

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Том 22
2016
№ 1

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

СОДЕРЖАНИЕ

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Кулагин В. П., Дубинин В. Н. Структурный анализ сетей Петри 3

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ

Барский А. Б., Желенков Б. В. Средства оптимизации информационного взаимодействия ресурсных процессоров для минимизации времени "облачных" вычислений. 14

Мандрикова О. В., Фетисова (Глушкова) Н. В., Полозов Ю. А. Моделирование временного хода параметров ионосферы и выделение аномалий. 21

Светушков Н. Н. Применение интегрального метода струн для расчета тепловых нагрузок на охлаждаемые стенки сопловой части ракетного двигателя. 32

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ И СЕТИ

Саак А. Э. Диспетчеризация заявок кругового типа в Grid-системах 37

Поршнев С. В., Божалкин Д. А., Копосов А. С. Опыт использования суперкомпьютера для обработки дампов сетевого трафика магистрального интернет-канала 42

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ИЗОБРАЖЕНИЙ

Наумов А. С. Метод автоматизированной реконструкции фрагментированных изображений на основе совокупности признаков фрагментов 48

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ

Щеглов К. А., Щеглов А. Ю. Интерпретация и моделирование угрозы атаки на информационную систему. Часть 2. Моделирование угрозы атаки 54

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В БИМЕДИЦИНСКИХ СИСТЕМАХ

Полосин В. Г., Бодин О. Н. Энтропийно-параметрический критерий проверки адекватности модели распределения ионных токов миокарда 65

Абдуллаев Н. Т., Гасанкулиева М. М. Информативные параметры и флуктуационная динамика электрокардиограмм при фликкер-шумовой спектроскопии 72

ПРИКЛАДНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Савченко В. В. Распознавание речевых команд методом фонетического декодирования слов с подавлением фонового шума 76

Главный редактор:
СТЕМПКОВСКИЙ А. Л.,
акад. РАН, д. т. н., проф.

Зам. главного редактора:
ИВАННИКОВ А. Д., д. т. н., проф.
ФИЛИМОНОВ Н. Б., д. т. н., с.н.с.

Редакционный совет:
БЫЧКОВ И. В., акад. РАН, д. т. н.
ЖУРАВЛЕВ Ю. И.,
акад. РАН, д. ф.-м. н., проф.
КУЛЕШОВ А. П.,
акад. РАН, д. т. н., проф.
ПОПКОВ Ю. С.,
чл.-корр. РАН, д. т. н., проф.
РУСАКОВ С. Г.,
чл.-корр. РАН, д. т. н., проф.
РЯБОВ Г. Г.,
чл.-корр. РАН, д. т. н., проф.
СОЙФЕР В. А.,
чл.-корр. РАН, д. т. н., проф.
СОКОЛОВ И. А., акад.
РАН, д. т. н., проф.
СУЕТИН Н. В., д. ф.-м. н., проф.
ЧАПЛЫГИН Ю. А.,
чл.-корр. РАН, д. т. н., проф.
ШАХНОВ В. А.,
чл.-корр. РАН, д. т. н., проф.
ШОКИН Ю. И.,
акад. РАН, д. т. н., проф.
ЮСУПОВ Р. М.,
чл.-корр. РАН, д. т. н., проф.

Редакционная коллегия:
АВДОШИН С. М., к. т. н., доц.
АНТОНОВ Б. И.
БАРСКИЙ А. Б., д. т. н., проф.
ВАСЕНИН В.А., д. ф.-м. н., проф.
ВИШНЕКОВ А. В., д. т. н., проф.
ГАЛУШКИН А.И., д. т. н., проф.
ДИМИТРИЕНКО Ю. И., д. ф.-м. н., проф.
ДОМРАЧЕВ В.Г., д. т. н., проф.
ЗАБОРОВСКИЙ В. С., д. т. н., проф.
ЗАГИДУЛЛИН Р. Ш., к. т. н., доц.
ЗАРУБИН В. С., д. т. н., проф.
КАРПЕНКО А. П., д. ф.-м. н., проф.
КОЛИН К. К., д. т. н., проф.
КУЛАГИН В. П., д. т. н., проф.
КУРЕЙЧИК В. М., д. т. н., проф.
КУХАРЕНКО Б. Г., к. ф.-м. н., доц.
ЛЬВОВИЧ Я. Е., д. т. н., проф.
МИХАЙЛОВ Б. М., д. т. н., проф.
НЕЧАЕВ В. В., к. т. н., проф.
ПОЛЕЩУК О. М., д. т. н., проф.
СОКОЛОВ Б. В., д. т. н., проф.
ТИМОНИНА Е. Е., д. т. н., проф.
УСКОВ В. Л., к. т. н. (США)
ФОМИЧЕВ В. А., д. т. н., проф.
ЧЕРМОШЕНЦЕВ С. Ф., д. т. н., проф.
ШИЛОВ В. В., к. т. н., доц.

Редакция:
БЕЗМЕНОВА М. Ю.
ГРИГОРИН-РЯБОВА Е. В.
ЛЫСЕНКО А. В.
ЧУГУНОВА А. В.

Информация о журнале доступна по сети Internet по адресу <http://novtex.ru/IT>.

Журнал включен в систему Российского индекса научного цитирования.

Журнал входит в Перечень научных журналов, в которых по рекомендации ВАК РФ должны быть опубликованы научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

CONTENTS

INTELLIGENT SYSTEMS AND TECHNOLOGIES

Kulagin V. P., Dubinin V. N. Structural Analysis of Petri Nets 3

MODELING AND OPTIMIZATION

Barskiy A. B., Zhelenkov B. V. Optimization Means for the Information Interaction of Resource Processors to Minimize the Time of "Cloud" Computing 14

Mandrikova O. V., Fetisova (Glushkova) N. V., Polozov Yu. A. Modeling of Ionospheric Parameter Time Variations and Detection of Anomalies 21

Svetushkov N. N. Integral Strings Method for Modeling Meat Loads in the Cooling Wall of the Rocket Engine Nozzle Part 32

COMPUTING SYSTEMS AND NETWORKS

Saak A. E. Circular-Typed Multiprocessor Tasks Scheduling in Grid Systems 37

Porshnev S. V., Bozhalkin D. A., Kuposov A. S. Experience of Using the Supercomputer for Processing of Traffic Dumps of Backbone Internet Channel 42

DIGITAL PROCESSING OF SIGNALS AND IMAGES

Naumov A. S. Method for Automated Reconstruction of the Fragmented Images Based on Multiple Fragment Features. 48

CRYPTOSAFETY INFORMATION

Shcheglov K. A., Shcheglov A. Yu. Informational System Attack Threat Modeling and Interpretation. Part 2. Attack Threat Modelling 54

INFORMATION TECHNOLOGIES IN BIOMEDICAL SYSTEMS

Polosin V. G., Bodin O. N. Entropy-Parametric Criterion for the Establishing the Adequacy of the Distribution Model of Ionic Currents of the Myocardium 65

Abdullayev N. T., Gasankuliyeva M. M. Informative Parameters and Fluctuation Dynamics of Electrocardiograms at Flicker-Noise Spectroscopy 72

APPLIED INFORMATION TECHNOLOGIES

Savchenko V. V. The Speech Recognition Method of Phonetic Decoding of Words with Background Noise Cancellation 76

Editor-in-Chief:

Stempkovsky A. L., Member of RAS,
Dr. Sci. (Tech.), Prof.

Deputy Editor-in-Chief:

Ivannikov A. D., Dr. Sci. (Tech.), Prof.
Filimonov N. B., Dr. Sci. (Tech.), Prof.

Chairman:

Bychkov I. V., Member of RAS,
Dr. Sci. (Tech.), Prof.
Zhuravljov Yu. I., Member of RAS,
Dr. Sci. (Phys.-Math.), Prof.
Kuleshov A. P., Member of RAS,
Dr. Sci. (Tech.), Prof.
Popkov Yu. S., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Rusakov S. G., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Ryabov G. G., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Soifer V. A., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Sokolov I. A., Member of RAS,
Dr. Sci. (Phys.-Math.), Prof.
Suetin N. V.,
Dr. Sci. (Phys.-Math.), Prof.
Chaplygin Yu. A., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Shakhnov V. A., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Shokin Yu. I., Member of RAS,
Dr. Sci. (Tech.), Prof.
Yusupov R. M., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.

Editorial Board Members:

Avdoshin S. M., Cand. Sci. (Tech.), Ass. Prof.
Antonov B. I.
Barsky A. B., Dr. Sci. (Tech.), Prof.
Vasenin V. A., Dr. Sci. (Phys.-Math.), Prof.
Vishnekov A. V., Dr. Sci. (Tech.), Prof.
Galushkin A. I., Dr. Sci. (Tech.), Prof.
Dimitrienko Yu. I., Dr. Sci. (Phys.-Math.), Prof.
Domrachev V. G., Dr. Sci. (Tech.), Prof.
Zaborovsky V. S., Dr. Sci. (Tech.), Prof.
Zagidullin R. Sh., Cand. Sci. (Tech.), Ass. Prof.
Zarubin V. S., Dr. Sci. (Tech.), Prof.
Karpenko A. P., Dr. Sci. (Phys.-Math.), Prof.
Kolin K. K., Dr. Sci. (Tech.)
Kulagin V. P., Dr. Sci. (Tech.), Prof.
Kureichik V. M., Dr. Sci. (Tech.), Prof.
Kukhareno B. G., Cand. Sci. (Phys.-Math.)
Ljvovich Ya. E., Dr. Sci. (Tech.), Prof.
Mikhailov B. M., Dr. Sci. (Tech.), Prof.
Nechaev V. V., Cand. Sci. (Tech.), Ass. Prof.
Poleschuk O. M., Dr. Sci. (Tech.), Prof.
Sokolov B. V., Dr. Sci. (Tech.)
Timonina E. E., Dr. Sci. (Tech.), Prof.
Uskov V. L. (USA), Dr. Sci. (Tech.)
Fomichev V. A., Dr. Sci. (Tech.), Prof.
Chermoshentsev S. F., Dr. Sci. (Tech.), Prof.
Shilov V. V., Cand. Sci. (Tech.), Ass. Prof.

Editors:

Bezmenova M. Yu.
Grigorin-Ryabova E. V.
Lysenko A. V.
Chugunova A. V.

Complete Internet version of the journal at site: <http://novtex.ru/IT>.

According to the decision of the Higher Certifying Commission of the Ministry of Education of Russian Federation, the journal is inscribed in "The List of the Leading Scientific Journals and Editions wherein Main Scientific Results of Theses for Doctor's or Candidate's Degrees Should Be Published"

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

INTELLIGENT SYSTEMS AND TECHNOLOGIES

УДК 519.1, 519.6

В. П. Кулагин, д-р техн. наук, профессор, руководитель научно-исследовательской лаборатории космических исследований, технологий и процессов, e-mail: vkulagin@hse.ru,

Московский институт электроники и математики

Национального исследовательского университета "Высшая школа экономики",

В. Н. Дубинин, д-р техн. наук, доц., e-mail: duvini@mail.ru,

Пензенский государственный университет

Структурный анализ сетей Петри

Предложен новый формализм — структуры сетей Петри (СП-структуры), предназначенный для исследования моделей сложных систем. Структурами, или решетками, называются частично упорядоченные множества, в которых для любых двух элементов существует точная нижняя и точная верхняя грани. В данной работе используется термин "структура". Введены операции преобразования сетей Петри (СП), предназначенные для анализа и синтеза новых СП-структур. Показано, что использование введенных операций позволяет разложить любую СП на множество линейных либо линейно-циклических фрагментов (ЛФ), которые одновременно будут относиться как к классу автоматных сетей Петри, так и к классу синхронизационных сетей Петри. Представленные операции позволяют анализировать лишь структурные свойства СП, в то время как существует ряд работ, в которых приводится анализ поведенческих свойств СП.

Описаны свойства СП-структур, введено квазиметрическое пространство структур СП. На пространстве СП-структур задано множество функций оценки, на основе которого введена система шкал, позволяющая проводить количественную оценку свойств СП-структур. Отличительной особенностью данного формализма является эффективное сочетание свойств аппарата СП с возможностями теории структур.

Практическое применение предложенных методов анализа и синтеза СП демонстрируется на классическом примере двух выталкивателей, предназначенных для транспортировки деталей. Показаны процедуры синтеза альтернативных вариантов СП-моделей системы выталкивателей, показаны их достоинства и недостатки, а также дана количественная оценка структур моделей с отображением в выбранных системах координат.

Ключевые слова: параллельные и распределенные вычисления и системы, моделирование, сложные системы, сети Петри, теория структур, системы координат

Введение

Среди большого числа понятий, которые возникли и исследуются в информатике и вычислительной технике, одним из наиболее общих является понятие сложной дискретной динамической системы. Примерами современных сложных дискретных систем могут служить компьютеры и вычислительные сети, специализированные процессоры различной структуры, программы и операционные системы, системы сбора и обработки данных, системы автоматического управления объектами и процессами, сенсорные сети, широкий спектр параллельных структур с недетерминированным поведением и др.

Изучение свойств дискретных систем и развитие методов их конструирования проводятся с помощью различных математических моделей и методов в зависимости от класса систем, степени детализации их структуры, а также от характера исследуемых проблем. Чтобы успешно решать задачи анализа и синтеза все более сложных систем, необхо-

димо дальнейшее развитие и совершенствование математических методов их исследования.

В последние годы наблюдается растущий интерес к использованию аппарата сетей Петри (СП) [1—6] для решения задач, связанных с исследованием сложных систем. Можно отметить, что современные подходы, используемые при анализе и синтезе структур сложных систем, основаны на методах декомпозиции, координации и агрегации, структурном подходе, подходе, основанном на теории сложности, и др. Однако противоречия между сложностью создаваемых современных систем и традиционными подходами к их проектированию определяют на сегодня одну из основных задач теории систем — разработка методологии и создание систем автоматизированного проектирования, решающих основные задачи синтеза. Данная проблема включает в себя разработку методов и алгоритмов, помогающих вести направленный поиск оптимальных характеристик системы, а также позволяющих контролировать изменение этих характеристик в процессе проектирования. Указанные потребности

привели к возникновению и интенсивному развитию системного подхода к проектированию сложных систем, в концепцию которого органически вписывается структурный подход, приобретающий еще большее значение на современном этапе.

В соответствии со структурным подходом выделяют следующие этапы:

- 1) выработка ряда гипотез, касающихся структур подсистем, из которых будет состоять проектируемая система;
- 2) формирование из данных подсистем законченных структур-кандидатов;
- 3) проведение анализа каждой структуры в целях определения характеристик, что позволяет сократить число структур-кандидатов и в конечном итоге выбрать окончательную структуру.

Механизм построения альтернативных вариантов на основе обобщенной модели позволяет проектировщику получить все множество возможных структур, из которых выбираются оптимальные. Однако *NP*-сложность задач получения множества альтернативных вариантов не позволяет широко использовать данный подход на практике. Кроме того, поиск наилучшего варианта структуры сопряжен с необходимостью количественной оценки каждой структуры-кандидата, а это, в свою очередь, требует наличия соответствующих математических теорий и методов.

Другим важным вопросом, требующим решения при проектировании сложных систем, является представление имеющихся данных и параллельных процессов в виде специальных формальных объектов, удобных для проведения над ними вычислительных и имитационных экспериментов. С большой долей достоверности можно отметить, что большинство современных дискретных динамических систем характеризуются такими свойствами, как параллелизм, недетерминированность, многоуровневость представления, сочетание синхронных и асинхронных процессов, однородность и др. Поэтому выбор формализованного языка, в наибольшей степени учитывающего особенности современных систем, является основной задачей начального этапа проектирования.

Преимущества СП при описании указанных сложных систем очевидны и широко описаны в литературе. Недостатком данного аппарата является отсутствие инструментария для количественной оценки моделей, выраженных в терминах СП. Вместе с тем, теория структур и теория сложности [7–9] предоставляют алгоритмы и методы, позволяющие строить оценочные шкалы для получения количественных характеристик элементов дискретного пространства.

В данной работе предлагается объединить теорию структур [8] и теорию СП в рамках *теории структур СП*. Данное сочетание позволяет в одном формализме объединить выразительную способность аппарата СП при описании распределенных систем, его возможности при исследовании структурных и

динамических свойств СП-моделей с возможностями теории структур. С практической точки зрения аппарат структур СП позволяет построить механизм количественной оценки моделей сложных систем и этим предоставить в распоряжение проектировщика инструментарий для сравнительной оценки синтезируемых СП-моделей.

Следует отметить, что современная алгебра изучает операции и отношения на множествах. Непустое множество с определенными на нем операциями и отношениями принято называть *алгебраической системой*. Простейшим примером такой системы является частично упорядоченное множество. Частично упорядоченное множество L , в котором для любых двух элементов a и b существует точная нижняя и точная верхняя грани, называется *структурой*, или *решеткой*. В данной работе используется термин "структура" в соответствии с терминологией, введенной в работе [8].

В представленной работе анализ касается статической структуры СП, в то время как, например, в работах [1, 2, 10] анализ связан с определением поведенческих свойств СП (таких как живость, безопасность, ограниченность, консервативность, повторяемость и т.д.) с использованием алгебраических средств, не предполагающих вычисление множества достижимых маркировок, но базирующихся на анализе уравнений состояния, матрицы инцидентности и т.д.

Основные определения

В сети Петри $N = (P, T, I, O, \mu_0)$ [1, 2], где P — множество позиций, T — множество переходов, $I: P \times T \rightarrow \{0, 1, 2, \dots\}$ и $O: T \times P \rightarrow \{0, 1, 2, \dots\}$ — матрицы инцидентности, μ_0 — начальная разметка, обозначим через $pre(t)$ и $post(t)$ множества входных и выходных позиций перехода t , а через $pre(p)$ и $post(p)$ — множества входных и выходных переходов позиции p . СП, в которой $P = \{p', p''\}$, $T = \{t\}$, $p' \in pre(t)$ и $p'' \in post(t)$, назовем *элементарной сетью*.

СП, состоящую из множества элементарных сетей, будем называть *примитивной системой* N_{PR} .

Для проведения анализа и синтеза сетевых моделей необходимы средства, позволяющие преобразовывать СП-модели. С этой целью рассмотрим операции разделения и объединения вершин СП.

Определение 1. Объединением переходов t_i и t_j , т. е. $(t_i + t_j)$, является переход t_{ij} , для которого справедливо:

$$pre(t_{ij}) = pre(t_i) \cup pre(t_j); post(t_{ij}) = post(t_i) \cup post(t_j).$$

Определение 2. Объединением позиций p_i и p_j , т. е. $(p_i + p_j)$, является позиция p_{ij} , для которой справедливо:

$$\begin{aligned} pre(p_{ij}) &= pre(p_i) \cup pre(p_j); \\ post(p_{ij}) &= post(p_i) \cup post(p_j); \\ \mu(p_{ij}) &= \max(\mu(p_i), \mu(p_j)). \end{aligned}$$

Делению подвергаются вершины, которые удовлетворяют следующим условиям:

$$|pre(t)| + |post(t)| > 2 \text{ и } |pre(p)| > 1 \text{ или } |post(p)| > 1.$$

Будем считать, что при определении вершин, которые должны подвергаться делению, *вес дуг не учитывается*. Рассмотрим следующие варианты операций деления позиций и переходов, которые будут использоваться при анализе СП-моделей.

Определение 3 (деление головной позиции). Если для позиции p_j выполняются условия: $pre(p_j) = \emptyset$ и $post(p_j) = \{t_{i1}, t_{i2}, \dots, t_{in}\}$, то позиция p_j делится на n позиций ($p_{j1}, p_{j2}, \dots, p_{jk}, \dots, p_{jn}$), для которых справедливо:

$$pre(p_{jk}) = \emptyset, post(p_{jk}) = \{t_{ik}\}, \mu(p_{jk}) = \mu(p_j), \text{ где } 1 \leq k \leq n.$$

Определение 4 (деление хвостовой позиции). Если для позиции p_j выполняются условия: $pre(p_j) = \{t_{i1}, t_{i2}, \dots, t_{in}\}$ и $post(p_j) = \emptyset$, то позиция p_j делится на n позиций ($p_{j1}, p_{j2}, \dots, p_{jk}, \dots, p_{jn}$), для которых справедливо:

$$pre(p_{jk}) = \{t_{ik}\}, post(p_{jk}) = \emptyset, \mu(p_{jk}) = \mu(p_j), \text{ где } 1 \leq k \leq n.$$

Определение 5 (деление позиции с одинаковым числом входных и выходных переходов). Если для позиции p_j выполняются условия: $pre(p_j) = \{t_{i1}, t_{i2}, \dots, t_{in}\}$ и $post(p_j) = \{t_{21}, t_{22}, \dots, t_{2n}\}$, то данная позиция делится на n позиций ($p_{j1}, p_{j2}, \dots, p_{jn}$), для которых справедливо:

$$pre(p_{jk}) = \{t_{1k}\}, post(p_{jk}) = \{t_{2k}\}, \text{ где } 1 \leq k \leq n.$$

Определение 6. Если для позиции p_j выполняются условия:

$$pre(p_j) = \{t_{11}, t_{12}, \dots, t_{1m}\}, post(p_j) = \{t_{21}, t_{22}, \dots, t_{2n}\} \text{ и}$$

а) $n > m$, то позиция p_j делится на позиции p'_j и p''_j , для которых справедливо:

$$pre(p'_j) = \{t_{11}, t_{12}, \dots, t_{1m}\}, post(p'_j) = \{t_{21}, t_{22}, \dots, t_{2m}\},$$

$$\mu(p'_j) = \mu(p_j),$$

$$pre(p''_j) = \emptyset, post(p''_j) = \{t_{2(m+1)}, \dots, t_{2n}\},$$

$$\mu(p''_j) = \mu(p_j);$$

б) $n < m$, то позиция p_j делится на позиции p'_j и p''_j , для которых справедливо:

$$pre(p'_j) = \{t_{11}, t_{12}, \dots, t_{1n}\}, post(p'_j) = \{t_{21}, t_{22}, \dots, t_{2n}\},$$

$$\mu(p'_j) = \mu(p_j),$$

$$pre(p''_j) = \{t_{1(n+1)}, \dots, t_{1m}\}, post(p''_j) = \emptyset,$$

$$\mu(p''_j) = \mu(p_j).$$

Можно показать, что используя введенные операции, любую СП можно разложить на множество линейных либо линейно-циклических фрагментов (ЛФ), которые одновременно будут относиться как к классу *автоматных сетей* (каждый переход сети имеет ровно одну входную и ровно одну выходную позиции), так и к классу *синхронизационных сетей* (каждая позиция имеет ровно один входной и

ровно один выходной переходы) [1]. Сети Петри, которые одновременно относятся к классу автоматных и к классу синхронизационных сетей, будем называть *автоматно-синхронизационными* сетями.

