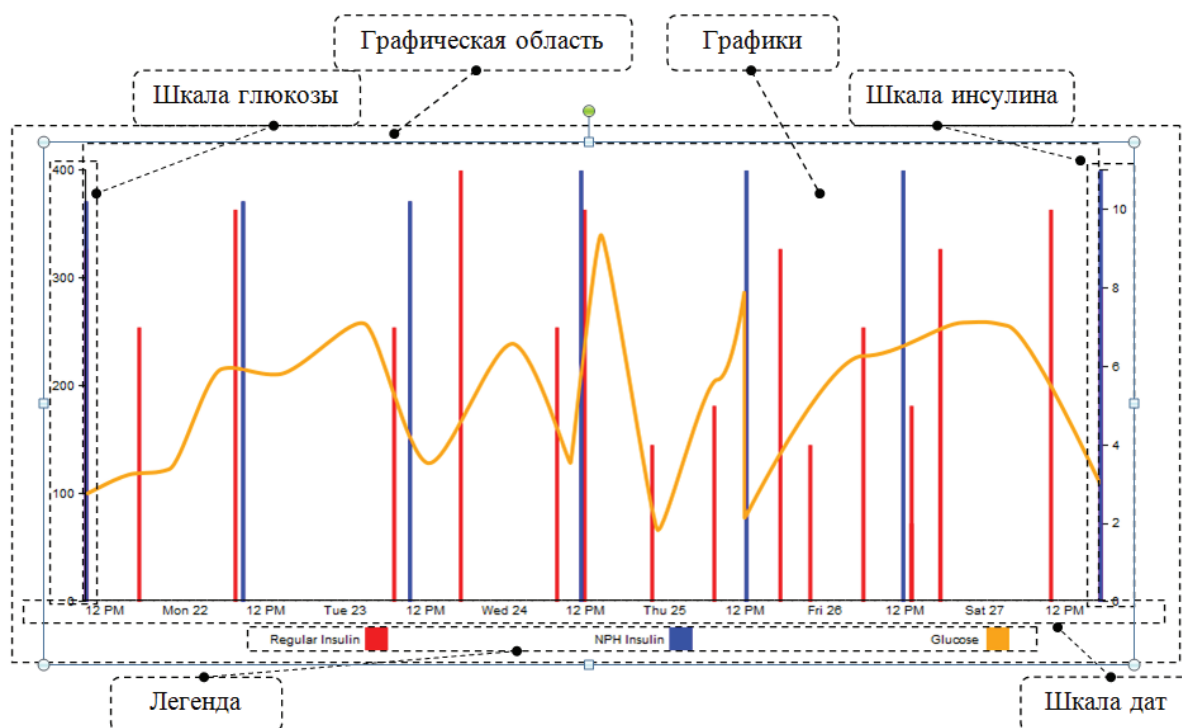


Рис. 6. Блоки библиотеки D3JS

```
var svg = d3.select("body").append("svg:svg")
    .attr("width", w + m[1] + m[3])
    .attr("height", h + m[0] + m[2])
    .append("svg:g")
    .attr("transform", "translate(" + m[3] + "," + m[0] + ")");
```

Рис. 7. Код элемента «графическая область»

```
svg.append("svg:g")
    .attr("class", "y axis")
    .attr("transform", "translate(" + w + ",0)")
    .call(yAxis);
```

Рис. 8. Код добавления оси ординат

```
svg.append("svg:g")
    .attr("class", "x axis")
    .attr("transform", "translate(0," + h + ")")
    .call(xAxis);
```

Рис. 9. Код добавления оси абсцисс

(рис. 10, 11), обрабатывающих хранящиеся данные. Для построения динамики изменения инсулина применяется гистограмма (bar chart), а для отображения динамики показателя глюкозы применяется кривая линия (curve line).

Область легенды, представленная на рис. 1, также является SVG элементом, состоящим из элементов прямоугольника и текстового элемента. На рис. 12 представлен код добавления легенды для обозначения представленных гистограмм инсулина и графика изменения глюкозы пациента.

Обсуждение

Представленный в настоящем исследовании подход, определяющий метод визуализации медицинских данных с применением графических библиотек JavaScript, несомненно, требует дальнейшего всестороннего изучения. Перспектива такого подхода заключается в удачном применении средств open source, определяющих графическое представление клинической информации для отображения медицинского документа, хранящегося в XML файле и определяющего содержание данно-

```

values.forEach(function (d) {
    d.Date = parse(d.Date);
    d.Insulin = +d.Insulin;});
x.domain([values[0].Date, values[values.length - 1].Date]);
y.domain([0, d3.max(values, function (d) { return d.Insulin;
}))).nice();
svg.selectAll(".data")
    .data(data)
    .enter().append("line")
    .attr("class", "bin")
    .attr("x1", function (d) { return x(d.Date); })
    .attr("x2", function (d) { return x(d.Date); })
    .attr("y1", h)
    .attr("y2", function (d) { return y(d.Insulin); });

```

Рис. 10. Код добавления гистограммы инсулина

```

var line = d3.svg.line()
    .interpolate("monotone")
    .x(function (d) { return x(d.Date); })
    .y(function (d) { return y(d.Glucose); });
values.forEach(function (d) {
    d.Date = parse(d.Date);
    d.Glucose = +d.Glucose; });
x.domain([values[0].Date, values[values.length - 1].Date]);
y.domain([0, d3.max(values, function (d) { return d.Glucose;
}))).nice();
svg.append("svg:clipPath")
    .attr("id", "clip")
    .append("svg:rect")
    .attr("width", w)
    .attr("height", h);
svg.append("svg:path")
    .attr("class", "line")
    .attr("clip-path", "url(#clip)")
    .attr("d", line(values));

```

Рис. 11. Код добавления кривой линии глюкозы

```

var legend = svg.selectAll(".legend")
    .data(color.domain())
    .enter().append("g")
    .attr("class", "legend")
    .attr("transform", "translate(-500," + (h + 20) + ")");
legend.append("rect")
    .attr("x", w - 80)
    .attr("width", 18)
    .attr("height", 18)
    .style("fill", "red");
legend.append("text")
    .attr("x", w - 84)
    .attr("y", 9)
    .attr("dy", ".35em")
    .style("text-anchor", "end")
    .text("Regular Insulin");
legend.append("rect")
    .attr("x", w + 410)
    .attr("width", 18)
    .attr("height", 18)
    .style("fill", "orange");
legend.append("text")
    .attr("x", w + 404)
    .attr("y", 9)
    .attr("dy", ".35em")
    .style("text-anchor", "end")
    .text("Glucose");

```

Рис. 12. Код добавления области легенды

го представления. В дальнейшем вышеописанный метод представления клинической информации может быть успешно применен в медицинских информационных, а также интеллектуальных медицинских системах, имеющих веб интерфейс и использующих стандартизированный подход к организации и хранению данных электронной истории болезни (ЭИБ) на основе формата ISO 13606. Так медицинский специалист сможет принимать эффективные клинические решения, основываясь на представленной динамике изменения медицинских показателей пациента, полученных лабораторными исследованиями или при проведении осмотров, с последующим занесением в ЭИБ. Помимо представленных в настоящей работе примеров визуализации динамики изменений таких показателей, как регулярный инсулин (regular insulin), инсулин ПНХ (NPH insulin), уровень сахара в крови (Glucose), пульс (Pulse), артериальное давление (Pressure), а также нормированных значений некоторых из них, также возможно представление интерфейса интерактивной ЭМК в целом. При реализации вышеприведенных примеров в качестве инструментов для создания графических представлений использовались разработки, представленные проектами «Data-Driven Documents» [16], а также «Flot» [17]. Также к подобным средствам можно отнести открытую библиотеку «Formatting Objects Processor» (FOP) [24]. Данная библиотека позволяет генерировать PDF документы с графическими представлениями медицинских данных в

формате «Scalable Vector Graphics» (SVG). Однако для работы совершения данных преобразований на платформе .NET в среде разработки MS Visual Studio необходима среда выполнения IKVM [25]. Данная среда также относится к компонентам open source. Описанный в настоящей работе метод был реализован для конкретных медицинских данных пациента с сахарным диабетом, применяемые для их визуализации шаблоны были настроены вручную.