При вводе основных понятий и определений структур СП будем использовать терминологию работы [8], в которой даются основы теории структур. Обозначим через $L_m = \{N_1, N_2, \dots, N_p\}$ множество СП, которое может быть получено путем объединения вершин СП, состоящей из m элементарных сетей.

Определение 7. Под "частично упорядоченным множеством" СП будем понимать систему $X = (L_m, \Theta)$, в которой на множестве L_m определено бинарное отношение $\Theta = "\geq"$ (больше или равно), удовлетворяющее следующим свойствам:

1. Для всех $N \in L_m : \Theta(N, N)$ (рефлексивность).
2. Если $\Theta(N_i, N_j)$ и $\Theta(N_j, N_k)$, то $N_i = N_k$ (антисимметричность).
3. Если $\Theta(N_i, N_j)$ и $\Theta(N_j, N_k)$, то $\Theta(N_i, N_k)$ (транзитивность).

Будем говорить, что СП N_i и N_j структурно изоморфны ($N_i = N_j$), если $P_i = P_j, T_i = T_j, I_i = I_j$ и $O_i = O_j$.

Рассмотрим физическую сущность отношения " \geq " для множества СП. Каждому элементу N_i множества L_m поставим в соответствие некоторое число k_i , которое определяется числом вершин СП N_j . Число вершин СП N_i будем обозначать как $|N_i|$ ($k_i = |N_i|$). Тогда СП $N_i \geq N_j$, если $|N_i| \geq |N_j|$. Очевидно, что если $N_i = N_j$, то $|N_i| = |N_j|$, однако обратное несправедливо.

Учитывая принятые условия получения элементов множества L_m , свойства определения 7 можно представить в следующем виде:

1. Для всех $N \in L_m : |N| \geq |N|$ (рефлексивность).
2. Если $|N_i| \geq |N_j|$ и $|N_j| \geq |N_k|$, то $N_i = N_k$ (антисимметричность).
3. Если $|N_i| \geq |N_j|$ и $|N_j| \geq |N_k|$, то $|N_i| \geq |N_k|$ (транзитивность).

Определение 8. Выражение "СП N_i покрывает СП N_j " означает, что $N_i > N_j$ и что $N_i > N_k > N_j$ не выполняется ни для какого N_k .

Операции объединения и разделения влияют на величину $|N_i|$ путем изменения числа вершин в СП N . Отсюда следует:

Утверждение 1. СП N_i покрывает СП N_j тогда и только тогда, когда $|N_i| - |N_j| = 1$. То есть N_j может быть получена из N_i в результате выполнения только одной операции объединения.

Проверим, удовлетворяет ли наша "интерпретация" отношения " \geq " свойствам определения 7.

1. Если число вершин в СП N обозначить через k , т.е. $k = |N|$, то очевидно, что $k \geq k$. Следовательно, рефлексивность соблюдается.

2. Рассмотрим СП N_i и СП N_j , причем N_i получена из N в результате объединения двух позиций, а N_j получена из N путем объединения двух переходов. Следовательно, можно составить следующие соотношения:

$$|N_i| = |N| - 1, |N_j| = |N| - 1, |N_i| = |N_j|.$$

Отсюда $N_i \geq N_j$ и $N_j \geq N_i$, но $N_i \neq N_j$ по условиям построения СП N_i и N_j . Следовательно, свойство антисимметричности не выполняется.

3. Доказательство выполнения свойства транзитивности очевидно.

Математические системы, обладающие отношением, для которого выполняются только свойства 1 и 3 определения 7, а свойство 2 не выполняется, называются *квазиупорядоченными* [8].

Используя определение 7 и утверждение 1, можно построить графическое представление любой системы X . Для этого каждому элементу множества L_m поставим в соответствие вершину графа α_i . При этом вершины графа расположим таким образом, чтобы вершина, соответствующая СП N_i , располагалась выше вершины, соответствующей СП N_j , если $N_i > N_j$. Вершины α_i и α_j соединим направленной дугой от α_i к α_j , если N_i покрывает N_j . Фигуру, полученную таким образом, назовем *диаграммой*.

Наименьшим элементом некоторого квазиупорядоченного подмножества СП $L' \subset L_m$ будем называть такой элемент $N_o \in L'$, что $N_o \leq N_i$ для всех $N_i \in L'$. Наибольшим элементом подмножества L' будем называть такой элемент $N_l \in L'$, что $N_l \geq N_i$ для всех $N_i \in L'$.

Утверждение 2. Минимальным элементом некоторого квазиупорядоченного подмножества СП $L' \subset L_m$ является такой элемент $N_{\min} \in L'$, что ни для какого $N_i \in L'$ не имеет места $N_{\min} \geq N_i$.

Аналогично максимальным элементом подмножества L' является такой элемент N_{\max} , что ни для какого $N_i \in L'$ не имеет места $N_{\max} \leq N_i$.

Под верхней гранью k -го уровня подмножества L' квазиупорядоченного множества СП L_m будем понимать подмножество СП $L_{ki} \subset L_m$ такое, что $|N_i| = k$ для всех $N_i \in L_{ki}$ и $k > |N_j|$ для всех $N_j \in L'$. Наименьшая верхняя грань есть верхняя грань с минимальным уровнем. Под нижней гранью l -го уровня подмножества L' будем понимать подмножество СП $L_{li} \subset L_m$ такое, что $|N_i| = l$ для всех $N_i \in L_{li}$ и $l < |N_j|$ для всех $N_j \in L'$. Наибольшая нижняя грань есть нижняя грань с максимальным уровнем.

Определение 9. Структура СП есть квазиупорядоченное подмножество СП $L' \subseteq L_m$, любой элемент которого, удовлетворяющий условию $N_l > N_i > N_o$, имеет наименьшую верхнюю грань и наибольшую нижнюю грань.

Если $L' = L_m$, то структура называется *полной*.

Приведем следующую теорему.

Теорема 1. Полная структура СП обладает только одним наибольшим и только одним наименьшим элементами.

Из данной теоремы следует, что: $|N| = m + n = m + 2m = 3m$, где m — число переходов, а n — число позиций в примитивной системе; $|N_{\min}| = 1 + 1 = 2$.

В дальнейшем число элементарных сетей, которое включает максимальный элемент полной структуры, будем называть *рангом структуры*.

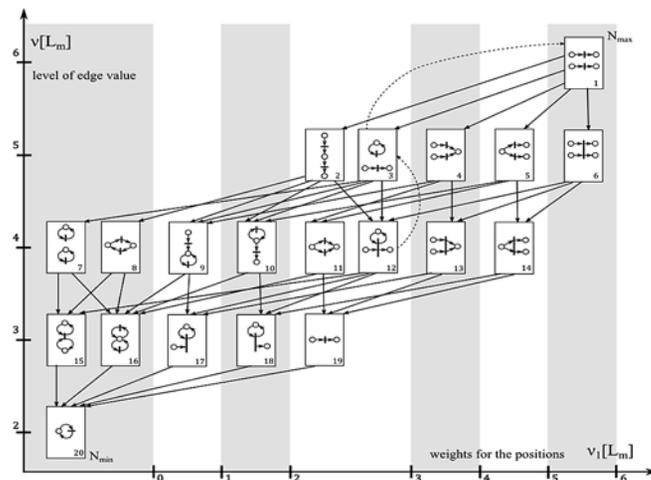


Рис. 1. Диаграмма структуры СП ранга 2:
 $v[L_m] = |N|$ (уровень грани); $v_1[L_m] = S_1 \Sigma p_i + S_2 \Sigma p_j$ (веса позиций), где Σp_i — число головных позиций; Σp_j — число хвостовых позиций; S_1, S_2 — веса головных и хвостовых позиций соответственно ($S_1 = 1, S_2 = 2$)

Рассмотрим пример. Пусть дана СП, обозначенная на рис. 1 под номером 12 (N_{12}). Построим для данной СП все возможные нижние и верхние грани. Для этого при построении элементов, принадлежащих нижним граням, будем выполнять операцию объединения, а при построении элементов, принадлежащих верхним граням, — операцию разделения. Одно из возможных преобразований N_{12} в N_{\max} (перемещение к верхним граням) показано пунктирной линией. Очевидно, что элемент, имеющий значение $k = 2$ ($|N_{20}| = 2$), является минимальным ($N_{20} = N_{\min}$), а элемент, имеющий $k = 6$ ($|N_1| = 6$), — максимальным ($N_1 = N_{\max}$). Следовательно, структура, изображенная на рис. 1, является полной. Наименьшей нижней гранью для N_1 является грань 5-го уровня, которая включает СП с N_2 по N_6 . Все данные СП могут быть получены из СП N_{\max} путем выполнения одной операции объединения.

Наибольшей нижней гранью для элемента N_{\min} является грань 3-го уровня, которая включает СП с N_{15} по N_{19} . Из рис. 1 видно, что ранг построенной структуры равен 2. Как подмножество, в данную структуру входит структура с рангом 1 (элементы N_{19} и N_{20}). Элементы данной структуры характеризуются тем, что они имеют наименьшую верхнюю грань либо наибольшую нижнюю грань. Такую структуру будем называть *вырожденной*.

Пространство структур сетей Петри

Ранее мы выделили множество СП L_m и отношение частичного порядка " \geq ". Дополнительно введем множество Q других отношений эквивалентности. Тогда пространством структур СП назовем тройку:

$$L = \langle L_m, F, Q \rangle,$$

где F — отношение частичного порядка " \geq ".

Любое из отношений эквивалентности $q \in Q$ может служить для разбиения множества L_m на смежные классы и построения некоторой шкалы путем сопоставления данных классов. Соответствие Θ , однозначное в одну сторону и удовлетворяющее условию: если $x \geq y$, то $\Theta(x) \geq \Theta(y)$, называется *изотонным*, или сохраняющим порядок [8]. Две шкалы Z и Z' будем называть *эквивалентными*, если существует такое взаимно однозначное отображение $f: Z \rightarrow Z'$, при котором из $z_i < z_j$, где $(z_i, z_j) \in Z$, следует $z'_i < z'_j$, где $(z'_i, z'_j) \in Z'$ (причем $z'_i = f(z_i)$, $z'_j = f(z_j)$) и обратно. Другими словами, эквивалентность шкал означает существование взаимно однозначного и изотонного отображения одной шкалы на другую.

На семействе смежных классов $L_m = \{L_i\}$ могут быть заданы различные функции оценки $v[L_i]$, определенные на L_m . Пусть N' , N'' и N''' — СП, принадлежащие подмножеству L_i . Оценка является *изотонной* тогда и только тогда, когда из условия $N' < N''$ следует, что $v[N'] \leq v[N'']$; *положительной* — тогда и только тогда, когда из $N' < N''$ следует, что $v[N'] < v[N'']$; *ограниченной сверху* — тогда и только тогда, когда существует конечное положительное число k_q и для всех цепей $N' < N'' < \dots < N^{(n)}$ имеет место неравенство:

$$v[N^{(n)}] - v[N'] \leq k_q.$$

Структура с изотонной оценкой $v[L_i]$ называется квазиметрической, а с положительной $v[L_i]$ — метрической [8].

Расстоянием Гливенко от $L \in L_m$ до $L' \in L_m$ произвольной квазиметрической (или метрической) структуры является функция

$$\delta(N, N') = v[N] - v[N']. \quad (1)$$

Квазиметрическое пространство (или пространство Гливенко) есть пространство, в котором может быть определена функция $\delta(N, N')$ со следующими свойствами:

- (1) $\delta(N, N) = 0$;
- (2) $\delta(N, N') > 0$ при $N' \neq N$;
- (3) $\delta(N, N') = \delta(N', N)$;
- (4) $\delta(N, N') + \delta(N', N'') \geq \delta(N, N'')$.

Очевидно, что оценка $v[L]$ является одинаковой для всех элементов $N_i \in L$, так как данная оценка является критерием для построения семейства смежных классов. Если функции оценки $v[L]$ поставить в соответствие число вершин в сетях Петри, принадлежащих некоторой структуре, и рассмотреть семейство $L_m = \{L_i\}$, то нетрудно заметить, что свойства (1)–(4) выполняются. Таким образом, пространство структур СП, разбитое на семейство смежных классов $L_m = \{L_i\}$, является пространством Гливенко.

Вернемся к диаграмме структуры СП, представленной на рис. 1. Здесь оценка $v[N] = |N|$ для всех $N \in L_m$ (при $m = 2$) разбивает множество L_m на

пять классов эквивалентности. Каждый класс объединяет множество СП с одинаковым числом вершин. Класс эквивалентности на рис. 1 отображается величиной уровня грани. Внутри каждого класса построена еще одна шкала $v_1[L_i]$, которая оценивает сумму весов головных и хвостовых позиций каждой СП $N \in L_i$, например, в соответствии со следующим выражением:

$$v_1[N] = S_1 \Sigma \bar{p}_i + S_2 \Sigma \bar{p}_j, \quad (2)$$

где $\Sigma \bar{p}_i$ — число головных позиций, $\Sigma \bar{p}_j$ — число хвостовых позиций; S_1, S_2 — веса головных и хвостовых позиций соответственно.

Видно, что оценка $v_1[L_i]$ также разбивает каждый класс эквивалентности на подклассы. Например, класс $L_3 = \{N_{15}, N_{16}, N_{17}, N_{18}, N_{19}\}$ разбит на подклассы: $L_{31} = \{N_{15}, N_{16}\}$, $L_{32} = \{N_{17}\}$, $L_{33} = \{N_{18}\}$, $L_{34} = \{N_{19}\}$.

Зададимся вопросом, можно ли считать пространство $L = \langle L_m, F, Q \rangle$ квазиметрическим, если на нем введена функция $\delta(L, L') = v_1[L] - v_1[L']$?

Если L, L' и L'' — классы эквивалентности, которые строятся в соответствии с отношением $q = v_1[L_m]$, то легко увидеть выполнимость свойств (1)–(4). Следовательно, пространство L_m , на котором введена функция $v_1[L_m]$, также является квазиметрическим.

Рассмотрим следующий вопрос: существуют ли инварианты при перемещении по шкале $v_1[L_m]$? На рис. 1 шкала $v_1[L_m]$ построена из предположения, что $S_1 = 1, S_2 = 2$. Для поиска инвариантов введем следующие веса для позиций и операций объединения вершин СП:

- если p — головная позиция, то $S(p) = S_1 = 1$;
- если p — хвостовая позиция, то $S(p) = S_2 = 2$;
- если p — внутренняя позиция, то $S(p) = S_3 = 0$;
- вес операции объединения λ позиций p_i и p_j определяется как:

$S(\lambda) = S(p_i)$ (или $S(p_j)$) — если объединяются только головные или только хвостовые позиции;
 $S(\lambda) = S(p_i) + S(p_j)$ — если объединяются разные позиции.

Рассмотрим выражение:

$$S(N) = S_1 \bar{n} + S_2 \bar{n} + \sum_{i=1}^w S(\lambda_i), \quad (3)$$

где \bar{n} — число головных позиций, \bar{n} — число хвостовых позиций в СП; w — число операций объединения, с помощью которых построена анализируемая СП; $S(N)$ — вес СП N .

Для всех СП $N_i \in L_m$, изображенных на рис. 1, вес СП, найденный в соответствии с выражением (3), является величиной постоянной и равен утроенному рангу структуры СП $3r$. Следовательно, величина $S(N)$ также является инвариантом.

На практике часто встречаются задачи, при решении которых необходимо определить, какие СП могут быть построены из ограниченного, наперед

заданного числа элементарных сетей. Попробуем ответить на данный вопрос.

Из свойств операции разделения переходов вытекает, что число элементов минимальной примитивной системы, которое получается в результате разделения переходов СП-структуры N , определяется по правилу:

$$\tilde{m} = \sum_{i=1}^{m'} \max(\text{pre}(t_i), \text{post}(t_i)), \quad (4)$$

где m' — число переходов в СП N ; \tilde{m} — число переходов в минимальной примитивной системе N_{PR} .

Справедлива следующая теорема.

Теорема 2. Для того чтобы СП N принадлежала множеству L_m , достаточно, чтобы размерность примитивной системы, построенной для N , не превышала ранга СП-структуры L_m .

Алгоритмы оценки интерпретирующих СП имеют экспоненциальную зависимость от числа вершин СП. Поэтому при анализе СП актуальным становится вопрос сокращения числа вершин.

Сокращение числа вершин СП (*редукция*) должно проводиться таким образом, чтобы язык, порождаемый измененной СП, являлся изоморфным языку исходной СП. Поэтому задача сокращения вершин СП в нашем случае ставится следующим образом.

Определить множество преобразований $G = \{G_1, G_2, \dots, G_n\}$ таким образом, чтобы выполнялись следующие условия:

- применение любого G_i сохраняет изоморфизм языков СП N и N' ;
- для любой СП N существует конечная (возможно пустая) последовательность преобразований такая, что к результирующей сети никакое преобразование из G неприменимо.

Обозначим через G^* множество конечных последовательностей элементов из G . Если $\alpha = G_i G_j \dots G_k$, то $N \xrightarrow{\alpha} N'$ означает, что сеть N' является результатом применения к сети N последовательности преобразований α . Сеть назовем несократимой, если к ней никакое $G_i \in G$ неприменимо.

Опишем некоторые преобразования G_i , входящие в множество G .

Преобразование G_1 (удаление эквивалентного перехода).

Если в СП имеется несколько эквивалентных переходов, то может быть оставлен только один из них.

Преобразование G_2 (удаление перехода-петли).

Если в СП имеется два перехода t_1 и t_2 такие, что: $\text{pre}(t_2) = \text{post}(t_2)$, $\text{pre}(t_2) \subseteq \text{pre}(t_1)$ (или $\text{post}(t_2) \subseteq \text{post}(t_1)$), то переход t_2 может быть удален.

Преобразование G_3 (удаление позиции-петли).

Если в СП имеется позиция p , для которой $\text{pre}(p) = \text{post}(p)$, то она может быть удалена.

Преобразование G_4 (удаление эквивалентных позиций).

Если в СП имеется несколько эквивалентных позиций, то может быть оставлена только одна из них.

Рассмотренные преобразования обладают свойствами конечности и однозначности.

Конечность. Множество преобразований G действует конечно, если для любой СП N найдется последовательность $\alpha \in G^*$ (возможно пустая) такая, что к СП N' , полученной из N ($N \xrightarrow{\alpha} N'$), никакое преобразование из G неприменимо.

Однозначность. Множество преобразований G действует однозначно, если несократимые сети N' и N'' , полученные в результате применения к СП N двух различных последовательностей α и β , оказываются структурно изоморфными.

Свойство однозначности позволяет нам не заботиться о порядке применения преобразований G_i .

Пример: система двух выталкивателей

В качестве иллюстративного примера рассмотрим систему из двух выталкивателей, предназначенную для транспортировки деталей (рис. 2). Данная система использовалась в ряде работ в качестве эталонного примера [11]. Как видно из рис. 2, траектория движения детали проходит через ряд характерных точек 1, 2 и 3. Исполнительный механизм системы позволяет выдвигать поршень выталкивателя и втягивать его. Для этого используется два управляющих сигнала — "Вперед" и "Назад". Каждый из выталкивателей снабжен двумя датчиками, сигнализирующими о нахождении поршня выталкивателя в одном из двух крайних положений — "Втянут" и "Выдвинут". В представленной системе двух выталкивателей имеется одна аварийная ситуация, когда оба выталкивателя находятся в выдвинутом состоянии (рис. 3).

СП-модель N_1 , описывающая систему двух выталкивателей, представлена на рис. 4. СП-модель учитывает, что выталкиватель может быть только в двух (крайних) состояниях "Втянут" и "Выдвинут". Ниже в табл. 1 и 2 дана интерпретация позиций и переходов, представленной на рис. 4 СП-модели.

Проведем анализ представленной СП-модели с точки зрения корректности работы (отсутствие тупиков, позиций с неограниченной разметкой, а также отсутствие состояний, при котором возникает аварийная ситуация). С этой целью разложим СП-модель N_1 на множество *автоматно-синхронизационных* сетей (множество линейных и линейно-

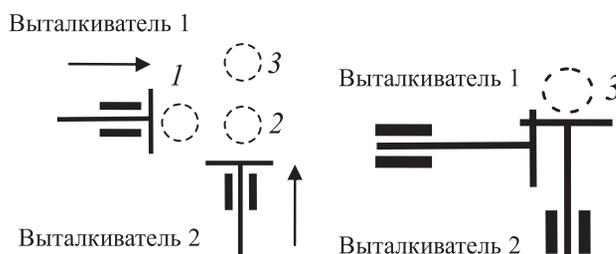


Рис. 2. Исходная ситуация (оба поршня втянуты)

Рис. 3. Аварийная ситуация (оба поршня выдвинуты)

Таблица 1

Имя позиции	Интерпретация позиции
$s<i>1$	i -й контроллер находится в начальном состоянии (состоянии ожидания прихода детали)
$s<i>2$	i -й контроллер находится в состоянии ожидания выдвижения i -го поршня ($i \in \{1,2\}$)
$s<i>3$	i -й контроллер находится в состоянии ожидания втягивания i -го поршня ($i \in \{1,2\}$)
$p<i>1$	i -й поршень втянут (исходное состояние) ($i \in \{1,2\}$)
$p<i>2$	i -й поршень выдвинут ($i \in \{1,2\}$)
$q<i>1$	i -м контроллером была выдана команда на выдвижение i -го поршня ($i \in \{1,2\}$)
$q<i>2$	i -м контроллером была выдана команда на втягивание i -го поршня ($i \in \{1,2\}$)
d1	Деталь находится в положении 1
d2	Деталь находится в положении 2
d3	Деталь находится в положении 3

Таблица 2

Имя перехода	Интерпретация перехода
$ts<i>1$	i -й контроллер выдал команду на выдвижение i -го поршня ($i \in \{1,2\}$)
$ts<i>2$	i -й контроллер выдал команду на втягивание i -го поршня ($i \in \{1,2\}$)
$ts<i>3$	i -й контроллер перешел в начальное состояние ($i \in \{1,2\}$)
$tr<i>1$	i -й поршень выдвинулся ($i \in \{1,2\}$)
$tr<i>2$	i -й поршень втянулся ($i \in \{1,2\}$)
td	После ухода транспортированной детали в системе появилась новая деталь

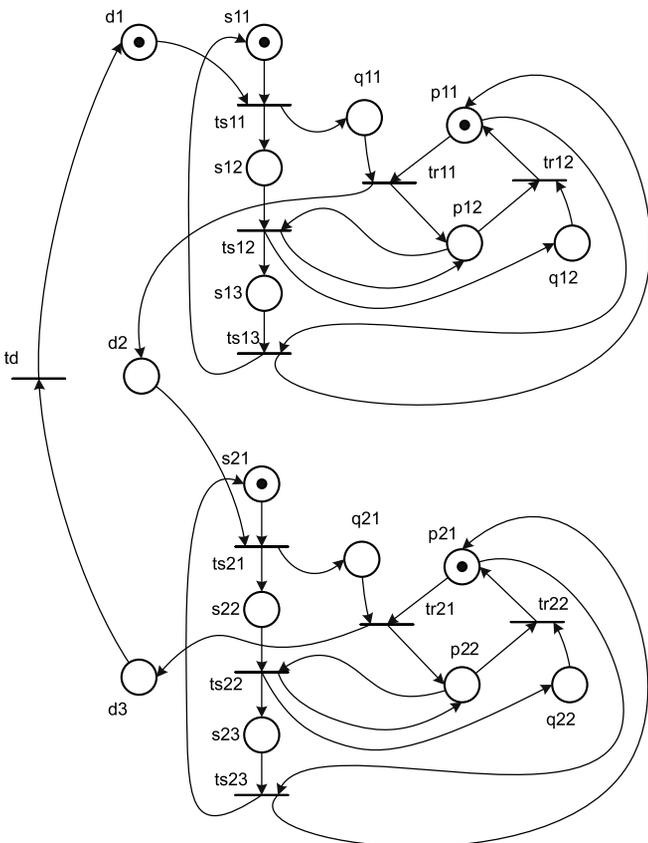


Рис. 4. СП-модель $M1$ системы двух выталкивателей

циклических фрагментов (ЛФ)), используя операции разделения вершин СП, описанных выше.

Предположим, что в результате деления вершин СП получена система ЛФ, представленная на рис. 5—8. На основе выражения (5) можно опреде-

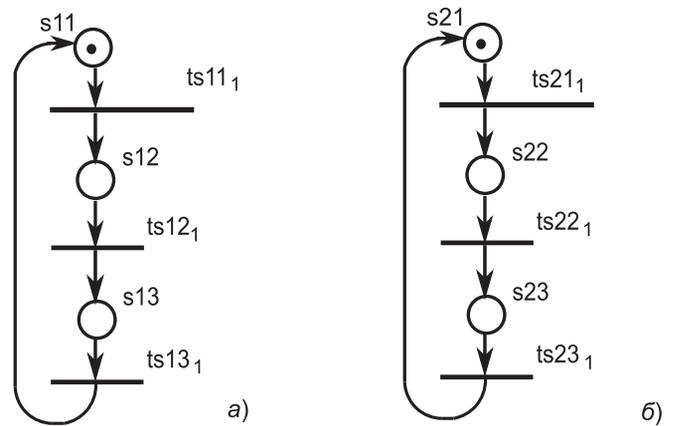


Рис. 5. Сетевые модели работы контроллеров

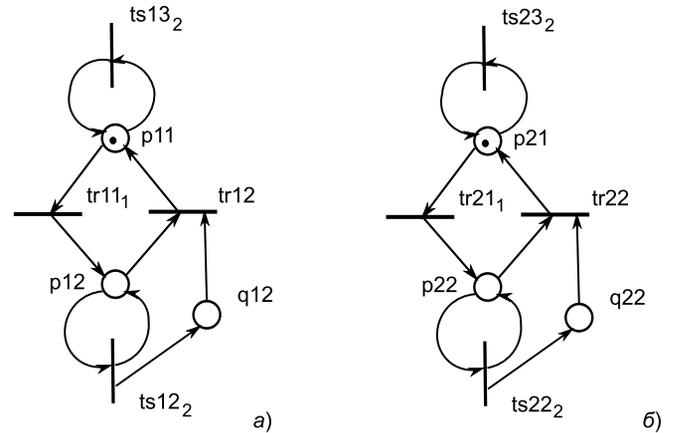


Рис. 6. Сетевые модели управления движением выталкивателей

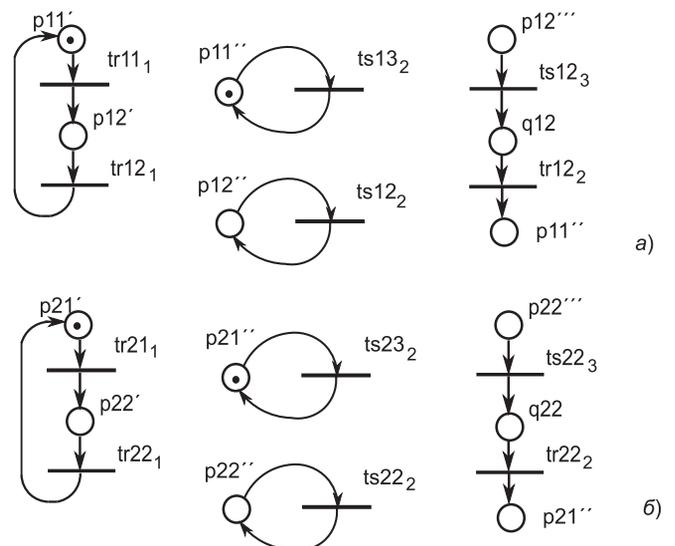


Рис. 7. Линейные и линейно-циклические фрагменты, образующие СП-модель управления движением выталкивателей

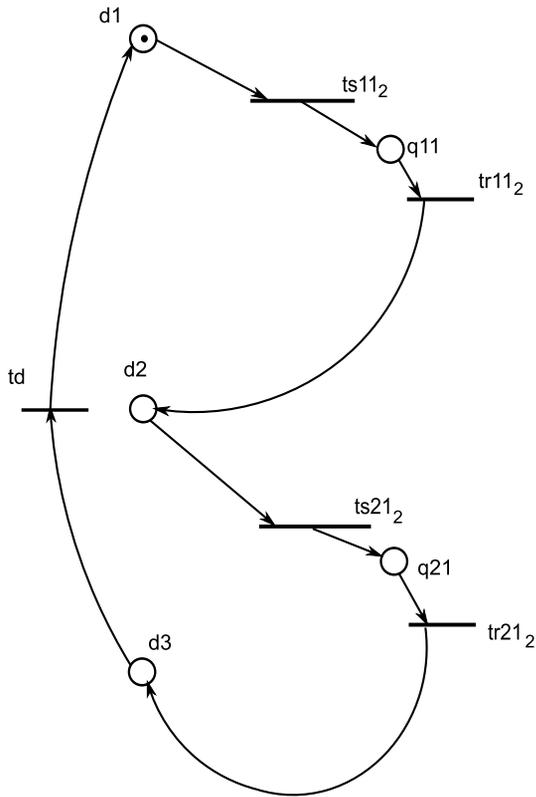


Рис. 8. Сетевая модель работы системы управления выдвиганием поршней

лечь, что для построения СП N_1 , представленной на рис. 4, необходимо 23 элементарные сети, что подтверждается видом ЛФ (рис. 5, 7, 8).

Программа синтеза СП-структуры N_1 на основе полученной системы ЛФ имеет следующий вид:

$$\begin{aligned}
 ts11 &= ts11_1 + ts11_2 & tr12 &= tr12_1 + tr12_2 \\
 ts12 &= ts12_1 + ts12_2 + ts12_3 & tr21 &= tr21_1 + tr21_2 \\
 ts13 &= ts13_1 + ts13_2 & tr22 &= tr22_1 + tr22_2 \\
 ts21 &= ts21_1 + ts21_2 & p11 &= p11' + p11'' + p11''' \\
 ts22 &= ts22_1 + ts22_2 + ts22_3 & p12 &= p12' + p12'' + p12''' \\
 ts23 &= ts23_1 + ts23_2 & p21 &= p21' + p21'' + p21''' \\
 tr11 &= tr11_1 + tr11_2 & p22 &= p22' + p22'' + p22'''
 \end{aligned}$$

Анализ СП-модели N_1 с использованием пакетов моделирующих программ *ViVe* [12] и *PIPE* [13] позволяет отметить, что данная СП-модель является живой и ограниченной (свойства *хорошо сформированной* сети). Однако СП-модель имеет недостаток, который заключается в том, что при срабатывании последовательности переходов $\sigma = ts11, tr11, ts21, tr21$ возникает разметка, соответствующая недопустимому состоянию системы двух выталкивателей (два поршня выдвинулись одновременно). Необходимо модифицировать исходную СП-модель N_1 таким образом, чтобы исключить описанное запрещенное состояние, но сохранить при этом свойства *хорошо сформированной* сети. С этой целью в СП-

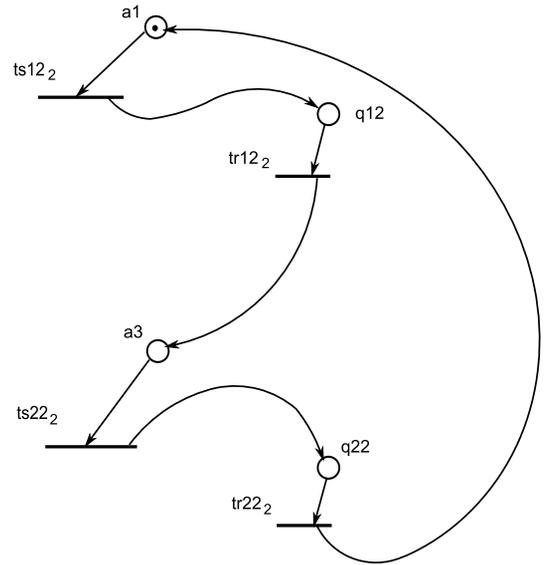


Рис. 9. Сетевая модель арбитра

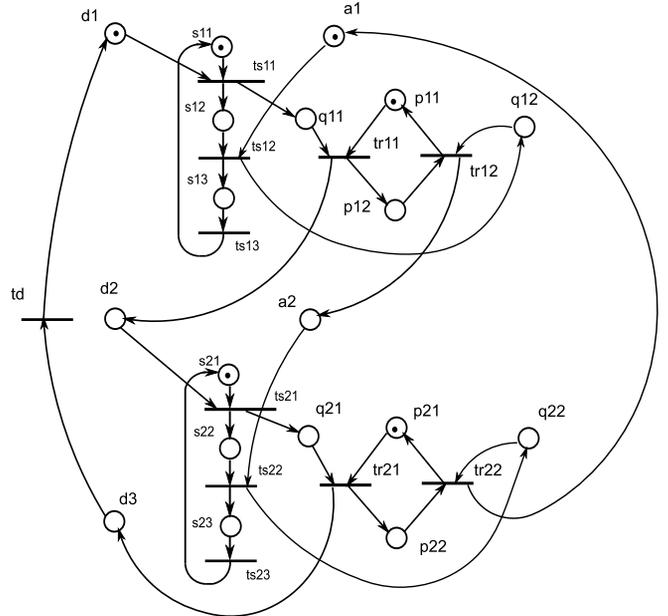


Рис. 10. СП-модель N_2 системы двух выталкивателей с арбитражем

модель N_1 предлагается включить арбитра (рис. 9) который позволяет выдвинуться первому толкателю только в том случае, если второй толкатель вдвинут, и наоборот. Анализ полученной СП N_2 (рис. 10) показывает, что сеть также хорошо сформирована и не содержит разметки, соответствующей запрещенному состоянию системы двух выталкивателей. Однако размерность примитивной системы (ранг структуры), необходимой для синтеза N_2 , превышает размерность системы для синтеза СП N_1 . Действительно, если для синтеза СП N_1 требуется 23 элементарные сети, то для синтеза СП N_2 необходимо 27 элементарных сетей ($r_{N_2} > r_{N_1}$).

Можно ли построить СП, моделирующую систему двух выталкивателей, используя меньшее число элементарных сетей? Разложим СП N_2 на систему ЛФ, которые представлены на рис. 5, 8, 9, 11.

Рассмотрим, возможно ли применить преобразования множества G над вершинами ЛФ СП N_2 ? Следует заметить, что преобразование $G2$ допустимо. Удалим из системы ЛФ СП N_2 циклические фрагменты $p11'' - ts13_2$, $p12'' - ts12_2$, $p21'' - ts23_2$ и $p22'' - ts22_2$. Построим новую СП N_3 путем выполнения представленной ниже программы синтеза на основе измененной системы ЛФ.

Программа синтеза СП N_3 :

$$\begin{aligned} ts11 &= ts11_1 + ts11_2 & tr12 &= tr12_1 + tr12_2 \\ ts12 &= ts12_1 + ts12_2 & tr21 &= tr21_1 + tr21_2 \\ ts13 &= ts13_1 & tr22 &= tr22_1 + tr22_2 \\ ts21 &= ts21_1 + ts21_2 & p11 &= p11' \\ ts22 &= ts22_1 + ts22_2 & p12 &= p12' \\ ts23 &= ts23_1 & p21 &= p21' \\ tr11 &= tr11_1 + tr11_2 & p22 &= p22' \end{aligned}$$

Полученная СП N_3 представлена на рис. 12. Анализ данной СП показывает, что она является хорошо сформированной и включает модель арбитра, запрещающего одновременное выдвигание двух выталкивателей. Кроме этого, для синтеза данной СП-модели необходимо всего лишь 19 элементарных сетей.

В целях последующего сравнительного анализа СП N_1 , N_2 и N_3 перейдем к единому рангу и возьмем за основу максимальный ранг, который характеризует СП N_2 ($r_2 = 27$).

Для количественной оценки полученных СП воспользуемся шкалами $v[L]$ и $v_1[L]$, описанными выше. На рис. 13, а приведена система координат в выбранных шкалах, в которой представлено положение СП N_1 , N_2 и N_3 . Видно, что отсутствие головных и хвостовых позиций исключило влияние шкалы $v[L]$ (характеризующие показатели СП по этой шкале равны 0). В целях более эффективной оценки свойств СП N_1 , N_2 , N_3 введем новые шкалы $v_2[L]$, $v_3[L]$ и $v_4[L]$, которые оценивают СП по числу операций объединения в целом ($v_4[L]$), а также по числу объединения позиций ($v_2[L]$) и переходов ($v_3[L]$) элементарных сетей при построении отдельных СП. Можно показать, что для оценок $v_2[L]$, $v_3[L]$ и $v_4[L]$ свойства (1)–(4) выполняются.

Для определения характеристических значений для N_1 , N_2 и N_3 в соответствии со шкалами $v_2[L]$, $v_3[L]$ и $v_4[L]$ воспользуемся соответствующими системами ЛФ.

Как было уже отмечено, для СП-модели N_1 ЛФ представлены на рис. 5, 7, 8. Очевидно, что для построения фрагментов, представленных на рис. 5, необходимо выполнить по три операции объединения позиций элементарных сетей $ts11_1$, $ts12_1$, $ts13_1$ и $ts21_1$, $ts22_1$ и $ts23_1$. Аналогично для построения фрагментов, представленных на рис. 7, а, б,

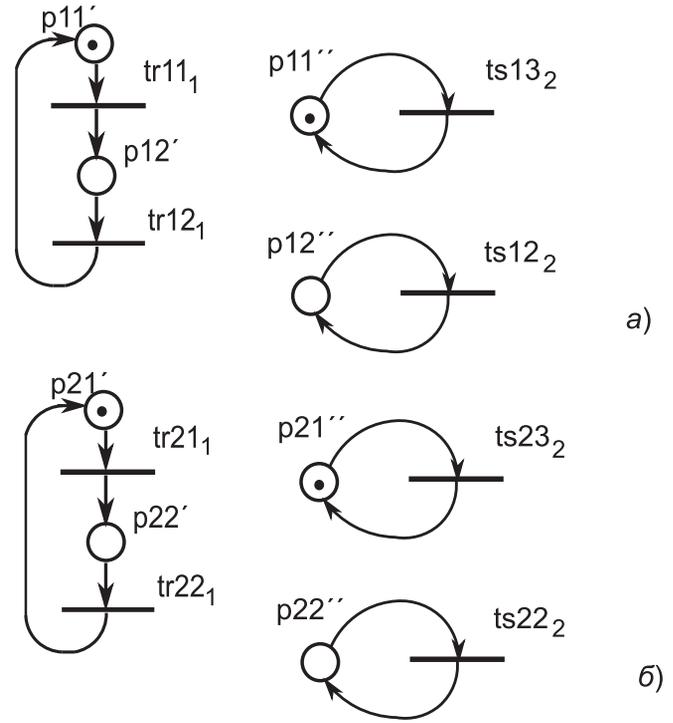


Рис. 11. Линейно-циклические фрагменты, образующие СП-модель управления движением поршней, для СП-модели N_2

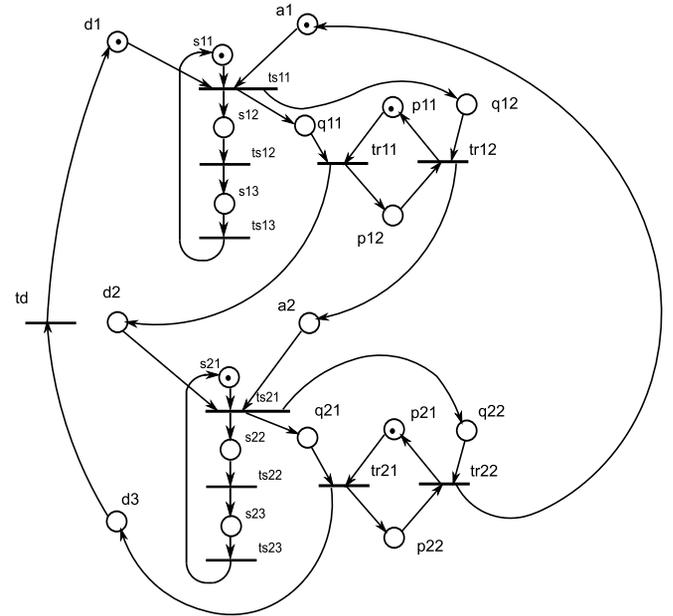


Рис. 12. Сетевая модель N_3 системы двух выталкивателей

необходимо выполнить в общей сложности 10 операций объединения позиций (5 для фрагмента (а) и 5 для фрагмента (б)), а для фрагмента, представленного на рис. 8, необходимо выполнить 5 операций объединения позиций. СП N_1 синтезируется в результате выполнения представленной выше программы, при выполнении которой необходимо выполнить 12 операций объединения переходов и 8 дополнительных операций по объединению пози-

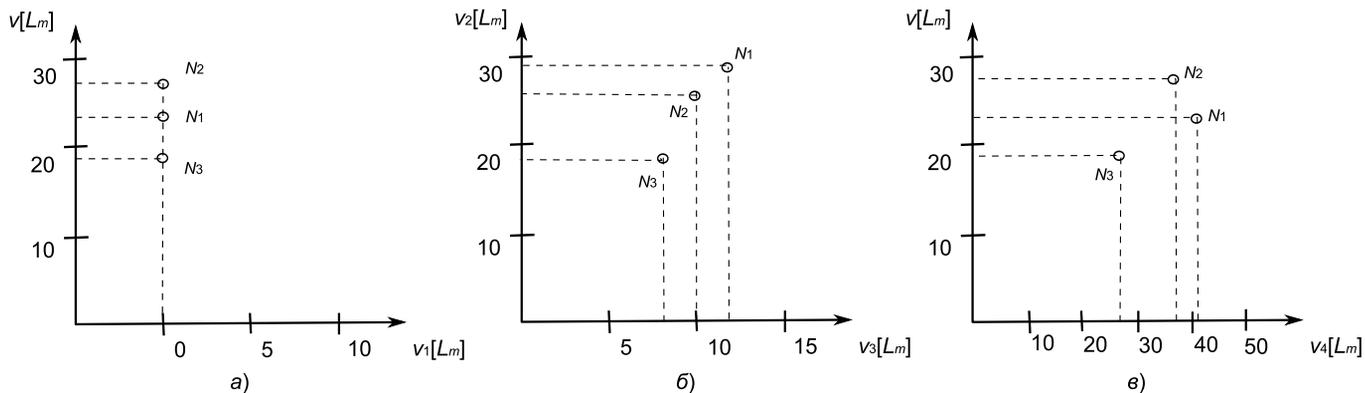


Рис.13. Примеры сравнительных оценок СП N_1, N_2, N_3 , представленных в различных системах координат

ций. В итоге для N_1 имеем: $v_2[N_1] = 29, v_3[N_1] = 12, v_4[N_1] = 41$.

Для СП-модели N_2 ЛФ представлены на рис. 5, 8, 9, 11. Рассуждая аналогично, что и при оценке СП N_1 , можно отметить, что число операций объединения позиций для синтеза фрагментов, представленных на рис. 5 и 8, остается 6 и 5 соответственно. Для построения фрагментов, представленных на рис. 9 и рис. 11, необходимо выполнить 4 и 8 операций объединения позиций соответственно. СП N_2 строится в результате выполнения следующей программы объединения позиций и переходов ЛФ.