Заключение

Полученные решения способствуют реализации гибких и тривиальных способов графического представления медицинских данных. Используя библиотеки JavaScript, разработчик получает мощный инструмент для создания графических элементов. Дополнительным достоинством средств open source является многообразие уже разработанных шаблонов – готовых решений. При данном подходе разработчику необходимо внести минимальные коррективы в исходный код для получения соответствующего графического представления клинической информации. Результат такого представления является простым и информативным, что особенно важно для пациентов, в частности для пациентов в престарелом возрасте. Таким образом, использование средств open source позволит значительно снизить финансовые затраты при создании EHR систем, а также сократить время разработки пользовательского интерфейса.

СПИСОК ЛИТЕРАТУРЫ

1. Sousa F., Viola L., Ferreira L. et al. An ecosystem of products and systems for ambient intelligence – the AAL4ALL users perspective // *Studies in health technology and informatics*. – 2012. – № 177. – P. 263–271.
2. Munoz A., Serrano E., Villa A. et al. An approach for representing sensor data to validate alerts in Ambient Assisted Living // *Sensors*. – 2012. – № 12 (5). – P. 6282–6306.
3. Fernandez-Llatas C., Garcia-Gomez J.M., Vicente J. et al. Behaviour patterns detection for persuasive design in Nursing Homes to help dementia patients // *IEEE Engineering in Medicine and Biology Society: Annual International Conference*. – Boston, Aug. 30 – Sept. 3, 2011. – P. 6413–6417.
4. Chen C., Haddad D., Selsky J. et al. Making sense of mobile health data: an open architecture to improve individual- and population-level health // *Journal of medical Internet research*. – 2012. – № 14 (4). – P. 112–117.
5. Aquilano M., Cavallo F., Bonaccorsi M. et al. Ambient Assisted Living and ageing: Preliminary results of RITA project // *IEEE Engineering in Medicine and Biology Society: Annual International Conference*. – San Diego, Aug. 28 – Sept. 1, 2012. – P. 5823–5826.
6. Demski H., Hildebrand C., Lopez Bolos J. et al. Technical requirements of a social networking platform for senior citizens // *Studies in health technology and informatics*. – 2012. – № 180. – P. 818–822.
7. Cascado D., Romero S.J., Hors S. et al. Virtual worlds to enhance Ambient-Assisted Living // *IEEE Engineering in Medicine and Biology Society: Annual International Conference*. – Buenos Aires, Aug. 31 – Sept. 4, 2010. – P. 212–215.
8. Veseli H., Kopanitsa G., Demski H. Standardized EHR interoperability – preliminary results of a German pilot project using the archetype methodology // *Studies in health technology and informatics*. – 2012. – № 180. – P. 646–650.
9. Model driven development of clinical information systems using openEHR / K. Atalag, H.Y. Yang, E. Tempero, J. Warren // *Studies in health technology and informatics*. – 2011. – № 169. – P. 849–853.
10. Cios K.J., Moore G.W. Uniqueness of medical data mining // *Artificial intelligence in medicine*. – 2002. – № 26 (1). – P. 1–24.
11. Van der Linden H., Austin T., Talmon J. Generic screen representations for future-proof systems, is it possible? There is more to a GUI than meets the eye // *Computer methods and programs in biomedicine*. – 2009. – № 95 (3). – P. 213–226.
12. Generic screen representations for future proof systems – is it possible? Two-model approach to a generic GUI / H. van der Linden, T. Schuler, R. Chen, J. Talmon // *Studies in health technology and informatics*. – 2007. – № 129 (P. 2). – P. 1122–1126.
13. Fonseca T., Ribeiro C., Granja C. Vital signs in intensive care: automatic acquisition and consolidation into electronic patient records // *Journal of medical systems*. – 2009. – № 33 (1). – P. 47–57.
14. Raghupathi W., Umar A. Upper-level ontologies for health information systems. Towards an archetype patterns approach // *Methods of information in medicine*. – 2011. – № 50 (3). – P. 285–295.
15. User interface of MUDR electronic health record / P. Hanzlicek, J. Spidlen, H. Heroutova, M. Nagy // *International journal of medical informatics*. – 2005. – № 74 (2). – P. 221–227.
16. Data-Driven Documents. URL: <http://d3js.org/> (дата обращения: 15.02.2013).

17. Attractive JavaScript plotting for jQuery. URL: <http://www.flotcharts.org/> (дата обращения: 15.02.2013).
18. Rinner C., Kohler M., Hubner-Bloder G. et al. Creating ISO/EN 13606 archetypes based on clinical information needs // *Studies in health technology and informatics*. – 2011. – № 165. – P. 43–48.
19. Duftschmid G., Wrba T., Rinner C. Extraction of standardized archetyped data from Electronic Health Record systems based on the Entity-Attribute-Value Model // *International journal of medical informatics*. – 2010. – № 79 (8). – P. 585–597.
20. Semantic validation of standard-based electronic health record documents with W3C XML schema / C. Rinner, S. Janzek-Hawlat, S. Sibinovic, G. Duftschmid // *Methods of information in medicine*. – 2010. – № 49 (3). – P. 271–280.
21. Mola E. Patient empowerment, an additional characteristic of the European definitions of general practice/family medicine // *The European journal of general practice*. – 2013. – № 19 (2). – P. 128–131.
22. Kopanitsa G. Standard based multiclient medical data visualization // *Studies in health technology and informatics*. – 2012. – № 180. – P. 199–203.
23. UC Irvine Machine Learning Repository. URL: <http://archive.ics.uci.edu/ml/datasets/Diabetes> (дата обращения: 10.02.2013).
24. The Apache FOP Project. URL: <http://xmlgraphics.apache.org/fop/> (дата обращения: 15.02.2013).
25. IKVM.NET HomePage. URL: <http://ikvm.net/> (дата обращения: 15.02.2013).

Поступила 02.09.2013 г.

UDC 004.22

APPLYING OPEN SOURCE DATA VISUALIZATION TOOLS TO STANDARD BASED MEDICAL DATA

Georgy D. Kopanitsa,

Cand. Sc., Tomsk Polytechnic University, 30, Lenin avenue, Tomsk, 634050, Russia; Tomsk State University of Architecture and Building, 2, Solyanaya sq., Tomsk, 634003, Russia. E-mail: georgy.kopanitsa@gmail.com

Maksim A. Taranik,

Tomsk Polytechnic University, 30, Lenin avenue, Tomsk, 634050, Russia. E-mail: taranik@tpu.ru

The urgency of the paper deals with the necessity of using flexible and simple software tools for standard-based medical data visualization.

The aim of the study: to implement ISO 13606 standard-based medical data visualization using open source tools.

The methods: ISO 13606 is an archetype clinical standard. A canonical approach and Open source JavaScript libraries are used for data transformation. MS Visual Studio is the development environment.

The results: For working with medical set of archetypes the XML informational model was developed. Using open source JavaScript libraries the insulin and blood sugar and dynamic blood sugar diagrams were constructed.

Conclusion: Applying the open source tools (JavaScript libraries) a developer gains a variety of prepared solutions which realize flexible and simple methods for standard-based medical data representation.