Программа синтеза СП N_2 на основе системы ЛФ:

$$\begin{aligned}
 ts11 &= ts11_1 + ts11_2 & tr12 &= tr12_1 + tr12_2 \\
 ts12 &= ts12_1 + ts12_2 & tr21 &= tr21_1 + tr21_2 \\
 ts13 &= ts13_1 + ts13_2 & tr22 &= tr22_1 + tr22_2 \\
 ts21 &= ts21_1 + ts21_2 & p11 &= p11' + p11'' \\
 ts22 &= ts22_1 + ts22_2 & p12 &= p12' + p12'' \\
 ts23 &= ts23_1 + ts23_2 & p21 &= p21' + p21'' \\
 tr11 &= tr11_1 + tr11_2 & p22 &= p22' + p22''
 \end{aligned}$$

В результате выполнения данной программы необходимо выполнить 10 операций объединения переходов и 4 дополнительные операции по объединению позиций. В итоге для N_2 имеем:

$$v_2[N_2] = 27, v_3[N_2] = 10, v_4[N_2] = 37.$$

Проведем анализ СП-модели N_3 . ЛФ, используемые для построения N_3 , представлены на рис. 5, 8, 9, 11. Число операций объединения позиций для синтеза фрагментов, представленных на рис. 5, 8 и 9, остается 6, 5 и 4 соответственно. Для построения фрагментов, представленных на рис. 11, необходи-

мо выполнить 4 операции объединения позиций (при синтезе N_3 мы не учитываем циклические фрагменты $r11'' - ts13_2, p12'' - ts12_2, p21'' - ts23_2$ и $p22'' - ts22_2$).

В результате синтеза СП N_3 на основе элементарных сетей необходимо выполнить 8 операций объединения переходов и 19 операций по объединению позиций. В итоге для N_3 имеем:

$$v_2[N_3] = 19, v_3[N_3] = 8, v_4[N_3] = 27.$$

Полученные численные значения СП N_1, N_2, N_3 сведены в табл. 3.

Использование шкал при оценке структурных особенностей СП имеет существенный практический смысл. Например, анализируя положение СП в системе координат $v_2[L], v_3[L]$ (рис. 13, б), можно дать оценку уровню распараллеливания процессов, протекающих в данной модели, а положение СП в системе координат, представленной на рис. 13, в, характеризует трудоемкость синтеза модели, которая, с одной стороны, зависит от числа используемых элементов (элементарных сетей) $v[L]$, а с другой — от числа выполненных операций объединения при синтезе СП-модели $v_4[L]$.

Заключение

В данной работе предложен новый формализм для исследования СП-моделей сложных систем — структуры СП. Описаны некоторые свойства СП-структур, введено квазиметрическое пространство структур СП. На пространстве СП-структур задано множество функций оценки, на основе которого введена система шкал, позволяющая проводить количественную оценку синтезируемых СП. Отличительной особенностью данного формализма является эффективное сочетание свойств аппарата СП с возможностями теории структур, которые показаны на примере.

Практическая значимость предложенного формализма заключается в том, что в виде системы шкал предоставляется инструмент количественной оценки и сравнительного анализа синтезируемых СП-моделей при исследовании сложных систем.

Таблица 3

СП-структура	$v[L]$	$v_1[L]$	$v_2[L]$	$v_3[L]$	$v_4[L] = v_2[L] + v_3[L]$
N_1	23	0	29	12	41
N_2	27	0	27	10	37
N_3	19	0	19	8	27

Работа выполнена в рамках гранта Российского научного фонда "Проведение фундаментальных научных исследований и поисковых научных исследований с привлечением молодых исследователей" (проект № 15-11-10010).

Список литературы

1. Котов В. Е. Сети Петри. М.: Наука, 1984.
2. Питерсон Дж. Теория сетей Петри и моделирование систем. М.: Мир, 1984.
3. Кулагин В. П. Проблемы анализа и синтеза структур параллельных вычислительных систем // Информационные технологии. 1997. № 1. С. 2—8.
4. 32nd International Conference on Application and Theory of Petri Nets and Concurrency, 2011. URL: <http://www.jaist.ac.jp/acsd-petrinets2011/>
5. 33rd International Conference on Application and Theory of Petri Nets and Concurrency (Petri Nets 2012), Hamburg, Germany,

June 25—29, 2012. URL: <http://www.informatik.uni-hamburg.de/TGI/PetriNets/meetings/pn2012/>

6. 34th International Conference on Application and Theory of Petri Nets and Concurrency (Petri Nets 2013), Milano, Italy, June 24—28, 2013. URL: <http://www.mc3.disco.unimib.it/petrinets2013/>
7. Анкундинов Г. И. Синтез структуры сложных объектов. Логико-комбинаторный подход. Л.: ЛГУ, 1986.
8. Биркгоф Г. Теория структур. М.: Иностранная литература, 1952.
9. Скорняков Л. А. Элементы теории структур. М.: Наука, 1982.
10. Cabasino M. P., Giua A., Seatzu N. Structural analysis of Petri nets // Control of Discrete-Event Systems. Lecture Notes in Computer Science. London: Springer-Verlag, 2013. № 433. P. 213—233.
11. Дубинин В. Н. Об одной методике проектирования supervizornogo upravleniia dlia diskretno-sobytiinykh sistem // Известия высших учебных заведений. Поволжский регион. Технические науки. 2009. № 3. С. 12—24.
12. ViVe — VisualVerifier Tool Framework. URL: <http://www.fb61499.com/license.html>
13. PIPE: Platform Independent Petri net Editor 2. URL: <http://pipe2.sourceforge.net/>

V. P. Kulagin, Head of the Research Laboratory, Professor, e-mail: vkulagin@hse.ru,
National Research University Higher School of Economics,
V. N. Dubinin, Associate Professor, Penza state University

Structural Analysis of Petri Nets

In this article we propose a new formalism — the structure of Petri nets (PN-structure) designed to study the models of complex systems. Structures or lattices are called partially ordered set in which any two elements there is a precise accurate lower and upper faces. In this paper we use the term "structure". Typed operations transformation Petri nets (PN) designed for the analysis and synthesis of new JV structures. It is shown that the introduced operations, allows us to decompose any SP into many linear or linear-cyclic fragments (LF), which simultaneously will be treated as a class of automata Petri nets and class synchronization Petri nets. Presents operations only allow you to analyze structural properties of the JV, while there are a number of studies in which the analysis of behavioral properties of the PN.

Describes the properties of PN-structures, introduced quasimatrices space structures PN. In the space of the PN-structures entities set multiple functions of evaluation, which introduced a system of scales, allowing a quantification of the properties of PN-structures. A distinctive feature of this formalism is an effective combination of properties of the device in PN with the possibilities of the theory of structures.

The practical application of the proposed methods of analysis and synthesis of PN is demonstrated by the classic example of two ejectors designed to transport parts. Showing the procedure of the synthesis of alternative options PN-models of the system of ejectors, showing their advantages and disadvantages, as well as a quantitative assessment of model structures with the display in the selected coordinate systems.

Keywords: parallel and distributed computing and systems, modeling, complex systems, Petri nets, theory of structures, system of coordinates

References

1. Kotov V. E. *Seti Petri*. Moscow: Nauka, 1984.
2. Peterson Dzh. *Teoriia setei' Petri i modelirovanie sistem*. Moscow: Mir, 1984.
3. Kulagin V. P. Problemy' analiza i sinteza struktur paralel'ny'kh vy'chislitel'ny'kh sistem, *Informatsionny'e tekhnologii*, 1997, no. 1, pp. 2—8.
4. 32nd International Conference on Application and Theory of Petri Nets and Concurrency, 2011. URL: <http://www.jaist.ac.jp/acsd-petrinets2011/>
5. 33rd International Conference on Application and Theory of Petri Nets and Concurrency (Petri Nets 2012), Hamburg, Germany, June 25—29, 2012. URL: <http://www.informatik.uni-hamburg.de/TGI/PetriNets/meetings/pn2012/>
6. 34th International Conference on Application and Theory of Petri Nets and Concurrency (Petri Nets 2013), Milano, Italy, June 24—28, 2013. URL: <http://www.mc3.disco.unimib.it/petrinets2013/>

7. Ankundinov G. I. *Sintez struktury' slozhny'kh ob'ektov. Logiko-kombinatorny'i' podhod*. Leningrad: LGU, 1986.

8. Birkhof G. *Teoriia struktur*. Moscow: Inostrannaia literatura, 1952.

9. Skorniakov L. A. *E'lementy' teorii struktur*. Moscow: Nauka, 1982.

10. Cabasino M. P., Giua A., Seatzu N. Structural analysis of Petri nets, *Control of Discrete-Event Systems. Lecture Notes in Computer Science*. London: Springer-Verlag, 2013, no. 433, pp. 213—233.

11. Dubinin V. N. Ob odnoi' metodike proektirovaniia supervizornogo upravleniia dlia diskretno-sobytiiny'kh sistem, *Izvestiia vy'sshikh uchebny'kh zavedenii'*. Povolzhskii' region. Tekhnicheskie nauki, 2009, no. 3, pp. 12—24.

12. ViVe — *VisualVerifier Tool Framework*. URL: <http://www.fb61499.com/license.html>

13. PIPE: *Platform Independent Petri net Editor 2*. URL: <http://pipe2.sourceforge.net/>

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ MODELING AND OPTIMIZATION

УДК 004.031.43

А. Б. Барский, д-р техн. наук, проф., e-mail: arkbarsk@mail.ru,
Б. В. Желенков, канд. техн. наук, доц., e-mail: boriszhv@gmail.com
МИИТ, Москва

Средства оптимизации информационного взаимодействия ресурсных процессоров для минимизации времени "облачных" вычислений

Исследуются возможности повышения скорости обмена данными при распределенных вычислениях в компьютерных сетях. Предлагается параллельный многоканальный прием данных одним компьютером от многих абонентов с помощью внешнего устройства, объединяющего несколько web-объектов. Применение такого устройства актуально при кластеризации ресурса для сборки результатов вычислений на головном компьютере при Grid- и "облачных" вычислениях по SPMD-технологии. Приводятся формулы оптимизации ресурсных параметров вычислительного процесса. Обсуждаются особенности закрепления "физических" ресурсов за виртуальными в глобальной сети. Утверждается, что значительное географическое удаление ресурсных процессоров друг от друга приводит к большим затратам времени на синхронизацию обмена, делая неэффективным применение многих web-объектов ввода. Рекомендуется компактное размещение ресурсных процессоров и применение локальной сети для распределенных вычислений. Показывается возможность реализации конвейера обработки потока поступающих на компьютер данных для перевода кадров из сетевого стандарта в компьютерное представление.

Ключевые слова: компьютерная сеть, "облачные" вычисления, виртуальный ресурс, сборка данных, кластеризация, конвейер операций

Введение

"Облачные" вычисления следует рассматривать в свете практической реализации Grid-технологии [1–3]. Эта технология предполагает наличие развитой сети Центров, объединенных и доступных пользователю через глобальную сеть (web-сеть), обладающих мощными вычислительными ресурсами для распределенного решения задач. Часть ресурсов, являющихся web-объектами, могут быть удаленными относительно Центра, а также быть общими для многих Центров. Использование абстрактных, виртуальных ресурсов всего доступного земного информационного пространства для оптимизации его производительной загрузки является главной идеей Grid- и "облачных" вычислений. Следует учесть и то, что в будущем сложные системы управления, в частности военного назначения, целесообразно реализовывать на основе практически неуязвимых виртуальных вычислительных ресурсов. Поэтому в общем случае, выделяя наиболее трудную часть проблемы, следует считать, что Центр на основе запросов пользователей либо сам организует решение задач на удаленных сетевых ресурсах, либо предоставляет эти ресурсы пользователю для самостоятельного планирования вычислительного процесса. В любом случае взаимодействие выделенных средств и их связь с пользователем или с головным процессором Центра, управляющим вычислитель-

ным процессом решения задачи, осуществляется на основе web-технологий.

Однако низкая скорость взаимодействия объектов сети превращает организацию распределенных вычислений в сложную оптимизационную задачу обоснования состава вычислительных средств. В частности, может оказаться, что распределение вычислений неэффективно: один компьютер справится быстрее, чем их медленно взаимодействующий коллектив.

Установлено, что основной, базовой, эффективной технологией "облачного" решения задач [4] является SPMD-технология (*Single Program — Multiple Data*). При организации вычислений по этой технологии основная проблема взаимодействия вычислительных средств в сети заключается в сборке результатов вычислений на одном процессоре для совместного анализа и выдачи пользователю. Для этого применяется принцип кластеризации [1, 3, 5, 6] ресурсных процессоров¹. Повышается актуальность разработки аппаратных и структурных средств поддержки параллельной, независимой передачи данных "от многих — одному" для минимизации времени вычислений.

¹ В соответствии с принятой терминологией, здесь, в отличие от кластера web-объектов, значение слова "кластер" соответствует объединению нескольких компьютеров для сбора результатов их счета на одном из них.

1. Общая схема "облачных" вычислений по SPMD-технологии и функции процессоров Grid-Центра

Организация "облачных" вычислений на базе SPMD-технологии отображена укрупненной схемой на рис. 1–4.

Очевидно, если пользователь запрашивает только ресурс, самостоятельно рассчитывая способ его использования, например, при моделировании сложной структурированной системы, то на него ложатся и хлопоты по достижению высокой производительности. В том числе это касается организации "быстрого" взаимодействия распределенных программных модулей. Однако если распределенное решение задачи полностью возлагается на Центр Grid-технологий с применением аппарата "облачных" вычислений, то минимизация времени решения становится главной функцией Центра. Минимум времени решения достигается оптимальным выбором числа ресурсных процессоров, способом кластеризации ресурсных процессоров при сборе результатов счета, а также средствами аппаратной и структурной поддержки распределенного вычислительного процесса.

2. Кластер web-объектов — внешнее устройство памяти хост-процессора

В работе [3] предложено в качестве промежуточного устройства памяти для сборки результатов счета ресурсных процессоров использовать модули оперативной памяти (ОП). В каждом таком модуле может формироваться область ОП ресурсного процессора, хранящая результаты вычислений.

За одним модулем ОП при решении задачи могут закрепляться несколько ресурсных процессоров. Таким образом, со стороны хост-процессора кластер таких модулей ОП входит в состав единого адресного пространства его оперативной памяти, разработанной по традиционной технологии, использующей матричный коммутатор. Со стороны web-пространства, т. е. со стороны ресурсных процессоров, каждый такой модуль ОП является web-объектом, имеющим свой адрес в глобальной сети. Ресурсные процессоры взаимодействуют с хост-процессором посредством закрепленных за ними модулей ОП — web-объектов построенного кластера.

Однако технология построения оперативной памяти, работающей в двух слабосовместимых направлениях, весьма сложна (хотя известны технологии построения многоходовой памяти). Обращение к ней не исключает конфликтов, разрешаемых арбитрами матричного коммутатора. Трудно достигается достаточный объем такой памяти.

Целесообразно считать, что модули, хранящие образы памяти ресурсных процессоров, выполнены по технологии построения внешних устройств, например дисковых. Типом данных для них становится файл [3], что соответствует управлению со стороны



Рис. 1. Алгоритм управления компьютера — администратора Центра



Рис. 2. Алгоритм управления, реализуемый хост-процессором

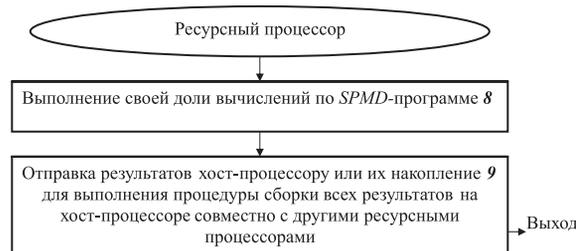


Рис. 3. Алгоритм работы ресурсного процессора

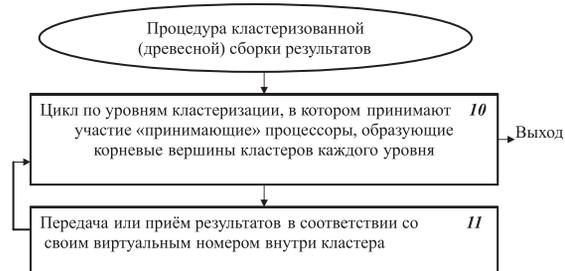


Рис. 4. Схема совместно выполняемой процедуры кластеризации

хост-процессора с помощью файловой подсистемы ОС. Кластер web-объектов как единое внешнее устройство может быть связан с хост-процессором одним каналом ввода-вывода при последовательном обслуживании web-объектов. Последовательное считывание результатов из образов памяти групп ресурсных процессоров, поступивших на web-объекты, окончательно формирует файл результатов, предоставляемый пользователю.

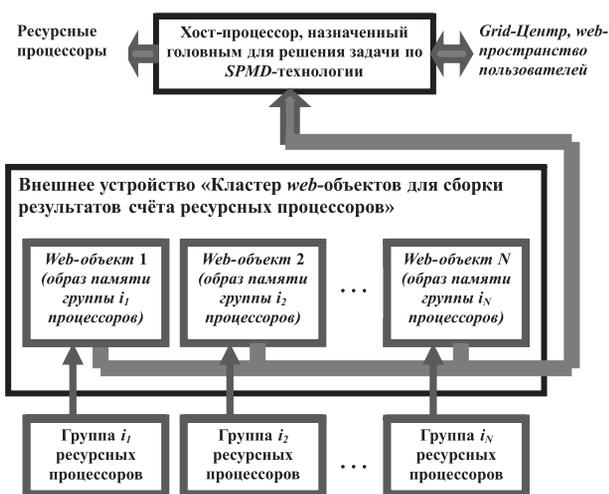


Рис. 5. Структура связей кластера web-объектов для сборки данных

На рис. 5 указаны связи, обусловленные сборкой результатов счета ресурсных процессоров. Однако связь в обратном направлении — от хост-процессора к ресурсным процессорам, обозначенная на рисунке отдельно, также может быть реализована через web-объекты кластера, если ресурсные процессоры распределены между ними.

Число M ресурсных процессоров может значительно превышать число N web-объектов в составе кластера, и за каждым web-объектом в общем случае может быть закреплено неодинаковое число ресурсных процессоров. Ведь возможности технической реализации весьма ограничены по сравнению с возможностями абстрактного виртуального выбора числа процессоров для распределенных вычислений. То есть, если $M > N$, однозначное закрепление web-объектов за ресурсными процессорами невозможно. Необходима кластеризация этих процессоров для объединения результатов их счета по "древесной" структуре [2, 4, 5]. На рис. 5 название (индекс) группы процессоров совпадает с номером ресурсного процессора, образующего корневую вершину дерева. Таким образом, на этих процессорах частично выполнена сборка результатов счета внутри группы, а окончательная сборка выполняется с помощью кластера web-объектов.

3. Оптимизация ресурсных параметров вычислительного процесса

Выбор оптимальной схемы вычислительного процесса заключается в таком подборе соотношений параметров, при котором достигается минимальное время решения задачи пользователя.

Этими параметрами являются:

N — число процессоров в кластере web-объектов;

M — число выделенных в задаче ресурсных процессоров;

n — размерность задачи или число решаемых вариантов.

Необходимо считать, что значение N находится на уровне построения системы обработки информации. Другие параметры выбирают для организации вычислительного процесса решения конкретной задачи. Так, согласно SPMD-технологии, параметр n определяется способом распараллеливания по информации [1, 2, 4, 8]. Он соответствует числу параллельных запусков программы (*single program*) для выполнения одного из вариантов обработки экспериментальных данных, нахождения корня системы линейных уравнений, варианта продолжения цепочки логического вывода и т. д. Число $M \leq n$ и состав ресурсных процессоров назначается компьютером — администратором Центра для минимизации или с учетом ограничения времени решения задачи.

Для определенности будем считать, что *single program* составлена для обработки одного варианта исходных данных и распределяется между однородными ресурсными процессорами в соответствии с распределением вариантов счета.

На рис. 6 представлена схема решения задачи при $N = 4$, $n = 48$, $M = 24$, отображающая использование выделенных ресурсных процессоров. Рассмотрена двухпроцессорная [7] схема кластеризации процессоров для сборки данных на внешнем устройстве памяти хост-процессора. Каждая группа, связанная с одним web-объектом, содержит шесть процессоров. Сборка внутри каждой группы отражена на рис. 6 снизу вверх, некоторые "собирающие" процессоры на третьем уровне указаны повторно.

Пусть $t_{\text{вар}}$ — время счета одного варианта; $t_{\text{пер}}$ — время выполнения сеанса передачи данных в сети; $t_{\text{вн}}$ — время считывания с одного web-объекта в оперативную память хост-процессора. Тогда время $T_{\text{реш}}$ распределенного решения задачи без учета подготовительных операций будет:

$$T_{\text{реш}} = \left\lceil \frac{n}{M} \right\rceil \left[t_{\text{вар}} + \left(\left\lceil \log_2 \frac{M}{N} \right\rceil + 1 \right) t_{\text{пер}} + N t_{\text{вн}} \right] \quad (1)$$

Построенная функция является ступенчатой. Для приближенной оценки оптимального соотношения

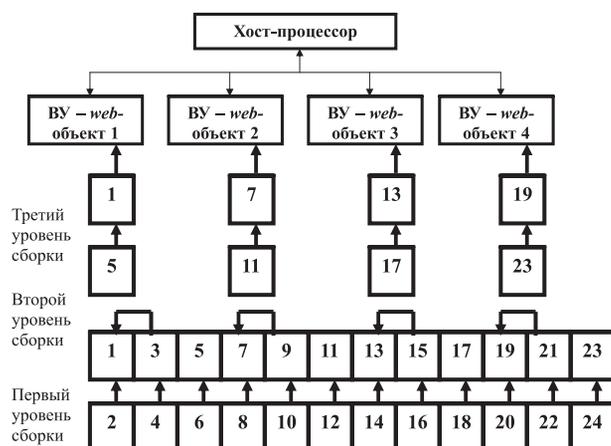


Рис. 6. Схема решения задачи с применением двухпроцессорной кластеризации для сборки результатов счета

шения между параметрами, а именно, для выбора значения M при назначении однородного ресурса администратором Центра, представим эту функцию непрерывной от этой переменной, отбросив требование целочисленности коэффициентов:

$$T_{\text{реш}} = \frac{nt_{\text{вар}}}{M} + (\log_2 M - \log_2 N + 1)t_{\text{пер}} + Nt_{\text{вн}}. \quad (2)$$

При этом считается, что $n \geq M \geq N$.