Key words:

Open source, visualization, archetypes, JavaScript, ISO 13606.

REFERENCES

1. Sousa F., Viola L., Ferreira L. An ecosystem of products and systems for ambient intelligence – the AAL4ALL users perspective. *Studies in health technology and informatics*, 2012, no. 177, pp. 263–271.
2. Munoz A., Serrano E., Villa A. An approach for representing sensor data to validate alerts in Ambient Assisted Living. *Sensors*, 2012, no. 12 (5), pp. 6282–6306.
3. Fernandez-Llatas C., Garcia-Gomez J.M., Vicente J. Behaviour patterns detection for persuasive design in Nursing Homes to help dementia patients. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Boston, Aug. 30 – Sept. 3, 2011. pp. 6413–6417.
4. Chen C., Haddad D., Selsky J. Making sense of mobile health data: an open architecture to improve individual- and population-level health. *Journal of medical Internet research*, 2012, no. 14 (4), pp. 112–117.
5. Aquilano M., Cavallo F., Bonaccorsi M. Ambient Assisted Living and ageing: Preliminary results of RITA project. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. San Diego, Aug. 28 – Sept. 1, 2012. pp. 5823–5826.
6. Demski H., Hildebrand C., Lopez Bolos J. Technical requirements of a social networking platform for senior citizens. *Studies in health technology and informatics*, 2012, no. 180, pp. 818–822.
7. Cascado D., Romero S.J., Hors S. Virtual worlds to enhance Ambient-Assisted Living. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Buenos Aires, Aug. 31 – Sept. 4, 2010. pp. 212–215.
8. Veseli H., Kopanitsa G., Demski H. Standardized EHR interoperability – preliminary results of a German pilot project using the archetype methodology. *Studies in health technology and informatics*, 2012, no. 180, pp. 646–650.
9. Atalag K., Yang H.Y., Tempero E., Warren J. Model driven development of clinical information systems using openEHR. *Studies in health technology and informatics*, 2011, no. 169, pp. 849–853.
10. Cios K.J., Moore G.W. Uniqueness of medical data mining. *Artificial intelligence in medicine*, 2002, no. 26 (1), pp. 1–24.

11. Van der Linden H., Austin T., Talmon J. Generic screen representations for future-proof systems, is it possible? There is more to a GUI than meets the eye. *Computer methods and programs in bio-medicine*, 2009, no. 95 (3), pp. 213–226.
12. Van der Linden H., Schuler T., Chen R., Talmon J. Generic screen representations for future proof systems – is it possible? Two-model approach to a generic GUI. *Studies in health technology and informatics*, 2007, no. 129 (P. 2), pp. 1122–1126.
13. Fonseca T., Ribeiro C., Granja C. Vital signs in intensive care: automatic acquisition and consolidation into electronic patient records. *Journal of medical systems*, 2009, no. 33 (1), pp. 47–57.
14. Raghupathi W., Umar A. Upper-level ontologies for health information systems. Towards an archetype patterns approach. *Methods of information in medicine*, 2011, no. 50 (3), pp. 285–295.
15. Hanzlicek P., Spidlen J., Heroutova H., Nagy M. User interface of MUDR electronic health record. *International journal of medical informatics*, 2005, no. 74 (2), pp. 221–227.
16. *Data-Driven Documents*. Available at: <http://d3js.org/> (accessed 15 February 2013).
17. *Attractive JavaScript plotting for jQuery*. Available at: <http://www.flotcharts.org/> (accessed 15 February 2013).
18. Rinner C., Kohler M., Hubner-Bloder G. Creating ISO/EN 13606 archetypes based on clinical information needs. *Studies in health technology and informatics*, 2011, no. 165, pp. 43–48.
19. Duftschmid G., Wrba T., Rinner C. Extraction of standardized archetyped data from Electronic Health Record systems based on the Entity-Attribute-Value Model. *International journal of medical informatics*, 2010, no. 79 (8), pp. 585–597.
20. Rinner C., Janzek-Hawlat S., Sabinovic S., Duftschmid G. Semantic validation of standard-based electronic health record documents with W3C XML schema. *Methods of information in medicine*, 2010, no. 49 (3), pp. 271–280.
21. Mola E. Patient empowerment, an additional characteristic of the European definitions of general practice/family medicine. *The European journal of general practice*, 2013, no. 19 (2), pp. 128–131.
22. Kopanitsa G. Standard based multiclient medical data visualization. *Studies in health technology and informatics*, 2012, no. 180, pp. 199–203.
23. *UC Irvine Machine Learning Repository*. Available at: <http://archive.ics.uci.edu/ml/datasets/Diabetes> (accessed 10 February 2013).
24. *The Apache FOP Project*. Available at: <http://xmlgraphics.apache.org/fop/> (accessed 15 February 2013).
25. *IKVM.NET Home Page*. Available at: <http://ikvm.net/> (accessed 15 February 2013).

УДК 004.75

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ЭМУЛЯЦИИ ПЕРЕДАЧИ МГНОВЕННЫХ ЗНАЧЕНИЙ ИЗМЕРЕНИЙ В СООТВЕТСТВИИ СО СТАНДАРТОМ МЭК 61850

Баранов Павел Федорович,

канд. техн. наук, инженер кафедры компьютерных измерительных систем и метрологии Института кибернетики Томского политехнического университета, Россия, 634050, г. Томск, пр. Ленина, д. 30. E-mail: bpf@tpu.ru

Муравьев Сергей Васильевич,

д-р техн. наук, зав. кафедрой компьютерных измерительных систем и метрологии Института кибернетики Томского политехнического университета, Россия, 634050, г. Томск, пр. Ленина, д. 30. E-mail: muravyov@tpu.ru

Сулайманов Алмаз Омурзакович,

канд. техн. наук, начальник научного отдела Энергетического института Томского политехнического университета, Россия, 634050, г. Томск, пр. Ленина, д. 30. E-mail: sao@tpu.ru

Леванова Вероника Сергеевна,

студент кафедры компьютерных измерительных систем и метрологии Института кибернетики Томского политехнического университета, Россия, 634050, г. Томск, пр. Ленина, д. 30. E-mail: levanova@tpu.ru

Актуальность работы обусловлена необходимостью анализа требований стандарта МЭК 61850–9-2 к передаче мгновенных значений измерений от цифровых трансформаторов тока и напряжения к устройствам защиты, автоматики, контроля и учета качества электрической энергии. Эти требования необходимо также учитывать при разработке программных и аппаратных средств для эмуляции передачи измеренных мгновенных значений параметров электрической энергии.

Цель работы: выявить особенности процесса передачи мгновенных значений измерений от цифровых трансформаторов тока и напряжения в соответствии с требованиями стандарта МЭК 61850–9-2 и разработать программное обеспечение для эмуляции данного процесса.

Методы исследования: моделирование систем и данных.

Результаты: представлено разработанное программное обеспечение для эмуляции передачи мгновенных значений измерений в соответствии со стандартом МЭК 61850–9-2 и рекомендациями МЭК 61850–9-2 LE, позволяющими эмулировать передачу мгновенных значений измерений с частотой 80 или 256 выборок на период для четырех синусоидальных токов и четырех синусоидальных напряжений, и результаты проверки правильности его функционирования с помощью двух независимых программ. Приведено описание протокола МЭК 61850–9-2 и полей данных, которое может быть использовано для разработки программного обеспечения цифровых подстанций.

Ключевые слова:

Мгновенные значения, эмуляция электрических систем, интеллектуальные электрические сети, стандарты МЭК, протоколы, цифровая подстанция.