Найдем первую производную (2) по M и приравняем ее нулю:

$$T'_{\text{реш}} \approx -\frac{nt_{\text{вар}}}{M^2} + \frac{t_{\text{пер}}}{M \ln 2} = \frac{-nt_{\text{вар}} \ln 2 + Mt_{\text{пер}}}{M^2 \ln 2} = 0. \quad (3)$$

Следовательно, экстремальное целое значение M определяется соотношением

$$M \approx 0,7n \frac{t_{\text{вар}}}{t_{\text{пер}}}. \quad (4)$$

Для проверки соответствия экстремальной точки (4) минимальному значению времени решения $T_{\text{реш}}$ найдем вторую производную (2):

$$T''_{\text{реш}} = \frac{2nt_{\text{вар}}}{M^3} - \frac{t_{\text{пер}}}{M^2 \ln 2} = \frac{1,4nt_{\text{вар}} - t_{\text{пер}}M}{M^3 0,7}.$$

В экстремальной точке $T''_{\text{реш}} = \frac{nt_{\text{вар}}}{M^3} > 0$. Положительное значение второй производной указывает на то, что в найденной по соотношению (4) экстремальной точке M достигается минимум времени решения задачи, так как первая производная меняет знак с "минуса" на "плюс".

Следует обратить внимание на то, что найденная рекомендация (4) администратору Центра по выбору числа однородных ресурсных процессоров на основе размерности задачи не зависит от наличия кластера *web*-объектов в составе внешних устройств хост-процессора. Весьма важным является отношение времени счета одного варианта к времени обмена данными в глобальной сети. Число *web*-объектов в соответствии с формулой (1) уменьшает время распределенного решения задачи.

На рис. 7 показана схема решения задачи при использовании трехпроцессорной кластеризации.

Как и ранее, каждая группа ресурсных процессоров, приходящаяся на один *web*-объект и внутри которой проводится "древесная" сборка, содержит шесть процессоров.

Трехпроцессорная кластеризация, обеспечивающая меньшее число уровней сборки результатов, целесообразна лишь в том случае, когда технически возможно совмещение двух синхронных пересылок одному из трех процессоров так, чтобы время сборки $t_{\text{сб}}$ данных на нем стало существенно меньше времени двух независимых, выполненных раздельно, пересылок: $t_{\text{сб}} < 2t_{\text{пер}}$.

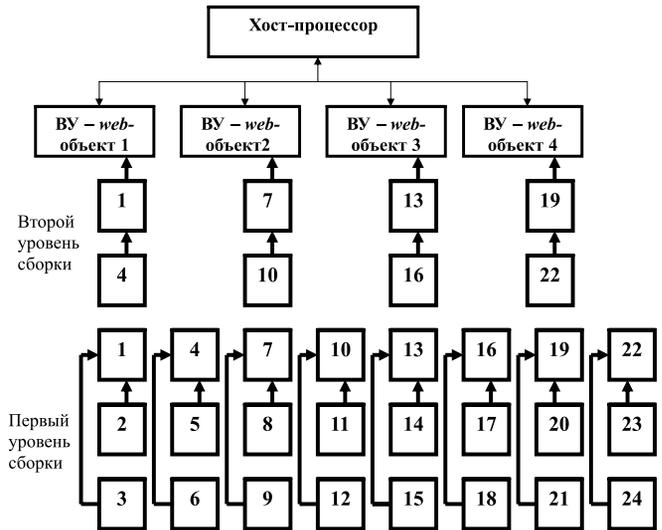


Рис. 7. Схема решения задачи при трехпроцессорной кластеризации

Время решения задачи находим следующим образом:

$$T_{\text{реш}} = \left] \frac{n}{M} [t_{\text{вар}} + \right] \log_3 \frac{M}{N} [t_{\text{сб}} + t_{\text{пер}} + Nt_{\text{вн}}. \quad (5)$$

Для получения приближенных оценок преобразуем ступенчатую функцию в непрерывную:

$$T_{\text{реш}} = \frac{nt_{\text{вар}}}{M^2} + (\log_3 M - \log_3 N)t_{\text{сб}} + t_{\text{пер}} + Nt_{\text{вн}}. \quad (6)$$

Найдем первую производную по M и приравняем ее нулю:

$$T'_{\text{реш}} = -\frac{nt_{\text{вар}}}{M} + \frac{t_{\text{сб}}}{M \ln 3} = \frac{-nt_{\text{вар}} \ln 3 + Mt_{\text{сб}}}{M^2 \ln 3} = 0. \quad (7)$$

Экстремальная точка приближенно определяется соотношением

$$M \approx 1,1n \frac{t_{\text{вар}}}{t_{\text{сб}}}. \quad (8)$$

Положительное значение второй производной в этой точке свидетельствует о достижении в ней минимума функции $T_{\text{реш}}$.

Пусть отсутствует техническая возможность хотя бы частичного совмещения двух пересылок данных одному процессору при трехпроцессорной кластеризации. Тогда $t_{\text{сб}} = 2t_{\text{пер}}$, и выбор трехпроцессорного способа кластеризации возможен при условии

$$2 \left] \log_3 \frac{M}{N} [< \left] \log_2 \frac{M}{N} [. \quad (9)$$

4. Проблемы практической реализации оптимизирующих средств передачи и сбора данных "облачных" вычислений

Внедрение распределенных вычислений в качестве одной из основных функций в такую развитую глобальную сеть, какой является Интернет, не может

не коснуться основ ее построения [8–10]. Прежде всего, это касается скорости информационного обмена между *web*-объектами, инициатором которого является не человек-пользователь, а компьютер или его внешнее устройство, вычислительная система, специализированное устройство связи и т.д. То есть внутри процесса решения задачи пользователя обмен является полностью автоматическим.

В то же время тенденция развития современной вычислительно-информационной среды направлена на расширение использования терминальных систем. Это обусловлено такими факторами, как:

- высокая степень интеграции схем на уровне аппаратной реализации средств вычислительной техники;
- высокая степень универсальности;
- высокая пропускная способность телекоммуникационных сетей.

Повышение степени интеграции схем на уровне аппаратной реализации средств вычислительной техники на начальном этапе привело к уходу от терминальных систем, давая возможность пользователю максимально абстрагироваться от внешних вычислительных ресурсов и выполнять решение задач локально. Но по мере увеличения степени универсализма вычислительных средств (компьютеров), расширился круг решаемых задач и увеличились объемы вычислений. Этот факт в сочетании с постоянным ростом пропускной способности телекоммуникационных сетей привел к новому витку развития терминальных систем. Высокая скорость передачи информации и практически неограниченная зона покрытия современных телекоммуникационных сетей в сочетании с компьютерами компакт-класса, позволяют использовать распределенные вычислительные ресурсы в сети общего доступа (в "облаке"), отправляя в него задачи и по-

лучая ответ. При этом в "облаке" создается набор виртуальных вычислительных ресурсов в зависимости от требований решаемой задачи и ограниченный времени получения результата.

Нет необходимости покупать дорогостоящий высокопроизводительный компьютер, а достаточно купить недорогое мобильное терминальное устройство с минимальными аппаратными возможностями и при необходимости арендовать виртуальные вычислительные ресурсы в "облаке".

Если раньше время решения задачи определялось системным временем, необходимым для решения задачи на однопроцессорном или мультипроцессорном компьютере, то теперь это время определяется временем решения задачи в "облаке", в основном зависящим от времени взаимодействия и синхронизации обмена общими данными.

В первых двух разделах уже описано взаимодействие объектов в такой вычислительной среде. Также рассмотрен процесс оптимизации ресурсных параметров вычислительного процесса при двухпроцессорной и трехпроцессорной кластеризации.

В общем случае физическое размещение виртуальных ресурсов зависит от администратора сети, закрепляющего "физические" процессоры за виртуальными. Тогда физические ресурсы, назначенные для поддержки одного вычислительного процесса, могут находиться в различных географически удаленных друг от друга частях "облака". Следовательно, результаты выполнения частных задач будут приходиться на хост-процессор настолько не одновременно, что кластеризация сборки потеряет смысл. Ведь на всех уровнях сборки потребуются синхронизация ожидания поступления всех необходимых результатов. В этом случае сборка решений может осуществляться последовательно, для чего достаточно одного внешнего *web*-объекта, как показано на рис. 8 для двухпроцессорной кластеризации. Тогда время $T_{\text{реш}}$ распределенного решения задачи без учета подготовительных операций можно найти по формуле

$$T_{\text{реш}} = \frac{nt_{\text{вар}}}{M} + (\log_2 M + 1)t_{\text{пер}} + t_{\text{вн}} \quad (10)$$

Рассмотрим ситуацию, противоположную предыдущей и допускающую практически синхронный обмен данными в сети, без специального применения средств синхронизации.

Допустим, что все компьютеры, предоставляющие свои вычислительные ресурсы, находятся в одной локальной сети или компактно и близко расположены на местности.

Тогда задания поступают всем компьютерам практически одновременно, и при приблизительно равных объемах распределяемых работ результаты также будут приходиться на хост-процессор практически одновременно. Это позволит ускорить процесс сборки за счет объединения в кластеры нескольких виртуальных машин при условии, что они будут запущены на разных компьютерах. Число

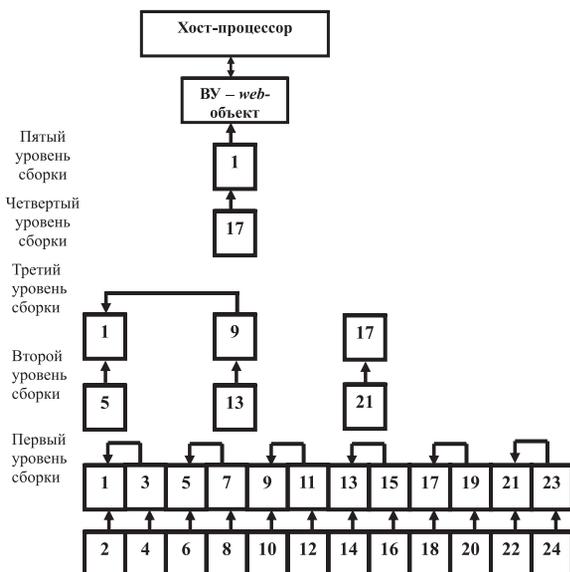


Рис. 8. Схема решения задачи с одним *web*-объектом приема данных при сильно рассогласованном, асинхронном обмене в сети

машин, которые можно объединить в кластер, будет зависеть от числа интерфейсов $k \geq 3$, которые можно установить на компьютер.

В таком случае выражение (5) примет следующий вид:

$$T_{\text{реш}} = \left\lceil \frac{n}{M} \left[t_{\text{вар}} + \right] \log_{k-1} \frac{M}{N} \left[t_{\text{сб}} + t_{\text{пер}} + N t_{\text{вн}} \right] \right\rceil \quad (11)$$

Из этого выражения можно выделить функцию определения числа L уровней сборки:

$$L = \left\lceil \log_{k-1} \frac{M}{N} \right\rceil. \quad (12)$$

На основе выражений (11) и (12) можно найти необходимое число процессоров в кластере или необходимое число интерфейсов k , как функцию от L .

5. Постановка задачи конвейеризации многоканального приема данных *web*-объектом от группы абонентов

Пропускная способность сетей связи и передачи данных сегодня достигает значительного объема. Однако в интересах стандартизации и унификации аппаратно-программных средств обмена между *web*-объектами, включая интерфейс пользователя и взаимодействие задач, используется пакетно-кадровая передача. Тогда, с учетом применения программных средств, реальная пропускная способность средств передачи данных в сети снижается на один-два порядка. При решении информационных задач, таких как социальные сети, торговые площадки, электронная почта, справочные системы и других на первый план выдвигаются требования высокого сервиса. Применяемых сегодня алгоритмов, дисциплин и протоколов обмена пока вполне достаточно для приемлемой пропускной способности сети.

Распределенные вычисления организуются как для структуризации имитационного моделирования, так и для минимизации времени вычислений за счет распараллеливания.

Для взаимодействия с пользователем и управления вычислительным процессом один из ресурсных процессоров пользователя является головным. Особую важность обретает поток, образуемый сборкой результатов вычислений или синхронизирующих сообщений на головном процессоре, поступающий от других ресурсных процессоров. Этот поток на принимающем процессоре образует "узкое горло", которое необходимо "расшить", используя принципы распараллеливания и конвейеризации.

Аналогичная проблема возникает при кластеризации, проводимой для реализации "древесной" структуры промежуточной сборки данных, например, при трех- (и выше) процессорной кластеризации ресурсных процессоров.

Распараллеливание приема данных рассмотрено ранее на основе введения нескольких внешних

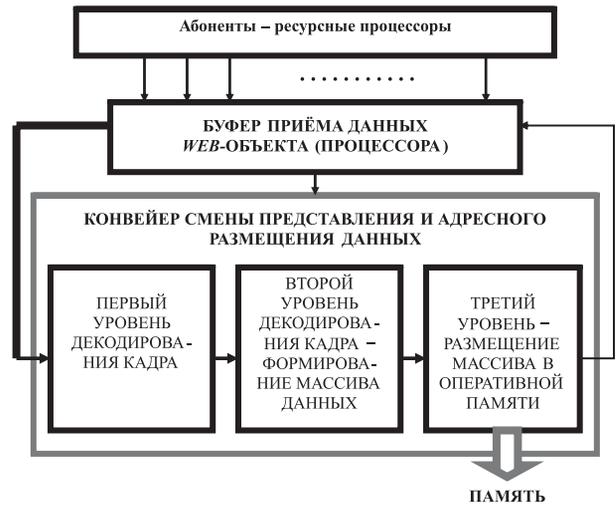


Рис. 9. Конвейеризация приема и размещения данных *web*-объекта

устройств обмена, каждый из которых является *web*-объектом, т.е. имеет адрес и достаточную самостоятельность для выполнения сетевых операций обмена данными.

Однако второй важный резерв ускорения обмена заключается в применении конвейерной обработки потока данных, поступающих на *web*-объект, перевода их стандартного представления в сети передачи данных в стандартное представление в памяти процессора. Такая обработка актуальна при использовании локальной вычислительной сети в качестве ресурса.

На рис. 9 первый и второй уровни переводят данные со стандартного, *кадрового* представления в линии передачи данных в тип *файл*. Третий уровень окончательно переводит данные в представление на уровне задачи, т.е. в тип *массив*, и выполняет адресное размещение массива в оперативной памяти.

Заключение

Слишком много случайных факторов налагаются друг на друга при совместном распределенном решении задачи в вычислительной сети: случайный разброс моментов времени получения заданий разными процессорами; случайный разброс времени выполнения заданий процессорами; случайный разброс времени начала асинхронной передачи результатов счета в соответствии с кластеризацией процессоров; случайный разброс времени асинхронной сборки результатов на одном процессоре; случайный разброс времени асинхронного обмена между уровнями кластеризации и, наконец, случайный разброс времени окончательного формирования вектора результатов на головном процессоре.

Тем не менее несмотря на возложение все более высокой вычислительной нагрузки на виртуальные процессоры компьютерных сетей, рост производительности этих сетей и пропускной способности сетей передачи данных, совершенствование про-

цедур обмена данными при компактном выборе ресурсов, процесс сборки данных будет обретать черты синхронности. Он будет следовать тем идеальным временным параметрам, которые соответствуют предполагаемой древесной структуре сборки. В этом случае ускорение совместной обработки потока принимаемых данных на одном *web*-объекте обретает особую актуальность.

В качестве аналогии следует обратить внимание на утвердившуюся практику построения специализированных (по операциям) скалярных и векторных конвейерных исполнительных устройств (ИУ) в составе АЛУ суперкомпьютеров. Применение скалярных конвейерных устройств, на которых время выполнения отдельных операций даже увеличивается, эффективно тогда, когда соответствующие устройству операции в программе образуют поток. В этом потоке следующая операция того же типа с большой вероятностью поступает на выполнение, когда выполнение предыдущей операции (этого типа) еще не закончено. Чем выше эта вероятность, т.е. плотность потока следования данной операции при выполнении программы, тем выше эффективность конвейерного скалярного ИУ при решении конкретной задачи. Чтобы максимально увеличить эту плотность, проводят векторизацию задач и используют синхронные векторно-конвейерные ИУ, способные в каждом такте выдавать результаты операции после "разгона" конвейера.

Такой же эффект для "пиковой" нагрузки следует ожидать при использовании многих *web*-объектов обмена и программно-аппаратной конвейерной поддержки кластеризации сборки данных в компьютерной сети.

Для более точной оценки эффективности мер ускорения обмена необходимо провести стохастиче-

ское моделирование информационных потоков на детерминированной модели глобальной и локальной компьютерной сети.

Список литературы

1. Барский А. Б. Параллельные информационные технологии. М.: ИНТУИТ; БИНОМ. Лаборатория знаний, 2007. 503 с.
2. Барский А. Б. Grid-вычисления. Организация, методы, планирование. Saarbrücken, Germany: LAP LAMBERT Academic Publishing, 2012. 349 с.
3. Барский А. Б. Алгоритмические, архитектурные и структурные методы организации управляющих процессов в виртуальном пространстве средств Grid-системы // Информационные технологии. 2012. № 5. С. 2—6.
4. Барский А. Б. Параллельные технологии решения оптимизационных задач // Информационные технологии. Приложение. 2001. № 2. 24 с.
5. Амиршахи Б. Кластеризация GRID-ресурсов для оптимизации информационного обмена при совместной обработке результатов распределенных вычислений // Информационные технологии. 2011. № 2. С. 22—28.
6. Амиршахи Б. GRID-технологии решения больших систем линейных уравнений на вычислительной сети и на суперкомпьютере кластерного типа // Информационные технологии. 2011. № 6. С. 17—22.
7. Барский А. Б. Двух- и трехкомпьютерная кластеризация Grid-ресурсов для оптимизации сборки результатов распределенных вычислений по SPMD-технологии // Информационные технологии. 2012. № 8. С. 43—46.
8. Загорский Г. С., Барский А. Б., Желенков Б. В. Grid-технологии — ресурсы супервозможностей // Мир Транспорта. 2008. № 3. С. 132—141.
9. Барский А. Б. Оптимизационные задачи в основе пакета параллельных прикладных программ и системы информационного обслуживания Центра Grid-технологий. // Информационные технологии. Приложение. 2010. № 10. 32 с.
10. Goldovsky Y. M., Zhgelenkov B. V., Safonova I. E., Cyganova N. A. To the question of evaluating the reliability of functioning of the elements of the telecommunication enterprise-wide networks // International Scientific — Practical Conference "Innovative Information Technologies". Part 2. Innovative Information Technologies in Science. Prague, 2014, April 21—25. P. 474—477.

A. B. Barskiy, Professor, e-mail: arkbarsk@mail.ru,

B. V. Zhelenkov, Associated Professor, e-mail: boriszhv@gmail.com,
Moscow State University of Railway Engineering (MIIT)

Optimization Means for the Information Interaction of Resource Processors to Minimize the Time of "Cloud" Computing

The possibilities of increasing the data rate when distributed computing in computer networks. The parallel multichannel reception of data from one computer to many subscribers via an external device that combines multiple web-sites. The use of such a device is important when clustering resource for assembling the results of calculations on the host computer when Grid- and "cloud" computing on the SPMD-technology. Formulas resource optimization parameters of the computational process. The features of fixing the "physical" resources for virtual WAN. It is argued that a significant geographical remoteness resource processors from each other leads to long time to sync exchange, making inefficient use of many web-input objects. Recommended compact arrangement of processors and application resource network for distributed computing. We show the feasibility of processing pipeline flow coming to a computer to transfer data frames from the network standard in the computer representation.

Keywords: computer network "cloud" computing, virtual resource, the assembly of data, clustering, pipeline operations

References

1. **Barskij A. B.** *Parallel'nye informacionnye tehnologii*, Moscow, INTUIT, BINOM. Laboratorija znaniy, 2007, 503 p.
2. **Barskij A. B.** *Grid-vychislenija: organizacija, metody, planirivanie*, Saarbrücken, Germany, LAP LAMBERT Academic Publishing, 2012. 349 p.
3. **Barskij A. B.** Algoritmicheskie, arhitekturnye i strukturnye metody organizacii upravljajushhijh processov v virtual'nom prostranstve sredstv Grid-sistemy, *Informacionnye tehnologii*, 2012, no. 5, pp. 2—6.
4. **Barskij A. B.** Parallel'nye tehnologii reshenija optimizacionnyh zadach, *Informacionnye tehnologii, Prilozhenie*, 2001, no. 2. 24 p.
5. **Amirshahi B.** Klasterizacija GRID-resursov dlja optimizacii informacionnogo obmena pri sovmestnoj obrabotke rezul'tatov raspredelyonnyh vychislenij, *Informacionnye tehnologii*, 2011, no. 2, pp. 22—28.
6. **Amirshahi B.** GRID-tehnologii reshenija bol'shijh sistem linejnyh uravnenij na vychislitel'noj seti i na superkomp'yutere klasterного типа, *Informacionnye tehnologii*, 2011, no. 6, pp. 17—22.
7. **Barskij A. B.** Dvuh- i tryohkomp'yuternaja klasterizacija Grid-resursov dlja optimizacii sborki rezul'tatov raspredelyonnyh vychislenij po SPMD-tehnologii, *Informacionnye tehnologii*, 2012, no. 8, pp. 43—46.
8. **Zagorskij G. S., Barskij A. B., Zhelenkov B. V.** Grid-tehnologii — resursy supervozmozhnostej, *Mir Transporta*, 2008, no. 3, pp. 132—141.
9. **Barskij A. B.** Optimizacionnye zadachi v osnove paketa parallel'nyh prikladnyh programm i sistemy informacionnogo obsluzhivaniya Centra Grid-tehnologij, *Informacionnye tehnologii. Prilozhenie*, 2010, no. 10. 32 p.
10. **Goldovsky Y. M., Zhelenkov B. V., Safonova I. E., Cyganova N. A.** To the question of evaluating the reliability of functioning of the elements of the telecommunication enterprise-wide networks, *International Scientific — Practical Conference "Innovative Information Technologies", Part 2, Innovative Information Technologies in Science. Prague, 2014, April 21—25*. P. 474—477.