Введение

Современные тенденции в области электроэнергетики направлены на построение цифровых подстанций и создание интеллектуальных распределенных сетей (Smart grid) [1–5]. Цифровая подстанция представляет собой комплекс интеллектуальных электронных устройств, реализующих задачи контроля и учета качества электрической энергии, релейной защиты и автоматики, а также регистрации аварийных событий [6–9].

Для решения задач стандартизации при построении цифровой подстанции Международная электротехническая комиссия (МЭК) опубликовала стандарт МЭК 61850 «Сети и системы связи на подстанциях» [10], в котором регламентируются схемные решения для подстанции, защиты, автоматики и измерений, а также конфигурация

устройств и вопросы передачи информации между отдельными устройствами. В части стандарта МЭК 61850-9-2 [11] регламентируется передача мгновенных значений измерений от цифровых трансформаторов тока и напряжения по цифровым сетям передачи данных к устройствам защиты, автоматики, контроля и учета качества электрической энергии.

Для настройки, тестирования и ввода в эксплуатацию узлов цифровых подстанций разработчику необходимо иметь программные и аппаратные средства для эмуляции передачи измеренных мгновенных значений параметров электрической энергии в соответствии с требованиями МЭК. Под эмулярованием будем понимать процесс воспроизведения сигналов и данных, передаваемых по интерфейсным линиям стандарта МЭК 61850, для

имитации работы реальных цифровых трансформаторов тока и напряжения.

Несмотря на то, что многие разработчики ведут работы в данном направлении [12–16], эмуляторы, реализующие полный набор описанных в стандарте функций, не существуют. Так, например, в существующих эмуляторах не поддерживаются требования к передаче данных с частотой 256 выборок на период, отсутствует возможность изменения коэффициента нелинейных искажений и т. д.

Цель данной статьи состоит в описании реализующего полный набор функций протокола МЭК 61850-9-2 для передачи мгновенных значений измерений, соответствующих полей данных и разработанного программного обеспечения (ПО) для эмулярования с учетом дополнительных рекомендаций МЭК 61850-9-2 LE [17].

Формирование пакета мгновенных значений измерений

Стандарт МЭК 61850 определяет пять типов сообщений для передачи данных:

- Sampled Values (SV) – передача мгновенных значений измерений тока и напряжения;
- Generic Object Oriented Substation Event (GOOSE) – передача объектно-ориентированных событий;
- Time Sync – передача меток синхронизации устройств;
- Manufacturing Messaging Specification (MMS) – передача сообщений внутри подстанции;
- Generic Substation Status Event (GSSE) – передача общих событий состояния на подстанции.

Для обмена данными в соответствии со стандартом МЭК 61850 используется сетевая модель OSI, в зависимости от типа сообщения используются раз-

личные коммуникационные профили, представленные на рис. 1.

Для SV, GOOSE и GSSE сообщений критично время доставки, поэтому они реализуются непосредственно на канальном уровне.

Сообщения MMS представляют собой стандартный клиент-серверный протокол поверх стека TCP/IP.

Сообщения Time Sync рассылаются всем интеллектуальным электронным устройствам на подстанции поверх стека UDP/IP.

Для передачи SV, GOOSE и GSSE сообщений формируется стандартизированный Ethernet кадр [17], структура которого показана на рис. 2.

Кадр начинается с преамбулы (Preamble), длина поля которой составляет 7 байт. Каждый байт содержит последовательность битов – 10101010. Преамбула используется для синхронизации приемопередатчиков. 8-ой байт кадра (Start of frame) содержит последовательность битов – 10101011 и указывает на начало кадра.

MAC адрес приемника (ов) (Destination address) – адрес устройства, которому направляется сообщение, длина поля 6 байт. Рассылка сообщений может быть одноадресной и многоадресной; для многоадресной рассылки SV сообщений стандарт МЭК 61850 регламентирует набор адресов от 01-0C-CD-04-00-00 до 01-0C-CD-04-01-FF.

MAC адрес источника (Source address) – уникальный адрес передающего интеллектуального электронного устройства (ИЭУ), длина поля 6 байт.

Идентификатор протокола (TPID, Tag Protocol Identifier) указывает, какой тип протокола используется. Длина поля 2 байта. Стандарт МЭК 61850 регламентирует использовать значение 0x8100.

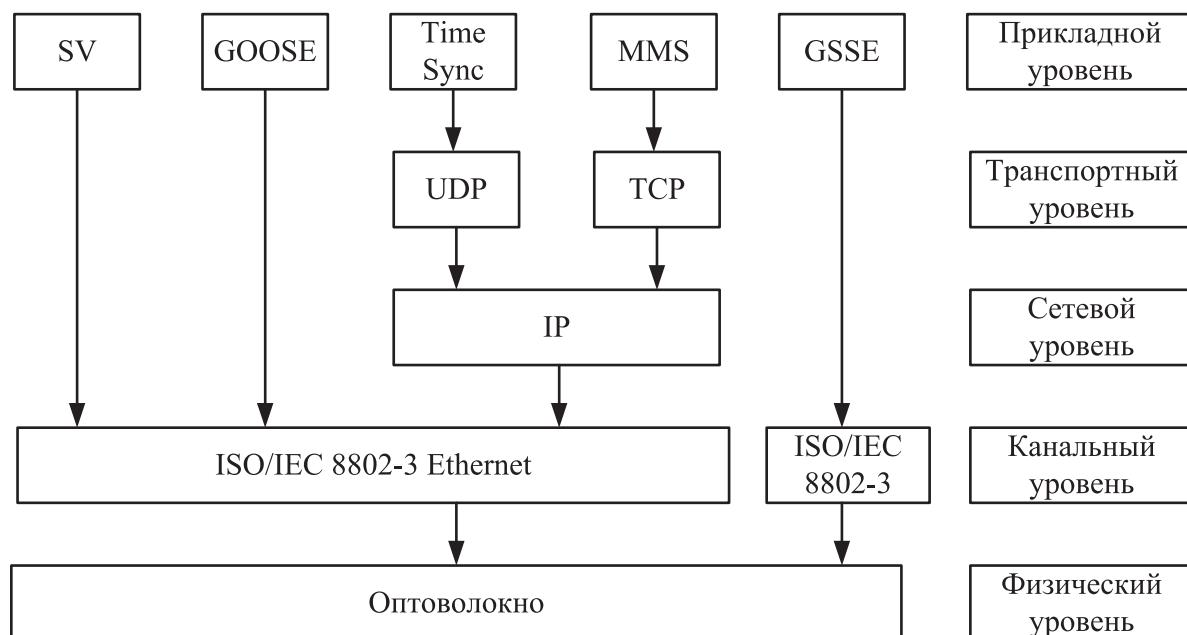


Рис. 1. Коммуникационные профили стандарта МЭК 61850

	8	7	6	5	4	3	2	1
1	Preamble							
2								
3								
4								
5								
6								
7								
8	Start of frame							
9	Destination address							
10								
11								
12								
13								
14								
15	Source address							
16								
17								
18								
19								
20								
21	TPID							
22								
23	User priority		CFI	VID				
24	VID							
25	Ethertype							
26								
27	APPID							
28								
29	Length							
30								
31	reserved 1							
32								
33	reserved 2							
34								
35	APDU							
.								
.	Frame check sequence							
.								
.								

Рис. 2. Формат стандартизированного Ethernet кадра

Приоритет (User priority) используется для задания приоритета передаваемого трафика. Длина поля 3 бита. Значение для SV сообщений устанавливается равным 0b100.

Индикатор канонического формата (CFI, Canonical Format Indicator) указывает на формат MAC-адреса. 0 – канонический, 1 – неканонический. Длина поля 1 бит.