УДК 519

О. В. Мандрикова^{1, 2}, д-р техн. наук, зав. лаб., e-mail: oksanaml@mail.ru,
Н. В. Фетисова (Глушкова)¹, мл. науч. сотр., e-mail: nv.glushkova@yandex.ru,
Ю. А. Полозов^{1, 2}, канд. техн. наук, ст. науч. сотр., e-mail: up_agent@mail.ru

¹ Институт космофизических исследований и распространения радиоволн ДВО РАН, с. Паратунка

² Камчатский государственный технический университет, г. Петропавловск-Камчатский

Моделирование временного хода параметров ионосферы и выделение аномалий

Описана новая многокомпонентная модель (МКМ) временного хода параметров ионосферы, основанная на совмещении вейвлетов с классом моделей авторегрессии проинтегрированного скользящего среднего. Модель позволяет изучать режим ионосферы и выделять аномалии, возникающие в периоды ионосферных возмущений. Построены МКМ критической частоты ионосферы для района Камчатки (станция "Паратунка", регистрацию данных выполняет Институт космофизических исследований и распространения радиоволн ДВО РАН). Выполнено сравнение полученных МКМ с эмпирической Международной справочной моделью ИРИ и методом скользящей медианы, которое показало перспективность предлагаемого авторами метода. Для детального изучения аномальных изменений в ионосфере предложено использовать непрерывное вейвлет-преобразование. На основе непрерывного вейвлет-преобразования разработаны вычислительные решения по выделению разномасштабных аномалий в ионосфере и оценке их параметров.

Анализ ионосферных параметров в периоды повышенной геомагнитной активности показал, что аномалии в ионосфере имеют крупные масштабы и достигают максимальной интенсивности в периоды наиболее сильных магнитных бурь. Замечено, что накануне главной фазы магнитной бури в анализируемом районе может наблюдаться повышение концентрации электронов.

Ключевые слова: вейвлет-преобразование, модель авторегрессии проинтегрированного скользящего среднего, критическая частота ионосферы, ионосферные возмущения

Введение

Работа направлена на создание средств анализа ионосферных параметров и выделения аномалий, возникающих в периоды ионосферных возмущений. Ионосфера Земли — область атмосферы, простирающаяся от 80 до 1000 км и влияющая на распространение радиоволн [1—3]. Структура ионосферы является изменчивой и неоднородной, ее изучение основано на анализе вариации регистрируемых параметров среды. Ионосферные параметры заметно меняются с высотой, зависят от цикла солнечной

активности, геомагнитных условий, географических координат и содержат характерные суточные и сезонные изменения [1—3]. В периоды ионосферных возмущений в ионосферных параметрах наблюдаются аномальные особенности, имеющие различную структуру и масштабы. В большинстве случаев ионосферные возмущения происходят вследствие повышения солнечной и геомагнитной активности, а в сейсмоактивных областях они также могут наблюдаться в периоды повышенной сейсмической активности [1, 3, 4].

Одной из важных задач обработки и анализа ионосферных параметров является контроль состояния ионосферы и выделение аномалий [1, 3], влияющих на многие аспекты нашей жизни, оказывая негативное влияние на работу спутниковых систем и распространение радиосвязи. Задачи анализа состояния ионосферы и выявления аномалий решаются большим числом авторов [3–6]. Среди основных подходов можно выделить традиционный метод скользящей медианы [1], эмпирические модели ионосферы [3, 5, 6], применение нейронных сетей [3, 7, 8] и вейвлет-преобразования [2, 7–10]. На сегодняшний момент наиболее разработанной эмпирической моделью ионосферы является Международная справочная ИРИ-модель [5, 6], которая основана на широком диапазоне наземных и космических данных. Ее точность для оценки параметров в конкретном регионе существенно зависит от наличия в нем регистрируемых данных [6]. С этим связан факт получения более точного прогноза на основе модели ИРИ для средних широт по сравнению с экваториальными и авроральными широтами. Также в работе [3] показано, что точность ИРИ-модели существенно зависит от уровня солнечной активности и убывает с ростом последней.

Изучение вариаций параметров ионосферы для анализируемого района Камчатки показало, что они имеют сложную нестационарную структуру. Это существенно ограничивает возможности и эффективность применения традиционных классических методов моделирования и анализа эмпирических данных. Как показывают последние исследования [2, 7–13], естественным и наиболее эффективным способом представления таких данных является построение нелинейных адаптивных аппроксимирующих схем. Основанные на этом подходе методы декомпозиции на эмпирические моды [12, 13] и адаптивные вейвлет-разложения [2, 7–11, 14] получают в настоящее время интенсивное развитие. Ввиду большого разнообразия ортогональных базисных вейвлетов с компактными носителями и наличия численно устойчивых быстрых алгоритмов преобразования данных вейвлет-разложение обеспечивает широкие возможности его применения для анализа сложных структур данных [15–17], в том числе в области геофизики [2, 7–11, 18].

В данной работе использовалось кратномасштабное вейвлет-разложение временного ряда параметров ионосферы, на основе которого временной ряд представлен в виде разномасштабных компонент более простой структуры, чем исходный ряд. Это позволило выделить стационарные компоненты и для их идентификации применить классические методы моделирования и анализа временных рядов. В работе использовался класс моделей авторегрессии интегрированного скользящего среднего (АРПСС) [19, 20]. Практика подтвердила мощь и гибкость методов АРПСС при решении многих прикладных

задач [19–21], в настоящее время эти методы получают развитие в области геофизики [9, 10, 21]. Но они имеют ограничения как на возможность их использования для отдельных временных рядов, так и на выявляемые при этом закономерности [9, 10, 20]. Оценка параметров моделей АРПСС, а также их диагностика и оптимизация полностью базируются на предположении, что данные имеют нормальное распределение, которое не всегда оправдано. Расширяя область применения этих методов, авторами предложена новая многокомпонентная модель (МКМ), основанная на совмещении вейвлетов и моделей АРПСС. Впервые данный подход был предложен для решения задачи выявления аномалий в данных подпочвенного радона, где показал свою эффективность [22]. В данной статье описан способ построения и оценки разработанной МКМ для параметров ионосферы. На примере ионосферных данных показана эффективность предлагаемого подхода, позволяющего подавить шум, упростить структуру данных и выделить стационарные составляющие, подлежащие идентификации. Сравнение полученных МКМ с ИРИ-моделью и методом скользящей медианы, широко применяемых для задач моделирования и анализа ионосферных параметров, показало перспективность предлагаемого авторами подхода.

Для детального изучения ионосферных параметров вместе с предложенным методом моделирования использовалось непрерывное вейвлет-преобразование. На основе непрерывного вейвлет-преобразования предложены вычислительные решения, позволившие выделить разномасштабные аномалии в ионосфере и оценить их моменты возникновения, временную длительность и интенсивность.

На основе моделирования изучен режим ионосферы в районе Камчатки, выполнен анализ суточных и сезонных вариаций параметров в периоды различной активности Солнца. В периоды повышенной солнечной и геомагнитной активности выделены аномалии в ионосфере. Детальный анализ показал, что максимальной интенсивности аномалии достигают в периоды наиболее сильных магнитных бурь и имеют крупные масштабы. Замечено, что накануне главной фазы магнитной бури в анализируемом районе может наблюдаться повышение концентрации электронов.

Описание метода моделирования и анализа параметров ионосферы

1.1. Построение МКМ

Временной ряд параметров ионосферы f_0 содержит стационарные составляющие, характеризующие режим ионосферы в районе регистрации данных, и случайный шум. Как показано в работах [9, 10], на основе кратномасштабного вейвлет-разложения временной ряд f_0 представляется в виде линейной комбинации разномасштабных составляющих: сгла-

женной $f[2^{-m}t]$ масштаба m и детализирующих $g[2^j t]$ масштабов $j = \overline{-1, -m}$:

$$f_0(t) = \sum_{j=-1}^{-m} g[2^j t] + f[2^{-m}t]. \quad (1)$$

В работе [16] показано, что когда масштабирующая функция ϕ имеет L нулевых моментов, т.е. $\int_{-\infty}^{+\infty} t^r \phi(t) dt = 0$, $r = \overline{1, L}$, и $f \in C^L$ (C^L — пространство функций, L раз непрерывно дифференцируемых), то для t вблизи $2^m k$ коэффициенты вейвлет-разложения

$$c_{-m, k} = \langle f, \phi_{-m, k} \rangle \approx 2^{-m/2} f(2^m k). \quad (2)$$

Из соотношения (2) следует, что в сглаженной компоненте $f[2^{-m}t] \in V_{-m}$ масштаба m мы имеем приближение аппроксимируемой функции f с разрешением 2^m . Высокочастотные особенности функции отображаются в пространства W_j .

Путем изменения уровня разложения m (см. соотношение (1)) мы можем получить различные представления ряда. Наша задача — определить *наилучшее представление*, которое выделит стационарные суточные и сезонные составляющие колебаний параметров ионосферы из шума и позволит получить для них адекватную модель АРПСС. Поскольку в соответствии с соотношением (2) в сглаженной компоненте $f[2^{-m}t]$ масштаба m мы имеем приближение аппроксимируемой функции f с разрешением 2^m (исходные данные масштаба $m = 0$ являются часовыми) и эти компоненты менее подвержены влиянию случайного фактора, чем высокочастотные детализирующие $g[2^j t]$, решение задачи построим на анализе сглаженных компонент.

1. Выполним кратномасштабные вейвлет-разложения ряда до уровней $m = \overline{1, M}$ (максимальный допустимый уровень разложения M определяется длиной ряда N : $M \leq \log_2 N$) и получим набор сглаженных компонент $f[2^{-m}t] = \sum_k c_{-m, k} \phi_{-m, k}(t)$, $m = \overline{1, M}$.

2. Из набора компонент $f[2^{-m}t]$, $m = \overline{1, M}$, определим стационарные компоненты. Для стационарных компонент $f[2^{-m}t]$, используя традиционные подходы [19, 20], выберем модели из класса АРПСС для их аппроксимации и получим представление каждой компоненты в виде

$$f_{-m}(t) = \sum_k s_{-m, k} \phi_{-m, k}(t),$$

где $s_{-m, k} = \sum_{l=1}^p \gamma_{-m, l} \omega_{-m, k-l}^{-1} - \sum_{n=1}^h \theta_{-m, n} a_{-m, k-n}^{-1}$ — оценочное значение сглаженной компоненты; $\omega_{-m, k} = \nabla^v c_{-m, k}$, ∇^v — оператор взятия разности порядка v ; $p, \gamma_{-m, l}$ — порядок и параметры авторегрес-

сии сглаженной компоненты; $h, \theta_{-m, n}$ — порядок и параметры скользящего среднего сглаженной компоненты; $a_{-m, k-n}$ — остаточные ошибки модели.

3. Выполним оценку погрешностей *моделей компонент*:

$$E_m = \sum_{k=1}^K \sum_{l=q=1}^Q e_{k+q}^m,$$

где $e_{k+q}^m = (s_{-m, k+q}^{\text{факт}} - s_{-m, k+q}^{\text{прогноз}})^2$ — ошибка *модели компоненты* в точке k с упреждением q ; $s_{-m, k+q}^{\text{факт}}$ — фактические значения компоненты ряда; $s_{-m, k+q}^{\text{прогноз}}$ — модельные значения компоненты ряда; Q — длина упреждения данных; K — длина компоненты ряда.

4. *Наилучшим представлением* временного ряда будем считать представление, соответствующее кратномасштабному вейвлет-разложению до уровня m^* , где m^* : $E_{m^*} = \min_m E_m$.

5. Из набора детализирующих компонент $g[2^j t]$, $j = \overline{-1, -m^*}$, определим стационарные компоненты. Для стационарных компонент $g[2^j t]$, используя традиционные подходы [19, 20], выберем модели из класса АРПСС для их аппроксимации.

6. Компоненты $g[2^j t]$, которые не являются стационарными, содержат локальные особенности и шум и будут исследованы другим методом.

7. Используя соотношение (1), объединим полученные *модели компонент* в общую многокомпонентную конструкцию, описывающую временной ход данных:

$$f_0(t) = \sum_{\mu=1, T} \sum_{k=1, N_j^\mu} s_{j, k}^\mu b_{j, k}^\mu(t), \quad (3)$$

где $s_{j, k}^\mu = \sum_{l=1}^{p_j^\mu} \gamma_{j, l}^\mu \omega_{j, k-l}^\mu - \sum_{n=1}^{h_j^\mu} \theta_{j, n}^\mu a_{j, k-n}^\mu$ — оценочное

значение μ -й компоненты; $p_j^\mu, \gamma_{j, l}^\mu$ — порядок и параметры авторегрессии μ -й компоненты; $h_j^\mu, \theta_{j, n}^\mu$ — порядок и параметры скользящего среднего μ -й компоненты; $\omega_{j, k}^\mu = \nabla^{v_j^\mu} \beta_{j, k}^\mu$; v_j^μ — порядок разности μ -й компоненты; $\beta_{j, k}^1 = c_{j, k}$, $\beta_{j, k}^\mu = d_{j, k}$, $\mu = \overline{2, T}$, T — число моделируемых компонент; $a_{j, k-n}^\mu$ — остаточные ошибки модели μ -й компоненты; N_j^μ — длина μ -й компоненты; $b_{j, k}^1 = \phi_{j, k}$ — масштабирующая функция; $b_{j, k}^\mu = \Psi_{j, k}$, $\mu = \overline{2, T}$, — вейвлет-базис μ -й компоненты.

Прогнозирование значения $s_{j,k+q}^\mu$, $q \geq 1$ определяет прогноз $s_{j,k}^\mu$ в точке k с упреждением q . Значение $s_{j,k+q}^\mu$ на основе модели μ -й компоненты определяется как

$$s_{j,k+q}^\mu = \sum_{l=1}^{p_j^\mu} \gamma_{j,l}^\mu \omega_{j,k+q-l}^\mu - \sum_{n=1}^{h_j^\mu} \theta_{j,n}^\mu a_{j,k+q-n}^\mu.$$

Остаточные ошибки модели μ -й компоненты определяются как разность между фактическими и прогнозными значениями в точке $k+q$: $a_{j,k+q}^\mu = s_{j,k+q}^{\mu, \text{факт}} - s_{j,k+q}^{\mu, \text{прогноз}}$.

Модель (3) описывает характерные изменения данных. В период аномальных изменений данных абсолютные значения остаточных ошибок моделей компонент возрастут, поэтому процедура выделения аномалий может быть построена, например, на проверке следующего условия:

$$\varepsilon_\mu = \sum_{q=1}^{Q_\mu} |a_{j,k+q}^\mu| > T_\mu, \quad (4)$$

где Q_μ — длина упреждения данных на основе модели μ -й компоненты; T_μ — пороговое значение μ -й компоненты, определяющее наличие аномалии.

Пороговое значение T_μ в соотношении (4) может быть определено либо:

1) на основе оценки дисперсии ошибок прогноза данных [19]:

$$T_\mu(Q_\mu) = \left\{ 1 + \sum_{q=1}^{Q_\mu-1} (\psi_{j,q}^\mu)^2 \right\}^{1/2} \sigma_{a_{j,k+q}^\mu},$$

где $\sigma_{a_{j,k+q}^\mu}$ — среднеквадратическое отклонение остаточных ошибок модели μ -й компоненты; $\psi_{j,q}^\mu$ — весовые коэффициенты модели μ -й компоненты, которые находятся приравниванием коэффициентов в выражении

$$\begin{aligned} & (1 - \gamma_{j,1}^\mu B - \gamma_{j,2}^\mu B^2 - \dots - \gamma_{j,p_j^\mu+v_j^\mu}^\mu B^{p_j^\mu+v_j^\mu}) \times \\ & \times (1 + \psi_{j,1}^\mu B + \psi_{j,2}^\mu B^2 + \dots) = \\ & = (1 - \theta_{j,1}^\mu B - \theta_{j,2}^\mu B^2 - \dots - \theta_{j,\eta_j^\mu}^\mu B^{\eta_j^\mu}), \end{aligned}$$

где B — оператор сдвига назад: $B^l \omega_{j,k}^\mu(t) = \omega_{j,k-l}^\mu(t)$,

$$\psi_{j,0}^\mu = 0,$$

2) либо с учетом вероятностных пределов

$$T_\mu(Q_\mu) = u_{\varepsilon/2} \left\{ 1 + \sum_{q=1}^{Q_\mu-1} (\psi_{j,q}^\mu)^2 \right\}^{1/2} \sigma_{a_{j,k+q}^\mu}, \quad (5)$$

где $u_{\varepsilon/2}$ — квантиль уровня $1 - \varepsilon/2$ стандартного нормального распределения.

1.2. Оценка МКМ параметров ионосферы для района Камчатки

В работе использовались часовые данные критической частоты ионосферного слоя F2 за период с 1969 по 2013 г. района Камчатки (п-ов Камчатка, с. Паратунка). Для определения степени геомагнитной возмущенности использовался К-индекс, который был вычислен на основе геомагнитных данных станции "Паратунка" (Камчатка, Россия, регистрацию выполняет ИКИР ДВО РАН). При построении моделей использовались данные ионосферы foF2 за периоды относительно спокойного геомагнитного поля (периоды, в которые суммарный за сутки К-индекс ΣK не превышал значения 23), не содержащие сильных сейсмических событий (периоды, в которые отсутствовали землетрясения с энергетическим классом $K_s \geq 12$, произошедшие в радиусе до 300 км от станции ионосферного зондирования, данные сейсмического каталога п-ова Камчатка были любезно предоставлены Камчатским филиалом геофизической службы РАН).

Кратномасштабные вейвлет-разложения выполнялись по вейвлет-базису Добеши 3-го порядка, который был определен на основе минимизации погрешности аппроксимации [9, 10]. Описанным в п. 1.2 способом были выделены стационарные компоненты ряда foF2 и определено его наилучшее представление, соответствующее кратномасштабному вейвлет-разложению до уровня $m^* = 3$:

$$f_0(t) = f[2^{-3}t] + g[2^{-3}t] + \sum_{j=-1}^{-2} g[2^j t],$$

где $f[2^{-3}t] = \sum_k c_{-3,k} \phi_{-3,k}(t)$ — сглаженная стационарная компонента, $g[2^{-3}t] = \sum_k d_{-3,k} \Psi_{-3,k}(t)$ — де-

тализирующая стационарная компонента, $g[2^j t] = \sum_k d_{j,k} \Psi_{j,k}(t)$, $j = -1, -2$ — детализирующие компоненты, содержат локальные особенности и шум.

На рис. 1 в качестве примера представлены автокорреляционные функции (АКФ) исходного ряда (анализировался период 19.12.2011—08.01.2012), его первой разности и выделенных стационарных компонент. Анализ рис. 1 показывает, что в данный период исходный ряд имеет сложную нестационарную структуру и непосредственное применение методов АРПСС не позволяет получить адекватную модель ряда. Выделенные компоненты ряда $f[2^{-3}t]$ и $g[2^{-3}t]$ имеют затухающие автокорреляционные функции и частные автокорреляционные функции порядка 3, что позволяет идентифицировать для них модели авторегрессии порядка 3 [19] и подтверждает эффективность предлагаемого подхода.

Оценка параметров АРПСС моделей для выделенных стационарных компонент выполнялась с учетом сезонного хода ионосферы и уровня сол-

нечной активности. МКМ были получены отдельно для каждого сезона и высокого и низкого уровней солнечной активности.

В соответствии с соотношением (3) диагностика полученных МКМ может быть выполнена на основе проверки адекватности образующих их моделей компонент. Она выполнялась на основе следующих методик, основанных на анализе остаточных ошибок.

1. На основе *совокупного критерия согласия* [19] подгоняемая модель удовлетворительна, если

$$Q^{\mu} = n \sum_{z=1}^Z r_z^2(a^{\mu})$$

распределено приближенно как $\chi^2(Z - h_j^{\mu} - p_j^{\mu})$, где Z — рассматриваемые первые автокорреляции остаточных ошибок модели μ -й компоненты; $r_z(a^{\mu})$ — автокорреляции ряда остаточных ошибок модели μ -й компоненты; $n = N - \vartheta$, N — длина ряда μ -й компоненты, ϑ — порядок разности модели μ -й компоненты.