Идентификатор виртуальной локальной компьютерной сети (VID, Virtual Local Area Network Identifier) указывает, какой сети принадлежит кадр. Длина поля 12 бит.

Идентификатор сообщения (Ethertype) – зарезервированное значение, непосредственно определяющее используемый тип сообщений. Длина поля 2 байта. Значение для SV сообщений устанавливается равным 0x88BA

Идентификатор приложения (APPID, Application Identifier) используется для разделения сообщений. Длина поля 2 байта. Значение для SV сообщений устанавливается равным 0x4000

Длина данных (Length) – суммарная длина полей APPID, Length, reserved 1, reserved 2 и APDU. Длина поля 2 байта.

Зарезервировано 1 (reserved 1) – зарезервированное поле, по умолчанию значение 0x0000. Если старший байт установлен в 1, то ИЭУ находится в режиме тестирования. Длина поля 2 байта.

Зарезервировано 2 (reserved 2) – зарезервированное поле для дополнительных параметров безопасности, по умолчанию значение 0x0000. Длина поля 2 байта.

Прикладной протокол данных (APDU, Application Protocol Data Unit) непосредственно содержит измерительную информацию.

Контрольная сумма (Frame check sequence) – контрольное значение, вычисляемое по алгоритму CRC-32. Длина поля 4 байта.

Структура прикладного протокола данных APDU для сообщений SV 256 выборок на период показана на рис. 3.

Синтаксис передачи сообщений имеет формат тройки TLV (Tag, Length, Value) (Метка, Длина, Значение), как показано на рис. 3. Все поля (T, L, V) являются последовательными октетами. Значение V может быть тройкой самой TLV, если оно составное. Для сообщений SV 256 информация об одном периоде измерений передается с помощью 32 пакетов по 8 мгновенных значений в каждом.

Поле savPdu – начало сообщения. Длина поля 4 байта, первый байт 0x60 – метка поля, второй байт 0x82, с третьего байта длина сообщения.

Поле noASDU – количество блоков данных (ASDU). Длина поля 3 байта, первый байт 0x80 – метка поля, второй байт 0x01 – длина значения, третий байт 0x08 – значение.

Поле Sequence of ASDU – начало блоков данных. Длина поля 4 байта, первый байт 0xA2 – метка поля, второй байт 0x82, с третьего байта общая длина блоков данных.

Поле Sequence ASDU 1 – идентификатор начала блока данных. Длина поля 2 байта, первый байт 0x30 – метка поля, второй байт – длина блока данных.

Поле svID – идентификатор мгновенных значений. Длина поля от 21 до 69 байт, первый байт 0x80 – метка поля, второй байт – длина значения идентификатора мгновенных значений (от 10 до 34 байт), с третьего байта значение идентификатора мгновенных значений длиной от 10 до 34 байт соответственно.

Поле smpCnt – номер выборки. Длина поля 4 байта, первый байт 0x82 – метка поля, второй

savPdu	60	L(751...943)					
noASDU		80	L (1)	8			
Sequence of ASDU		A2	L(744...936)				
Sequence ASDU1			30	L (91...115)			
svID			ASDU 1	80	L (10...34)	Values	
smpCnt				82	L (2)	Values	
confRev				83	L (4)	Values	
smpSynch				85	L (1)	Values	
Sequence of Data				87	L (64)	Values	
						Data set	Values
							Values
							Values
							Values
							Values
Sequence ASDU 2			30	L (91...115)	ASDU 2		
Sequence ASDU 3			30	L (91...115)	ASDU 3		
Sequence ASDU 4			30	L (91...115)	ASDU 4		
Sequence ASDU 5			30	L (91...115)	ASDU 5		
Sequence ASDU 6			30	L (91...115)	ASDU 6		
Sequence ASDU 7			30	L (91...115)	ASDU 7		
Sequence ASDU 8			30	L (91...115)	ASDU 8		

Рис. 3. Структура протокола данных APDU для сообщений SV 256 выборки на период

байт – 0x02 длина значения, с третьего байта значение номера выборки.

Поле confRev – номер конфигурации. Длина поля 6 байт, первый байт 0x83 – метка поля, второй байт – 0x04 длина значения, с третьего байта значение номера конфигурации.

Поле smpSynch – метка наличия синхронизации. Длина поля 3 байта, первый 0x85 – метка поля, второй байт – 0x01 длина значения, третий байт – значение.

Поле Sequence of Data – последовательность мгновенных значений. Длина поля 66 байт, первый байт 0x87 – метка поля, второй байт – 0x40 длина последовательности мгновенных значений, с третьего байта последовательность мгновенных значений. Поле Sequence of Data содержит информацию о мгновенных значениях токов и напряжений фаз А, В, С и нейтрали N. Каждое измеренное значение кодируется 8-ми байтным кодом, формат кадра для одного мгновенного значения тока показан на рис. 4.

Поле InnxTCTR1.Amp.instMag.i – амплитудная величина мгновенных значений тока, где x – фаза А, В или С. Масштабный коэффициент: 1 бит = 1 мА. Длина поля 4 байта.

	8	7	6	5	4	3	2	1
1	InnxTCTR1.Amp.instMag.i							
2								
3								
4								
5	InnxTCTR1.Amp.q							
6								
7			der	OpB	Test	Src	DetailQual	
8	DetailQual						validity	

Рис. 4. Формат кадра значения тока

Поле InnxTCTR1.Amp.q – дополнительная информация о мгновенных значениях тока, где x – фаза А, В или С. Длина поля 2 байта, значение по умолчанию 0x0000.

Поле der – идентификатор значения. Длина поля 1 бит. Значение показывает, являются ли мгновенные значения результатами измерений (0b0) или результатами расчета (0b1).

Поле OpB – идентификатор блокировки, устанавливается в случае, если оператор блокирует дальнейшее обновление значений. Значением является информация, полученная перед блокировкой.

кой. Длина поля 1 бит. Если этот идентификатор установлен, должен быть также установлен флаг oldData идентификатора detailQual.

Поле Test – идентификатор для указания тестового значения, которое не предназначено для использования в операционных целях. Длина поля 1 бит.

Поле Src – идентификатор происхождения значения. Длина поля 1 бит. Значение может быть получено из процесса (process, 0b0) или может быть установленным пользователем (substituted, 0b1).

Поле DetailQual – дополнительный идентификатор качества. Длина поля 1 байт. Устанавливаемые флаги идентификатора DetailQual приведены в таблице.

Таблица. Идентификатор DetailQual

Бит	Флаг
1	overflow (переполнение)
2	outOfRange (за пределами диапазона)
3	badReference (требуется калибровка)
4	oscillatory (колебательный)
5	failure (повреждение)
6	oldData (старые данные)
7	inconsistent (несогласованный)
8	inaccurate (неточный)

Поле validity – идентификатор качества может быть хорошим (good), сомнительным (questionable) или недействительным (invalid). Длина поля 2 бита. Значение по умолчанию good – 0b00.

При наличии одного из флагов идентификатора DetailQual и идентификатор Validity должен быть настроен на значение invalid (0b01) или questionable (0b11).

В случае если величины тока и/или напряжения не измеряются, то они транслируются в канал

передачи данных с нулевыми значениями. Кроме того, для этих значений должен быть установлен идентификатор качества invalid.

Программное обеспечение для эмулирования передачи мгновенных значений измерений

Для эмулирования процесса передачи мгновенных значений измерений и тестирования оборудования цифровых подстанций в среде графического программирования LabVIEW было разработано специальное ПО – виртуальный прибор «IED Emulator».

Лицевая панель разработанного виртуального прибора представлена на рис. 5.