2. На основе нормированной кумулятивной периодограммы

$$C^{\mu}(f_{\beta}) = \frac{\sum_{i=1}^{\beta} I(f_i)}{ns^2},$$

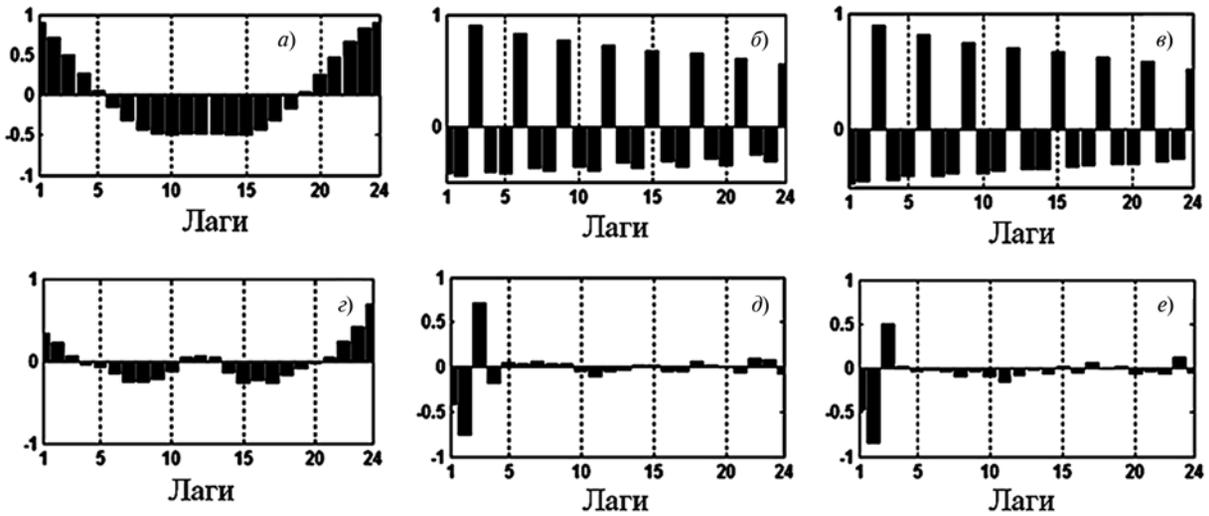


Рис. 1. АКФ ряда foF2 (а); АКФ первой разности ряда foF2 (б); АКФ компоненты $f[2^{-3}t]$ (в); частная АКФ компоненты $f[2^{-3}t]$ (г); АКФ компоненты $g[2^{-3}t]$ (д); частная АКФ компоненты $g[2^{-3}t]$ (е)

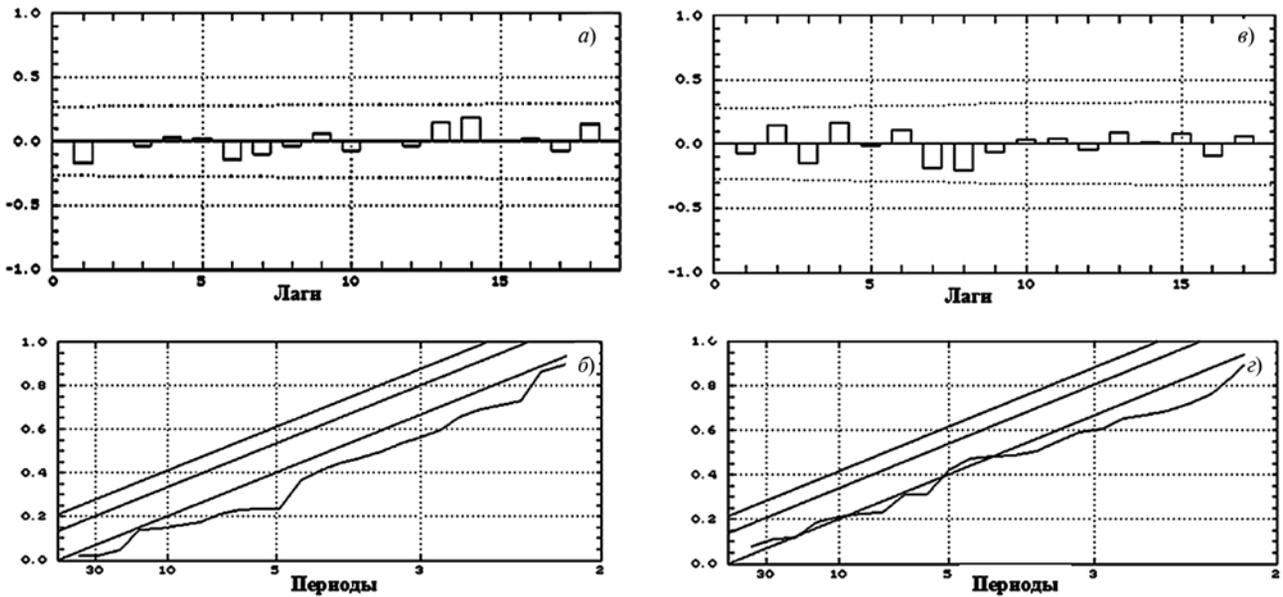


Рис. 2. Результаты диагностики моделей компонент: а — АКФ остаточных ошибок модели компоненты $f[2^{-3}t]$; б — кумулятивная периодограмма остаточных ошибок модели компоненты $f[2^{-3}t]$; в — АКФ остаточных ошибок модели компоненты $g[2^{-3}t]$; г — кумулятивная периодограмма остаточных ошибок модели компоненты $g[2^{-3}t]$

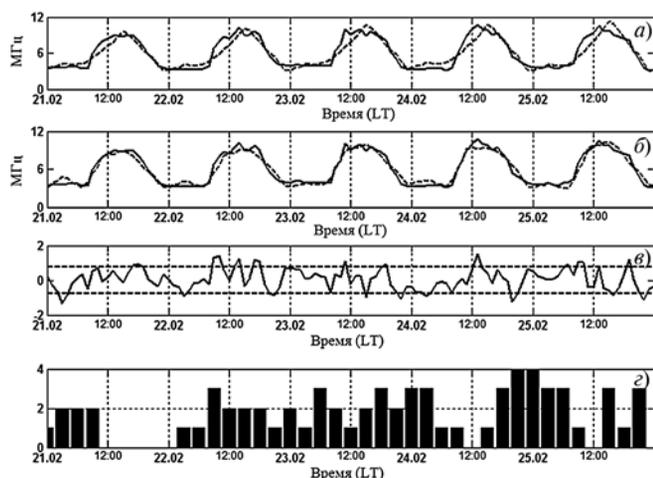


Рис. 3. Моделирование данных foF2 (ст. "Паратунка") за период с 21.02.1999 по 25.02.1999, время местное:

a — foF2 (сплошная линия), моделирование данных foF2 на основе модели компоненты $f[2^{-3}t]$ (штриховая линия); *b* — foF2 (сплошная линия), моделирование данных foF2 на основе полученной МКМ (штриховая линия); *в* — ошибки моделирования МКМ; *г* — K-индекс геомагнитной активности (ст. "Паратунка"). На графике *в* штриховой линией показаны стандартные отклонения ошибок МКМ

где $I(f_i)$ — периодограмма ряда остаточных ошибок модели μ -й компоненты a_k^μ , $k = \overline{1, n}$, n — длина ряда a_k^μ :

$$I(f_i) = \frac{2}{n} \left[\left(\sum_{k=1}^n a_k^\mu \cos 2\pi f_i k \right)^2 + \left(\sum_{k=1}^n a_k^\mu \sin 2\pi f_i k \right)^2 \right],$$

$f_i = i/n$ — частота; s^2 — оценка $\sigma_{a^\mu}^2$ ряда остаточных ошибок модели μ -й компоненты.

Результаты диагностики моделей компонент, представленные на рис. 2 для периода 19.12.2011—08.01.2012, подтверждают их адекватность моделируемому данным. Аналогичные результаты были получены в анализе данных за другие годы.

Процесс моделирования данных foF2 в период относительно спокойного геомагнитного поля (рис. 3) показывает хорошие аппроксимирующие свойства модели и ее сходимость к процессу.

Сравнение МКМ с методом скользящей медианы и ИРИ-моделями, которые получены для различных сезонов и уровней солнечной активности (рис. 4, 5, таблица), показывает, что МКМ для ана-

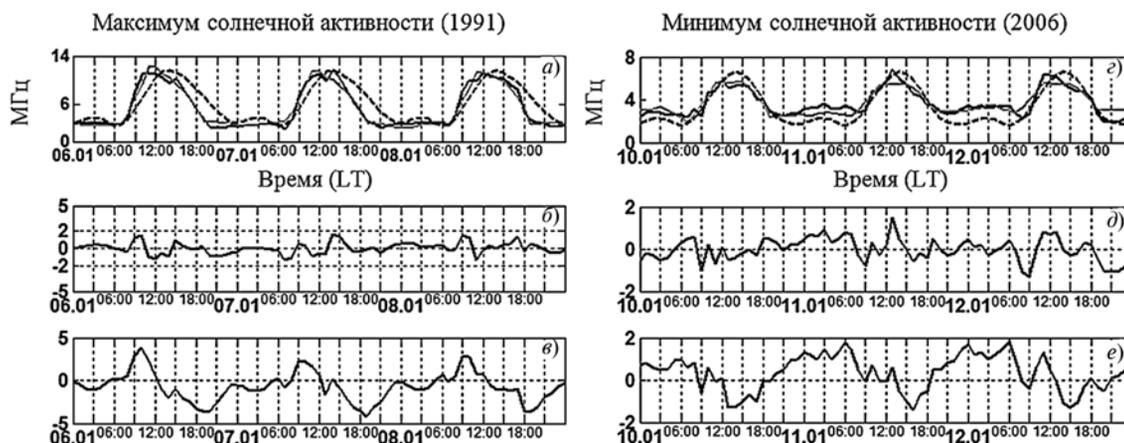


Рис. 4. Прогнозирование данных foF2 зимнего периода времени (время местное):

a, г — foF2 (черная сплошная линия), прогноз по МКМ (серая сплошная линия), прогноз по ИРИ-модели (черная штриховая линия); *б, д* — ошибки МКМ; *в, е* — ошибки ИРИ-модели

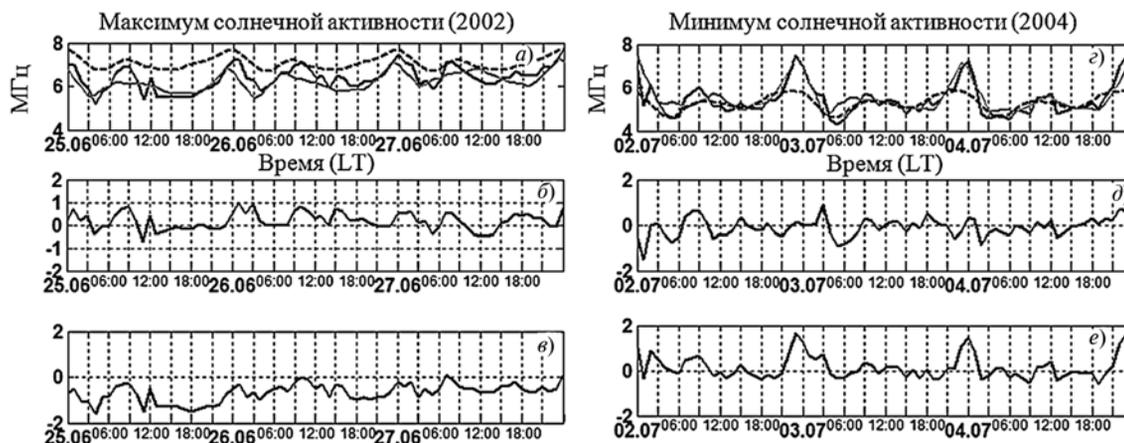


Рис. 5. Прогнозирование данных foF2 летнего периода времени (время местное):

a, г — foF2 (черная сплошная линия), прогноз по МКМ (серая сплошная линия), прогноз по ИРИ-модели (черная штриховая линия); *б, д* — ошибки МКМ; *в, е* — ошибки ИРИ-модели

Оценка погрешностей, полученных МКМ и ИРИ-моделями

Анализируемый период	Сумма квадратов ошибок/ коэффициент корреляции	
	МКМ	ИРИ-модель
06.01—16.06.1991	132,39/1,00	719,7/0,89
03.01—13.01.2006	131,4/0,87	236,7/0,48
22.06—03.07.2002	91,4/0,69	183,8/0,49
29.06—14.07.2004	104,72/0,76	156,31/0,62

лизируемого района позволяют получить более точный прогноз, особенно в период максимума солнечной активности. Летом, в период максимума солнечной активности, ИРИ завышает значения foF2 (рис. 5, а), в период минимума — занижает (рис. 5, з). В период максимума солнечной активности существенное возрастание ошибок ИРИ-моделей наблюдаются с 09:00 до 00:00 по местному времени (см. рис. 4, в), в период минимума — с 21:00 до 03:00 (рис. 5, е), что согласуется с результатами работы [3]. Наблюдаемое наличие корреляции ошибок ИРИ-моделей ставит под сомнение их адекватность. В свою очередь, ошибки МКМ яв-

ляются белым шумом, что подтверждено результатами диагностики.

При выполнении прогноза в соответствии с равенством (5) мы можем с заданной вероятностью оценить его доверительный интервал. Когда ошибки *моделей компонент* выходят за этот доверительный интервал, мы можем фиксировать аномалию в ионосфере, что затруднительно сделать для ИРИ-модели и скользящей медианы. На рис. 6 представлены результаты моделирования данных в периоды возмущенного геомагнитного поля. В моменты повышения геомагнитной активности ошибки *моделей компонент* возрастают и выходят за пределы стандартного отклонения (доверительная вероятность более 70 %), что свидетельствует о возникновении аномальных изменений во временном ходе ряда foF2. Оценка медианных значений ряда foF2 (рис. 6, а, серая линия) показывает наибольшие отклонения как в моменты повышенной геомагнитной активности (05.02.11, 15.02.11), так и в моменты спокойного геомагнитного поля (12.02.11). ИРИ-модель также не позволяет выделить аномальные периоды в ионосфере и показывает незначительное возрастание ошибок в анализируемый период (геомагнитное поле незначительно возмущено, рис. 6, е).

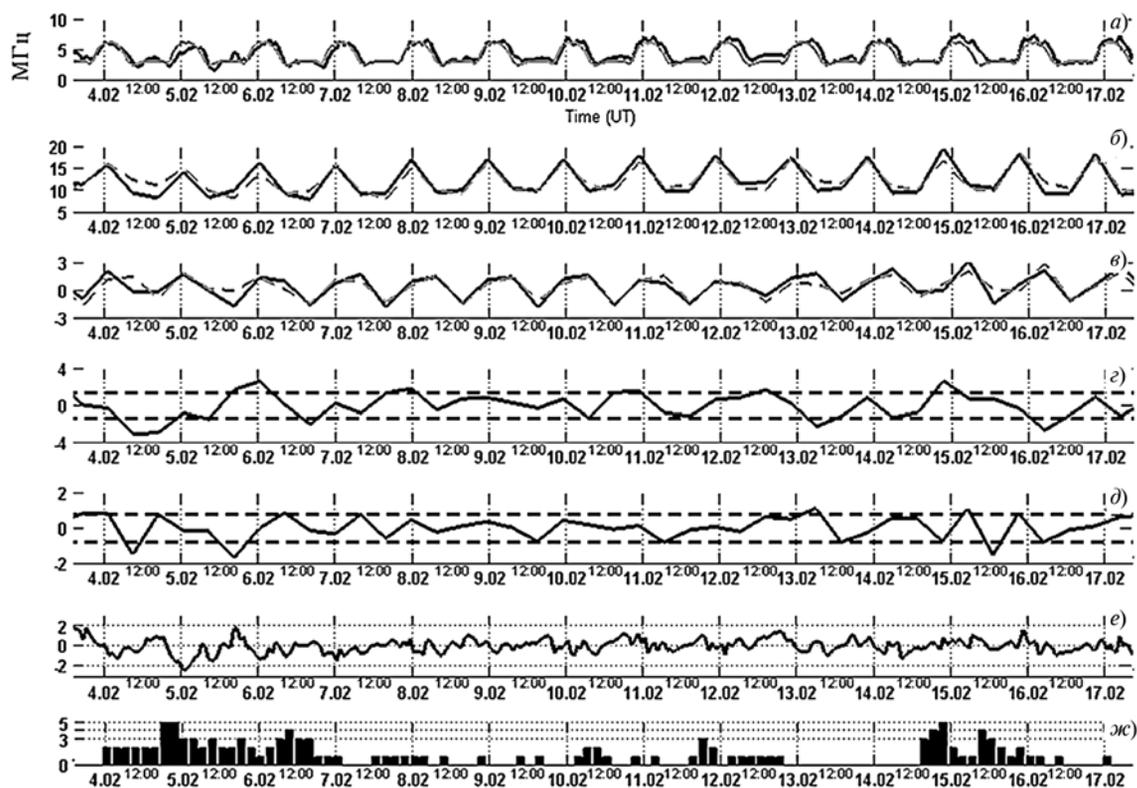


Рис. 6. Результаты моделирования компонент ряда foF2 за период 4.02.2011—17.02.2011 г.:

а — foF2 (черная линия), медианные значения ряда foF2 (серая линия); б — фактические значения компоненты $f[2^{-3}f]$ (черная сплошная линия), модельные значения компоненты $f[2^{-3}f]$ (серая штриховая линия); в — фактические значения компоненты $g[2^{-3}f]$ (черная сплошная линия), модельные значения компоненты $g[2^{-3}f]$ (серая штриховая линия); г — ошибки модели компоненты $f[2^{-3}f]$; д — ошибки модели компоненты $g[2^{-3}f]$; е — ошибки ИРИ-модели; ж — К-индекс (ст. "Паратунка"). На графиках з, д штриховыми линиями показаны стандартные отклонения ошибок моделей компонент

1.3. Выделение аномалий в ионосфере на основе непрерывного вейвлет-преобразования и пороговых функций

Относительно каждого базисного вейвлета Ψ непрерывное вейвлет-преобразование определяется формулой [16, 17]

$$W_{\Psi}f_{b,a} := |a|^{-1/2} \int_{-\infty}^{\infty} f(t)\Psi\left(\frac{t-b}{a}\right) dt,$$

$$f \in L^2(R), a, b \in R, a \neq 0.$$

Убывание амплитуды коэффициентов $|W_{\Psi}f_{b,a}|$ в зависимости от масштаба a связано с равномерной и точечной гладкостью Липшица функции f [15, 16]. Когда масштаб a убывает, амплитуды коэффициентов $|W_{\Psi}f_{b,a}|$ имеют быстрое убывание до нуля в областях, где функция f гладкая и не имеет локальных особенностей [15, 23]. Основываясь на этом свойстве вейвлет-преобразования, для выделения локальных особенностей во временном ряде foF2, свидетельствующих о возникновении аномалий в ионосфере, использовалась пороговая функция:

$$P_{T_a}(W_{\Psi}f_{b,a}) = \begin{cases} W_{\Psi}f_{b,a}, & \text{если } |W_{\Psi}f_{b,a} - W_{\Psi}f_{b,a}^{med}| \geq T_a; \\ 0, & \text{если } |W_{\Psi}f_{b,a} - W_{\Psi}f_{b,a}^{med}| < T_a. \end{cases}$$

Порог $T_a = U \cdot St_a$ определяет наличие аномалии на масштабе a вблизи точки ξ , содержащейся в носителе $\Psi_{b,a}$; U — коэффициент порога;

$$St_a = \sqrt{\frac{1}{\Phi-1} \sum_{k=1}^{\Phi} (W_{\Psi}f_{b,a} - \overline{W_{\Psi}f_{b,a}})^2}, \quad \overline{W_{\Psi}f_{b,a}} \text{ и}$$

$W_{\Psi}f_{b,a}^{med}$ — среднее значение и медиана, определяемые в скользящем временном окне длины Φ .

С учетом суточного хода ионосферных данных среднее значение $\overline{W_{\Psi}f_{b,a}}$ и медиана $W_{\Psi}f_{b,a}^{med}$ вычислялись отдельно для каждого часа.

В силу случайной природы данных использование любого порога T_a (определяющего правило выбора решения о наличии либо отсутствии аномалии) неминуемо связано с возможностью ошибочных решений. В работе в качестве критерия качества выбора решения использовался критерий наименьшей частоты ошибок (оценивался и минимизировался апостериорный риск [24]), который при предполагаемых априорных данных представляет наиболее полную их характеристику. При оценке апостериорного риска для определения состояния ионосферы использовались данные ионограмм (данные ионосферной станции "Паратунка", Камчатка, Россия, регистрацию выполняет ИКИР ДВО РАН), которые также сопоставлялись с геомагнитными данными (использовался К-индекс) и данными каталога землетрясений п-ова Камчатка. Исследование показало зависимость порога T_a от уровня солнечной активности и его возрастание в периоды высокой активности Солнца. Поэтому оценка порогов для периодов лет с высокой и низкой активностью Солнца выполнялась отдельно.

Если вейвлет Ψ имеет компактный носитель, равный $[-\Omega, \Omega]$, то множество пар точек (b, a) таких,

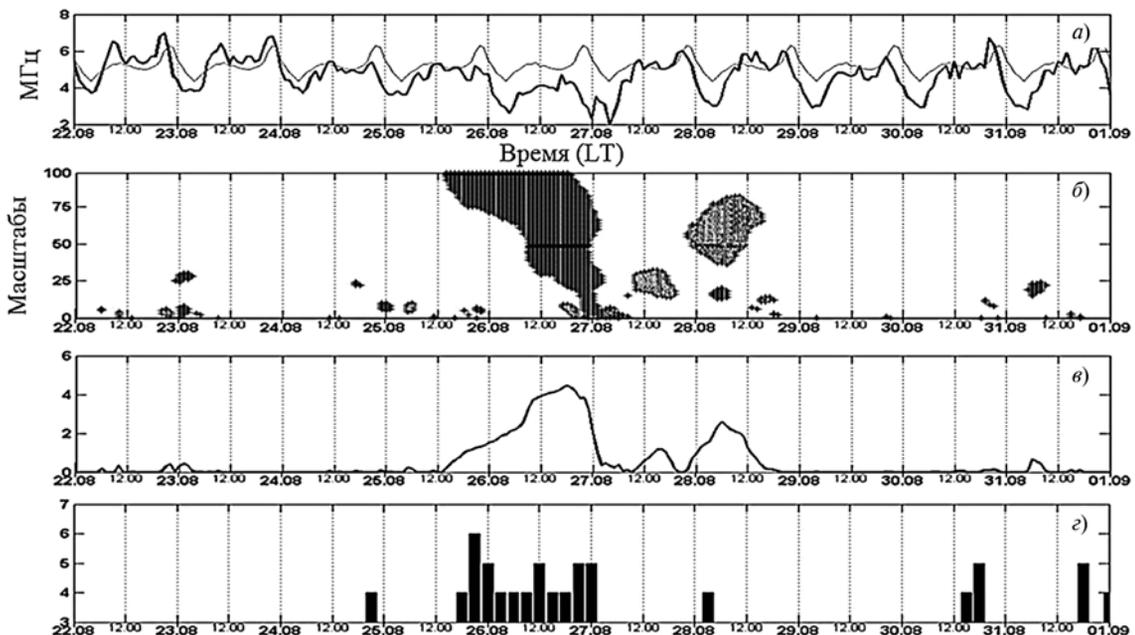


Рис. 7. Результаты обработки ионосферных данных за период 22.08—31.08.1987 (станция "Паратунка", Камчатка): а — foF2 (черная линия), медианные значения ряда foF2 (серая линия); б — выделенные аномалии, коэффициент порога $U = 2,3$, длина скользящего временного окна $\Phi = 336$ ч; в — оценка интенсивности аномалий; з — значения К-индекса более 3 (станция "Паратунка", Камчатка)

что точка ξ содержится в носителе $\Psi_{b,a}$, определяют конус влияния точки ξ [15]. Так как носитель $\Psi_{b,a}$ на масштабе a равен $[b - \Omega a, b + \Omega a]$, то конус влияния точки ξ на масштабе a определяется неравенством $|b - \xi| \leq \Omega a$.