Виртуальный прибор позволяет выбрать сетевой адаптер, через который будет формироваться передача мгновенных значений, настроить параметры передачи (Destination address, TPID, User priority, CFI, VID, svID, reserved 1, reserved 2, smpSynch, confRev), эмулировать передачу мгновенных значений измерений с частотой 80 или 256 выборок на период для четырех синусоидальных токов и четырех синусоидальных напряжений.

Виртуальный прибор позволяет непосредственно во время эмулирования сигналов изменять их амплитуды и спектральный состав, значения флагов качества и флага синхронизации по времени и другие параметры. Благодаря этому, в отличие от других существующих эмуляторов, разработанное ПО позволяет тестировать оборудование подстанции в динамическом режиме, т. е. отслеживать результаты воздействия передаваемых параметров на устройства защиты, автоматики, контроля и учета качества электрической энергии.

Для работы непосредственно с канальным уровнем на языке программирования C++ разработана динамически подключаемая библиотека (DLL) с использованием библиотеки WinPcap (Windows Packet Capture) [18].

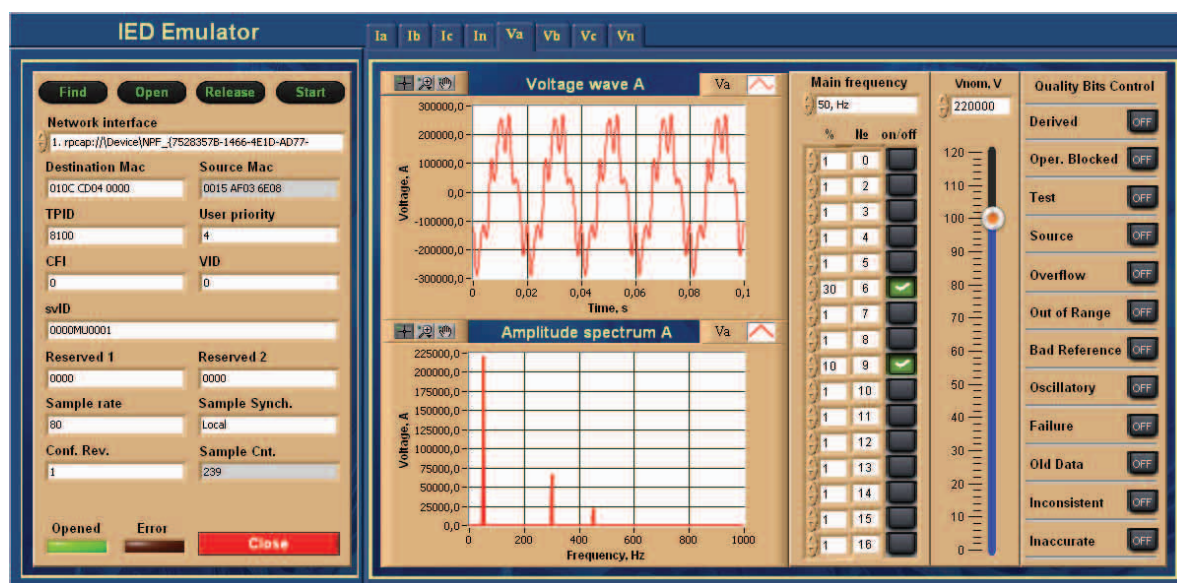


Рис. 5. Лицевая панель виртуального прибора «IED Emulator»

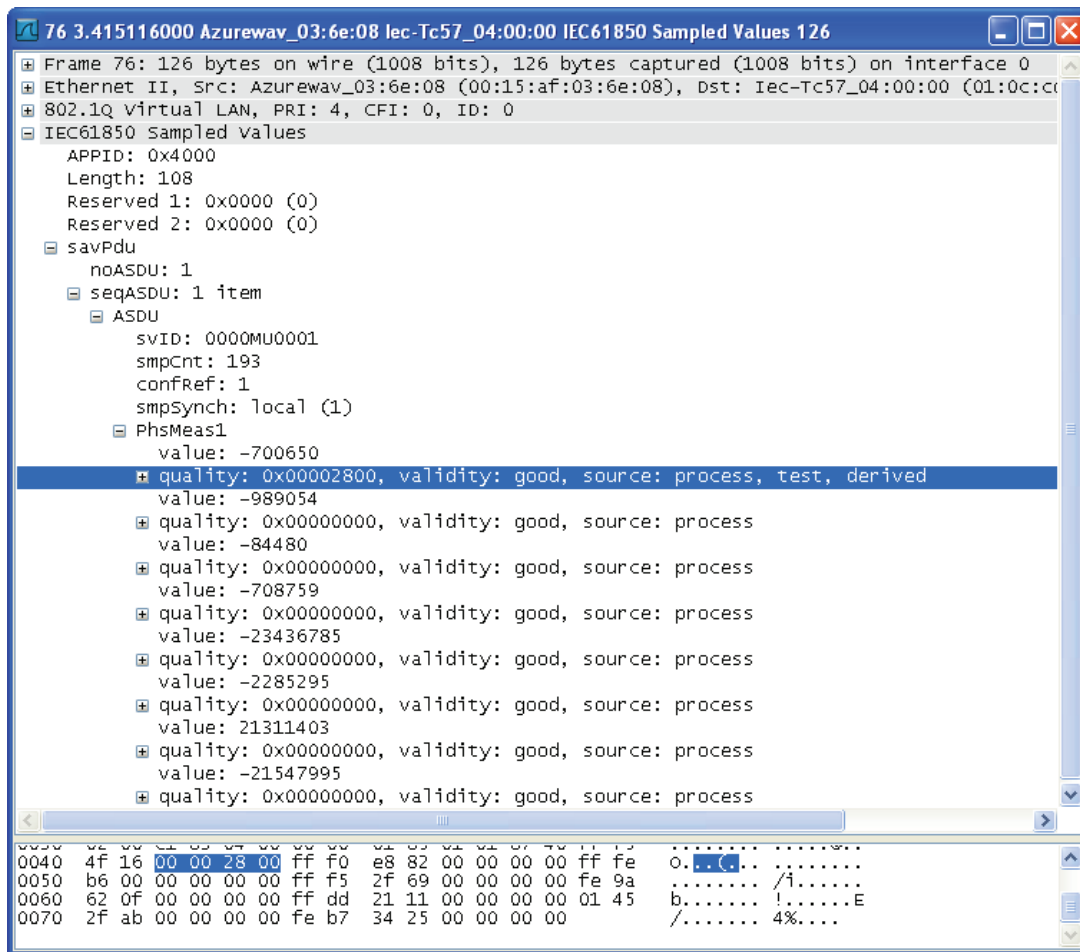


Рис. 6. Декодированный первый блок данных

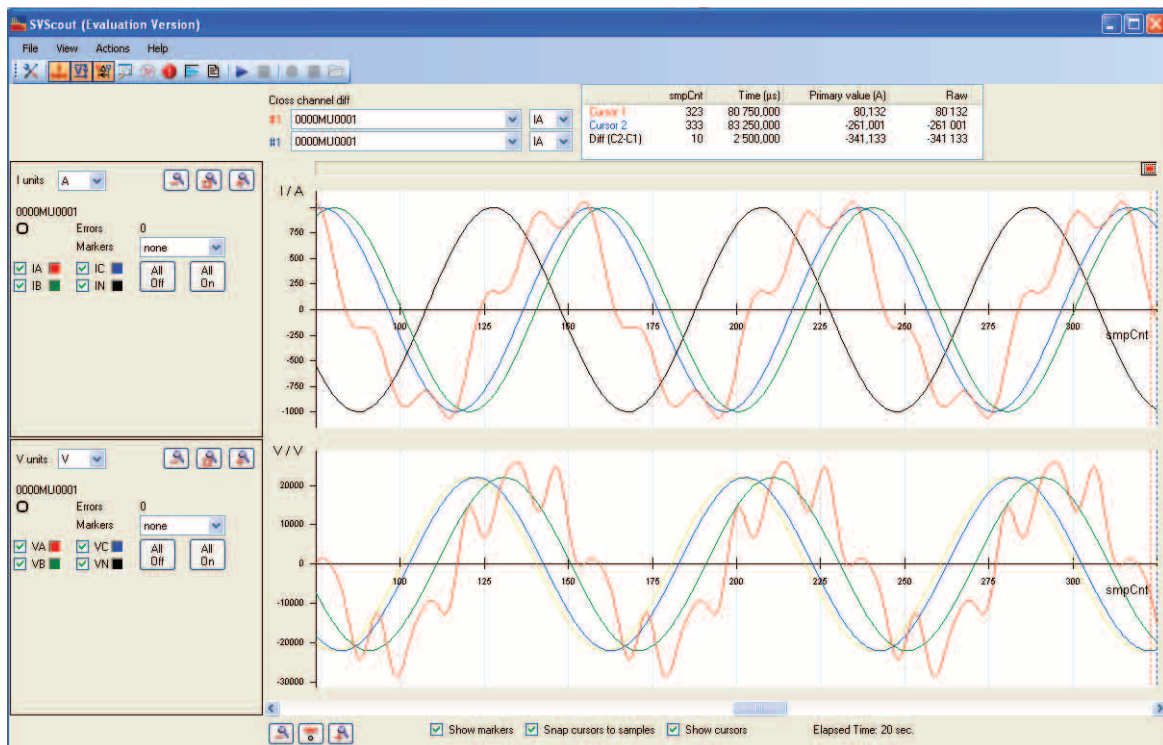


Рис. 7. Визуальный анализ передаваемого пакета данных

Проверка разработанного программного обеспечения

Для проверки правильности функционирования виртуального прибора использовался сетевой анализатор Wireshark [19]. Данный анализатор поддерживает протоколы стандарта МЭК 61850 и позволяет регистрировать соответствующий трафик. На рис. 6 приведен результат анализа передаваемых мгновенных значений измерений.

Для визуального анализа правильности функционирования разработанного виртуального прибора использовалось ПО SVScout [20]. SVScout декодирует потоки выборочных значений измерений с частотой 80 выборок на период от интеллектуальных устройств и отображает формы колебаний напряжений и токов в окне осциллографа. Результаты анализа эмулированных сигналов с помощью ПО SVScout представлены на рис. 7.

Результаты проверки правильности функционирования виртуального прибора с помощью двух независимых программ показали, что пакет данных сформирован и транслируется без ошибок, данные однозначно декодируются. Интервалы времени между передаваемыми пакетами не превышали 120 мкс.

СПИСОК ЛИТЕРАТУРЫ

1. Concept for intelligent distributed power system automation with IEC 61850 and IEC 61499 / N. Higgins, V. Vyatkin, N.C. Nair, K. Schwarz // Proceedings of IEEE International Conference on Systems, Man and Cybernetics. – Singapore, October 12–15, 2008. – P. 36–41.
2. Gungor V.C., Sahin D., Kocak T., et al. Smart Grid Technologies: Communication Technologies and Standards // Industrial Informatics. – 2011. – V. 7. – № 4. – P. 529–539.
3. Apostolov A. IEC 61850 9-2 process bus applications and benefits // Proceedings of XX IET International Conference on Developments in Power System Protection, DPSP 2010. – United Kingdom, Manchester, March 29 – April 1, 2010. – P. 48–53.
4. Ruihua Z., Yumei D., Yuhong L. New challenges to power system planning and operation of smart grid development in china // Proceedings of International conference on power system technology: Technological innovations making power grid smarter, POWERCON 2010. – China, Hangzhou, October 24–28, 2010. – P. 1–8.
5. Smart grid communications: Overview of research challenges, solutions, and standardization activities / Z. Fan, P. Kulkarni, S. Gormus, C. Efthymiou, G. Kalogridis, M. Sooriyabandara, et al. // IEEE Communications Surveys and Tutorials. – 2013. – V. 15. – № 1. – P. 21–38.
6. McGhee J., Goraj M. Smart High Voltage Substation Based on IEC 61850 Process Bus and IEEE 1588 Time Synchronization // Proceedings of First IEEE International Conference Smart Grid Communications. – USA, Gaithersburg, October 4–6, 2010. – P. 489–494.
7. Andersson L., Brunner C., Engler F. Substation automation based on IEC 61850 with new process-close technologies // Proceedings of IEEE Bologna Power Tech Conference. – Italy, Bologna, June 23–26, 2003. – P. 271–276.
8. Ikbali Ali Mini S. Sunil Gupta. Substation communication architecture to realize the future Smart Grid // Journal of energy technologies and policy. – 2011. – V. 1. – № 4. – P. 25–35.
9. Improvements in dependability and usability for a substation automation system with redundancy / H. Ito, K. Kaneda, K. Hamamatsu, T. Tanaka, K. Nara // Proceedings of XII WSEAS Inter-

Выводы

1. Для эмулирования процесса передачи мгновенных значений измерений в соответствии со стандартом МЭК 61850 с частотой 80 или 256 выборок на период разработано специальное программное обеспечение – виртуальный прибор «IED Emulator».
2. Проверка правильности функционирования виртуального прибора с помощью двух независимых программ показала, что виртуальный прибор работает без ошибок и пригоден для тестирования оборудования на соответствие стандарту МЭК 61850 в части приема и декодирования мгновенных значений измерений.
3. Разработанный виртуальный прибор можно использовать для тестирования программного обеспечения и приборов учета электрической энергии на подстанции.
4. Приведенные в статье описания протокола МЭК 61850–9-2 и полей данных могут быть использованы для разработки программного обеспечения цифровых подстанций.

Работа выполнена в рамках федеральной целевой программы «Исследования и разработка по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 гг.» (государственный контракт № 14.516.12.0009).

- national Conference on Systems. – Greece, Heraklion, July 22–24, 2008. – P. 713–721.
10. ГОСТ Р 54835–2011/IEC/TR 61850–1:2003. Сети и системы связи на подстанциях. Ч. 1. Введение и обзор. – Введ. с 2012–09–01 – М.: Стандартинформ, 2012. – 45 с.
11. IEC 61850–9-2 Sampled values over ISO/IEC 8802–3. – Geneva, Switzerland: IEC Central Office, 2004. – 34 p.
12. Kuffel R., Ouellette D., Forsyth P. Real time simulation and testing using IEC 61850 // Proceedings of the International Symposium Modern Electric Power Systems (MEPS). – Poland, Wroclaw, September 20–22, 2010. – P. 1–8.
13. Triangle MicroWorks – Protocol Solutions for IEC 61850. URL: <http://www.trianglemicroworks.com/ProductPage.aspx> (дата обращения: 05.07.2013).
14. OMICRON – Испытание с протоколами обмена данными. URL: <http://www.omicron.at/ru/products/pro/communication-protocols/> (дата обращения: 07.07.2013).
15. TEKVEL – IEC 61850. URL: <http://tekvel.ru> (дата обращения: 02.08.2013).
16. PETOM-61850 – Устройство испытательное. URL: <http://www.dynamics.com.ru/production/retom-61850/> (дата обращения: 12.07.2013).
17. IEC 61850–9-2 LE: Implementation Guideline for Digital Interface to Instrument Transformers Using IEC 61850–9-2. URL: <http://www.ucainter-national.org> (дата обращения: 05.07.2013).
18. Winpcap – Industry-standard tool for link-layer network access in Windows. URL: <http://www.winpcap.org> (дата обращения: 11.07.2013).
19. Wireshark – Network Protocol Analyzer. URL: <http://www.wireshark.org> (дата обращения: 15.07.2013).
20. SVScout – Программные средства для визуального представления результатов работы IEC 61850 Sampled Values. URL: <http://www.omicron.at/ru/products/pro/communication-protocols/svscout/> (дата обращения: 15.07.2013).

Поступила 17.09.2013 г.

UDC 004.75

SOFTWARE FOR EMULATING THE SAMPLED VALUES TRANSMISSION IN ACCORDANCE WITH IEC 61850 STANDARD

Pavel F. Baranov,

Cand. Sc., Tomsk Polytechnic University, 30, Lenin Avenue,
Tomsk, 634050, Russia. E-mail: bpf@tpu.ru

Sergey V. Muravyov,

Dr. Sc., Tomsk Polytechnic University, 30, Lenin Avenue,
Tomsk, 634050, Russia. E-mail: muravyov@tpu.ru

Almaz O. Sulaymanov,

Cand. Sc., Tomsk Polytechnic University, 30, Lenin Avenue,
Tomsk, 634050, Russia. E-mail: sao@tpu.ru

Veronika S. Levanova,

Tomsk Polytechnic University, 30, Lenin Avenue,
Tomsk, 634050, Russia. E-mail: levanova@tpu.ru

The importance of the work is conditioned by the need to analyze the requirements of IEC 61850-9-2 for transmission of measured sampled values from the non-conventional current and voltage transformers to units of relay protection and automation, control and electric power fiscal accounting. These requirements should be also taken into account when developing both software and hardware for emulation of transmission of measured instantaneous values of electric energy.

The main aim of the study: to identify the features of the sampled values transmission from non-conventional current and voltage transformers in accordance with IEC 61850-9-2 and to develop software to emulate this process.

The methods used in the study: modeling of systems and data.

The results: The paper introduces the developed software for emulating transmission of sampled values at 80 or 256 samples per period for phase and neutral line sinusoidal currents and voltages in accordance with IEC 61850-9-2 and IEC 61850-9-2LE and the results of verification of the software operation using two independent programs. The authors have described the protocol of IEC 61850-9-2 and appropriate data fields that can be used to develop the software of digital substations.

Key words:

Sample values, power system emulation, smart power grids, IEC standards, protocols, digital substation.

REFERENCES

- Higgins N., Vyatkin V., Nair N. C., Schwarz K. Concept for intelligent distributed power system automation with IEC 61850 and IEC 61499. *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*. Singapore, October 12–15, 2008. pp. 36–41.
- Gungor V.C., Sahin D., Kocak T. Smart Grid Technologies: Communication Technologies and Standards. *Industrial Informatics*, 2011, vol. 7, no. 4, pp. 529–539.
- Apostolov A. IEC 61850 9-2 process bus applications and benefits. *Proceedings of XX IET International Conference on Developments in Power System Protection, DPSP 2010*. United Kingdom, Manchester, March 29 – April 1, 2010. pp. 48–53.
- Ruihua Z., Yumei D., Yuhong L. New challenges to power system planning and operation of smart grid development in china. *Proceedings of International conference on power system technology: Technological innovations making power grid smarter, POWERCON 2010*. China, Hangzhou, October 24–28, 2010. pp. 1–8.
- Fan Z., Kulkarni P., Gormus S., Efthymiou C., Kalogridis G., Soriyabandara M. Smart grid communications: Overview of research challenges, solutions, and standardization activities. *IEEE Communications Surveys and Tutorials*, 2013, vol. 15, no. 1, pp. 21–38.
- McGhee J., Goraj M. Smart High Voltage Substation Based on IEC 61850 Process Bus and IEEE 1588 Time Synchronization. *Proceedings of First IEEE International Conference Smart Grid Communications*. USA, Gaithersburg, October 4–6, 2010. pp. 489–494.
- Andersson L., Brunner C., Engler F. Substation automation based on IEC 61850 with new process-close technologies. *Proceedings of IEEE Bologna Power Tech Conference*. Italy, Bologna. June 23–26, 2003. pp. 271–276.
- Ikbali Ali Mini S., Sunil Gupta. Substation communication architecture to realize the future Smart Grid. *Journal of energy technologies and policy*, 2011, vol. 1, no. 4, pp. 25–35.
- Ito H., Kaneda K., Hamamatsu K., Tanaka T., Nara K. Improvements in dependability and usability for a substation automation system with redundancy. *Proceedings of XII WSEAS International Conference on Systems*. Greece, Heraklion, July 22–24, 2008. pp. 713–721.
- GOST R 54835–2011/IEC/TR 61850–1:2003. *Seti i sistemy svyazi na podstancijah. Ch. 1. Vvedenie i obzor* [State Standard R 54835–2011/IEC/TR 61850–1:2003. Communication networks and systems in substations. P. 1. Introduction and overview]. Moscow, Standartinform, 2012. 45 p.
- IEC 61850-9-2 Sampled values over ISO/IEC 8802-3. Geneva, Switzerland: IEC Central Office, 2004. 34 p.
- Kuffel R., Ouellette D., Forsyth P. Real time simulation and testing using IEC 61850. *Proceedings of the International Symposium Modern Electric Power Systems (MEPS)*. Poland, Wroclaw, September 20–22, 2010. pp. 1–8.
- Triangle MicroWorks – Protocol Solutions for IEC 61850. Available at: <http://www.trianglemicroworks.com/ProductPage.aspx> (accessed 5 July 2013).
- OMICRON – Ispytanie s protokolami obmena dannymi [Test with communication protocols]. Available at: <http://www.omicron.com>

- ron.at/ru/products/pro/communication-protocols/ (accessed 7 July 2013).
15. *TEKVEL – IEC 61850*. Available at: <http://tekvel.ru> (accessed 2 August 2013).
16. *RETOM-61850 – Ustroystvo ispytatelnoe* [Testing instrument]. Available at: <http://www.dynamics.com.ru/production/retom-61850/> (accessed 12 July 2013).
17. *IEC 61850–9-2 LE: Implementation Guideline for Digital Interface to Instrument Transformers Using IEC 61850–9-2*. Available at: <http://www.ucainter-national.org> (accessed 5 July 2013).
18. *Winpcap – Industry-standard tool for link-layer network access in Windows*. Available at: <http://www.winpcap.org> (accessed 11 July 2013).
19. *Wireshark – Network Protocol Analyzer*. Available at: <http://www.wireshark.org> (accessed 15 July 2013).
20. *SVScout – Programmnye sredstva dlya vizualnogo predstavleniya rezultatov raboty* [Software Tool for Visualizing IEC 61850 Sampled Values]. Available at: <http://www.omicron.at/ru/products/pro/communication-protocols/svscout/> (accessed 15 July 2013).

Редактирование *А.С. Глазырин*
Компьютерная верстка *О.Ю. Аршинова*
Перевод на англ. язык и корректура *С.В. Жаркова*

Подписано к печати 04.06.2014. Формат 60х84/8. Бумага «Снегурочка».
Печать XEROX. Усл. печ. л. 16,28. Уч.-изд. л. 14,73.
Заказ 527-14. Тираж 500 экз.



Национальный исследовательский Томский политехнический университет
Система менеджмента качества
Издательства Томского политехнического университета
сертифицирована в соответствии с требованиями ISO 9001:2008



ИЗДАТЕЛЬСТВО  ТПУ. 634050, г. Томск, пр. Ленина, 30.
Тел./факс: 8(3822) 563-291, www.tpu.ru, izv@tpu.ru