Следовательно, временная длительность аномалии на масштабе a определяется конусом влияния точки ξ и равна

$$H_a = 2\Omega a.$$

Интенсивность аномалии в момент времени $t = b$ можно определить как

$$Y_b = \sum_a \frac{P_{T_a}(W_{\Psi} f_{b,a})}{\|W_{\Psi} f_{b,a}\|_2},$$

где норма $\|W_{\Psi} f_{b,a}\|_2 = \sqrt{\sum_a (P_{T_a}(W_{\Psi} f_{b,a}))^2}$, N_a — длина ряда на масштабе a .

Результаты выделения аномалий в ионосфере и оценки их интенсивности в период магнитной бури 25–26 августа 1987 г. показаны на рис. 7. Если значения вейвлет-коэффициентов $W_{\Psi} f_{b,a}$ превышают на величину порога T_a соответствующие им медианные значения $W_{\Psi} f_{b,a}^{med}$, то аномалию считают положительной (характеризует увеличение концентрации электронов в ионосфере по сравнению с характерным уровнем, показаны на рис. 7, б серым цветом). Если медианные значения $W_{\Psi} f_{b,a}^{med}$ превышают на величину порога T_a соответствующие им значения вейвлет-коэффициентов $W_{\Psi} f_{b,a}$, считаем аномалию отрицательной (характеризует уменьшение концентрации электронов по сравнению с характерным уровнем, показаны на рис. 7, б черным цветом). Анализ результатов показывает, что в период магнитной бури в ионосфере возникает отрицательная аномалия, длительностью более суток. Интенсивность аномалии возрастает с момента начала магнитной бури и достигает наибольших значений в период основной фазы бури. После магнитной бури концентрация электронов повышается, о чем свидетельствуют: положительные аномалии, наблюдаемые 28 августа 1987 г. В период бури также возникают мелкомасштабные аномалии, связанные с

локальными колебаниями электронной плотности ионосферы. Вычисление медианных значений ряда foF2 (показано на рис. 7, а серым цветом), в отличие от предлагаемых решений, не позволяет получить количественные оценки возмущений в ионосфере и локализовать аномальный период. Наибольшие отклонения медианных значений ряда foF2 наблюдаются как в период магнитной бури, так и в моменты спокойного геомагнитного поля, преимущественно в ночное время.

2. Анализ данных в периоды магнитных бурь

На рис. 8 показаны результаты анализа ионосферных и геомагнитных данных в периоды магнитной бури 17 марта 2013 г., имеющей резкое начало. Сопоставление параметров солнечного ветра с результатами обработки геомагнитных и ионосферных данных показывает наличие общего характера процессов. В период существенного возрастания скорости солнечного ветра с 05.25 по 05.55 UT до 710 км/с в геомагнитном поле сформировались возмущения, которые достигли наибольших значений с 06.15 по 19.50 UT. Одновременно в ионосфере возникла крупномасштабная отрицательная аномалия дли-

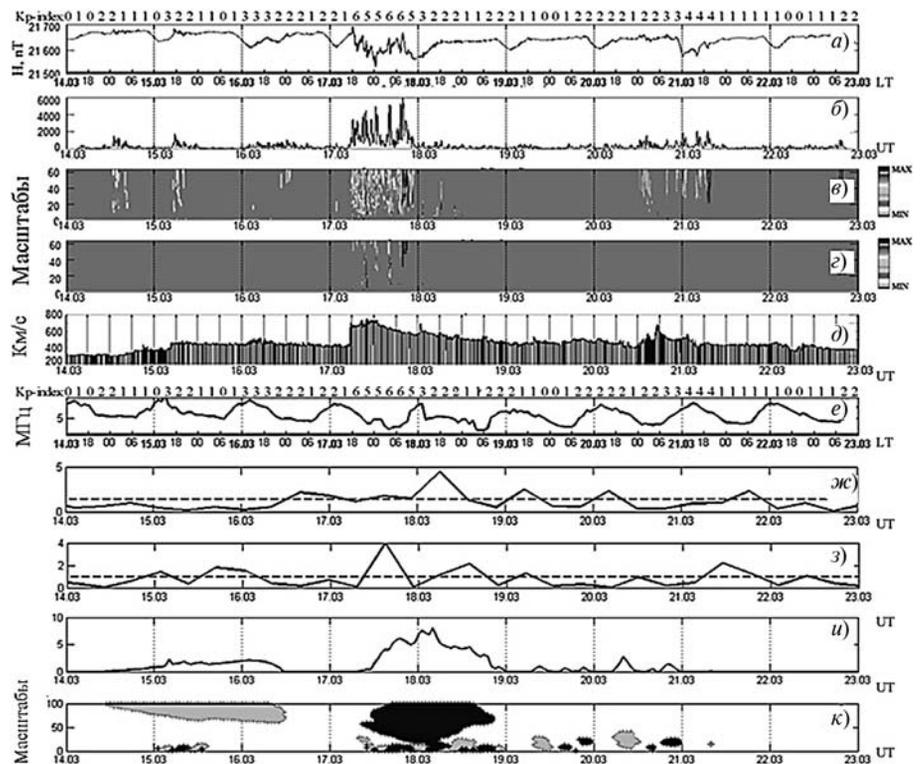


Рис. 8. Результаты обработки данных за период 14–22 марта 2013 г.: а — N-компонента магнитного поля Земли; б — оценка интенсивности геомагнитных возмущений; в — выделение периодов слабых и сильных геомагнитных возмущений; г — выделение периодов сильных геомагнитных возмущений; д — скорость солнечного ветра; е — данные foF2; ж — абсолютные значения ошибок МКМ сглаженной компоненты; з — абсолютные значения ошибок МКМ детализирующей компоненты; и — оценка интенсивности аномалий; к — выделенные аномалии, коэффициент порога $U = 2,5$, длина скользящего временного окна $\Phi = 336$ ч. Штриховая линия на графиках ж, з — стандартные отклонения ошибок моделей компонент

тельностью около суток и достигшая максимальной интенсивности преимущественно в дневное время в период 6:00—18:00 18 марта LT. Накануне магнитной бури (15 и 16 марта 2013 г.) наблюдаются локальные возрастания скорости солнечного ветра, которые сопровождаются слабыми возмущениями в геомагнитном поле, положительной крупномасштабной аномалией в ионосфере (длительностью более суток) и мелкомасштабными аномалиями, связанными с локальными колебаниями электронной плотности ионосферы.

Выводы

На основе предложенного авторами метода для района Камчатки построены модели вариаций критической частоты ионосферы. Сравнение полученных моделей с эмпирической Международной справочной моделью ИРИ и методом скользящей медианы показало перспективность предлагаемого авторами метода и возможность его использования для получения более достоверной информации о состоянии ионосферы в анализируемом районе. Разработанные вычислительные решения, основанные на непрерывном вейвлет-преобразовании, позволили в периоды возмущений выделить в ионосфере разномасштабные аномалии, оценить их длительность и интенсивность.

Выполненный анализ ионосферных и геомагнитных данных в период сильной магнитной бури 17 марта 2013 г. показал картину протекания процессов и их особенности накануне и в период события. Периоды существенного возрастания скорости солнечного ветра накануне главной фазы магнитной бури сопровождались формированием в геомагнитном поле возмущений и возникновением крупномасштабных отрицательных аномалий в ионосфере высокой интенсивности. В моменты локальных незначительных возрастаний скорости солнечного ветра в геомагнитном поле выделены слабые возмущения, которые сопровождались разномасштабными аномальными изменениями в параметрах ионосферы. Накануне события в ионосфере выделена крупномасштабная положительная аномалия, свидетельствующая о повышении концентрации электронов, а также аномалии мелких масштабов, связанные с локальными колебаниями электронной плотности ионосферы. Полученные результаты важны для задач оценки состояния околоземного космического пространства и прогноза сильных магнитных бурь.

Работа поддержана грантом РНФ № 14-11-00194, грантом стипендии Президента РФ СП-2976.2013.5 и грантом "УМНИК" № 11754р/17262 от 05.04.2013.

Авторы благодарят институты, поддерживающие станции регистрации ионосферных параметров и магнитные обсерватории, данные которых были использованы в исследовании, а также выражают признательность сотрудникам Камчатского филиала

ла геофизической службы РАН, обеспечившим качество регистрации сейсмических данных.

Список литературы

1. **Афраймович Э. Л., Первалова Н. П.** GPS-мониторинг верхней атмосферы Земли. Иркутск: ГУ НУ РВХ ВСНЦ СО РАН. 2006. 480 с.
2. **Kato H.** Development of automatic scaling software of ionospheric parameters // Journal of the National Institute of Information and Communications Technology. 2009. Vol. 56. P. 465—474.
3. **Nakamura M., Maruyama T., Shidama Y.** Using a neural network to make operational forecasts of ionospheric variations and storms at Kokubunji, Japan // Journal of the National Institute of Information and Communications Technology. 2009. Vol. 56. P. 391—406.
4. **Klimenko M. V.** Modeling of local disturbance formation in the ionosphere electron concentration before strong earthquakes // Earth, Planets and Space. 2012. Vol. 64, N 6. P. 441—450.
5. **Klimenko M. V.** Numerical modeling of the global ionospheric effects of storm sequence on September 9—14, 2005 — comparison with IRI model // Earth, Planets and Space. 2012. Vol. 64, N 6. P. 433—440.
6. **Bilitza D., Reinisch B. W.** International Reference Ionosphere 2007: Improvements and new parameters // Advances in space research. 2008. Vol. 42. P. 599—609.
7. **Мандрикова О. В., Полозов Ю. А.** Метод выделения аномальных особенностей в данных критической частоты ионосферы на основе совмещения вейвлет-преобразования и нейронных сетей // Цифровая обработка сигналов. 2012. № 2. С. 29—35.
8. **Мандрикова О. В., Полозов Ю. А.** Аппроксимация и анализ ионосферных параметров на основе совмещения вейвлет-преобразования с коллективами нейронных сетей // Информационные технологии. 2014. № 7. С. 61—65.
9. **Мандрикова О. В., Глушкова Н. В., Полозов Ю. А.** Алгоритмы выделения и анализа аномалий в параметрах критической частоты ионосферы foF2 на основе совмещения вейвлет-преобразования и авторегрессионных моделей // Цифровая обработка сигналов. 2013. № 1. С. 47—53.
10. **Mandrikova O. V., Glushkova N. V., Zhivet'ev I. V.** Modeling and analysis of ionospheric parameters by a combination of wavelet transform and autoregressive models // Geomagnetism and Aeronomy. 2014. Vol. 54, N 5. P. 593—600. DOI:10.1134/S0016793214050107.
11. **Zaourar N.** Wavelet-based multiscale analysis of geomagnetic disturbance // Harm Planets Space. 2013. Vol. 65, N 12. P. 1525—1540.
12. **Yu Z.** Modeling and simulation of the horizontal component of the geomagnetic field by fractional stochastic differential equations in conjunction with empirical mode decomposition // Journal of geophysical research. 2010. Vol. 115, N A10. DOI: 10.1029/2009JA015206.
13. **Klionsky D. M., Oreshko N. I., Geppener V. V.** Applications of empirical mode decomposition for processing nonstationary signals // Pattern recognition: and image analysis. 2008. Vol. 13, N 3. P. 390—399.
14. **Туровский А. Я., Кургулин С. Д., Семенов А. Г.** Исследование динамики максимумов локальных вейвлет-спектров вызванных зрительных потенциалов головного мозга // Информационные технологии. 2013. № 10. С. 46—50.
15. **Mallat S.** A wavelet tour of signal processing. London: Academic Press. 1999. P. 637.
16. **Daubechies I.** Ten lectures on wavelets / CBMS-NSF Lecture Notes nr. 61, SIAM. 1992. P. 377.
17. **Chui C. K.** An introduction to wavelets. Boston: Academic Press. 1992. P. 264.
18. **Мандрикова О. В., Соловьев И. С.** Вейвлет-технология вариации магнитного поля Земли // Информационные технологии. 2011. № 1. С. 34—38.
19. **Box G., Jenkins G.** Time series analysis: Forecasting and control. San Francisco: Holden-Day. 1970. P. 553.
20. **Marple S.** Digital spectral analysis with applications. New-Jersey: Prentice-Hall. 1987. P. 492.
21. **Mabrouk A., Abdallah N., Dhifaoui Z.** Wavelet decomposition and autoregressive model for time series prediction // Applied mathematics and computation. 2008. Vol. 199. P. 334—340.
22. **Геппенер В. В., Мандрикова О. В.** Совмещение параметрического и непараметрического подходов к построению моделей нестационарных временных рядов, имеющих сложную структуру, с целью повышения качества их обработки // Известия СПбГЭТУ "ЛЭТИ". Сер. "Информатика, управление и компьютерные технологии". 2003. Вып. 2. С. 14—17.
23. **Jaffard S.** Pointwise smoothness, two-microlocalisation and wavelet coefficients // Publications Mathématiques. 1991. Vol. 35. P. 155—168.
24. **Левин Б. Р.** Теоретические основы статистической радиотехники. Изд. 2-е. М.: Советское радио. 1975. 392 с.

O. V. Mandrikova^{1, 2}, Head of Laboratory, oksanaml@mail.ru,
N. V. Fetisova (Glushkova)¹, Junior Researcher, nv.glushkova@yandex.ru,
Yu. A. Polozov^{1, 2}, Senior Researcher, up_agent@mail.ru

¹ Institute of Cosmophysical Researches and Radio Wave Propagation FEB RAS, Paratunka, Russia

² Kamchatka State Technical University, Petropavlovsk-Kamchatsky, Russia

Modeling of Ionospheric Parameter Time Variations and Detection of Anomalies

A new multicomponent model of ionosphere parameter time variation (MCM) was described. It is based on the combination of wavelets with autoregressive-integrated moving average model class. The model allows us to study ionosphere mode and to determine anomalies occurring during ionospheric disturbances. MCM-models of the ionosphere critical frequency for Kamchatka region (Paratunka station, registration is carried out at the Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS) were developed. Comparison of the obtained MCMs with the Empiric International Reference IRI model and moving median method was performed that showed that the suggested method is prospective. To investigate anomalous changes in the ionosphere in detail, the authors suggested applying continuous wavelet-transform. On the basis of continuous wavelet-transform, computational solutions to detect different scale anomalies in the ionosphere and to estimate their parameters were developed.

The analysis of ionospheric parameters during increased geomagnetic activity showed that ionospheric anomalies have large scales and achieve the maximum intensity during the strongest magnetic storms. It was noticed that local increase of electron concentration may be observed in the analyzed region before the main phase of a magnetic storm.

Keywords: wavelet-transform, autoregressive-integrated moving average model, ionosphere critical frequency, ionospheric disturbances

References

1. Afrajmovich Je. L., Perevalova N. P. *GPS-monitoring verhnjej atmosfery Zemli*. Irkutsk: GU NU RVH VSNC SO RAMN, 2006. 480 p.
2. Kato H. Development of automatic scaling software of ionospheric parameters. *Journal of the National Institute of Information and Communications Technology*, 2009, vol. 56, pp. 465–474.
3. Nakamura M., Maruyama T., Shidama Y. Using a neural network to make operational forecasts of ionospheric variations and storms at Kokubunji, Japan. *Journal of the National Institute of Information and Communications Technology*, 2009, vol. 56, pp. 391–406.
4. Klimenko M. V. Modeling of local disturbance formation in the ionosphere electron concentration before strong earthquakes. *Earth, Planets and Space*, 2012, vol. 64, pp. 6, pp. 441–450.
5. Klimenko M. V. Numerical modeling of the global ionospheric effects of storm sequence on September 9–14, 2005 — comparison with IRI model. *Earth, Planets and Space*, 2012, vol. 64, no. 6, pp. 433–440.
6. Bilitz D., Reinisch B. W. International Reference Ionosphere 2007: Improvements and new parameters. *Advances in space research*, 2008, vol. 42, pp. 599–609.
7. Mandrikova O. V., Polozov Ju. A. Metod vydelenija anomal'nyh osobennostej v dannyh kriticheskoj chastoty ionosfery na osnove sovmeshhenija vejvlet-preobrazovanija i avtoregionnyh setej, *Cifrovaja obrabotka signalov*, 2012, no. 2, pp. 29–35.
8. Mandrikova O. V., Polozov Ju. A. Approksimacija i analiz ionosfernyh parametrov na osnove sovmeshhenija vejvlet-preobrazovanija s kollektiyami nejronnyh setej, *Informacionnye tehnologii*, 2014, no. 7, pp. 61–65.
9. Mandrikova O. V., Glushkova N. V., Polozov Ju. A. Algoritmy vydelenija i analiza anomalij v parametrah kriticheskoj chastoty ionosfery foF2 na osnove sovmeshhenija vejvlet-preobrazovanija i avtoregionnyh modelej, *Cifrovaja obrabotka signalov*. 2013, no. 1, pp. 47–53.
10. Mandrikova O. V., Glushkova N. V., Zhivet'ev I. V. Modeling and analysis of ionospheric parameters by a combination of wavelet transform and autoregressive models, *Geomagnetism and Aeronomy*, 2014, vol. 54, no. 5, pp. 593–600, DOI: 10.1134/S0016793214050107.
11. Zaourar N. Wavelet-based multiscale analysis of geomagnetic disturbance, *Earth Planets Space*, 2013, vol. 65, no. 12, pp. 1525–1540.
12. Yu Z. Modeling and simulation of the horizontal component of the geomagnetic field by fractional stochastic differential equations in conjunction with empirical mode decomposition, *Journal of geophysical research*, 2010, vol. 115, no. A10. DOI: 10.1029/2009JA015206.
13. Klionsky D. M., Oreshko N. I., Geppener V. V. Applications of empirical mode decomposition for processing nonstationary signals, *Pattern recognition and image analysis*, 2008, vol. 13, no. 3, pp. 390–399.
14. Turovskij A. Ja., Kurgalin S. D., Semenov A. G. Issledovanie dinamiki maksimumov lokal'nyh vejvlet-spektpov vyzvannyh zpitel'nyh potencialov golovnogo mozga, *Informacionnye tehnologii*, 2013, no. 10, pp. 46–50.
15. Mallat S. *A wavelet tour of signal processing*. London: Academic Press, 1999. 637 p.
16. Daubechies I. Ten lectures on wavelets. CBMS-NSF Lecture Notes nr. 61, SIAM. 1992. 377 p.
17. Chui C. K. *An introduction to wavelets*. Boston: Academic Press, 1992. 264 p.
18. Mandrikova O. V., Solov'ev I. S. Vejvlet-tehnologija variacij magnitnogo polja Zemli, *Informacionnye tehnologii*, 2011, no. 1, pp. 34–38.
19. Box G., Jenkins G. *Time series analysis: Forecasting and control*. San Francisco: Holden-Day, 1970. 553 p.
20. Marple S. *Digital spectral analysis with applications*. New-Jersey: Prentice-Hall, 1987. 492 p.
21. Mabrouk A., Abdallah N., Dhifaoui Z. Wavelet decomposition and autoregressive model for time series prediction, *Applied mathematics and computation*, 2008, vol. 199, pp. 334–340.
22. Geppener V. V., Mandrikova O. V. Sovmeshhenie parametricheskogo i neparametricheskogo podhodov k postroeniju modelej nestacionarnykh vremennykh rjadov, imejushhii slozhnuju strukturu, s cel'ju povyshenija kachestva ih obrabotki, *Izvestija SPbGJeTU "LJeTI". Ser. "Informatika, upravlenie i komp'juternye tehnologii"*, 2003, vol. 2, pp. 14–17.
23. Jaffard S. Pointwise smoothness, two-microlocalisation and wavelet coefficients, *Publications Mathématiques*, 1991, vol. 35, pp. 155–168.
24. Levin B. R. *Teoreticheskie osnovy statisticheskoj radiotehniki*. Izd. 2-e, Moscow, Sovetskoe radio, 1975. 392 p.

Н. Н. Светушков, канд. техн. наук, доц., e-mail: svt.n.n@mail.ru,
Московский авиационный институт (национальный исследовательский университет),
факультет "Прикладная математика и физика"

Применение интегрального метода струн для расчета тепловых нагрузок на охлаждаемые стенки сопловой части ракетного двигателя

В настоящее время имеется большое количество как коммерческих, так и находящихся в свободном доступе программных продуктов, позволяющих проводить тепловые расчеты. Наиболее известным среди них является пакет ANSYS, который часто применяется на промышленных предприятиях. Однако часто основным недостатком таких расчетов является большая погрешность вычислений, доходящая до 50 %. Поэтому часто полученные таким образом численные результаты несут лишь иллюстративную информацию о происходящих тепловых процессах без реальной оценки точности расчетов. В данной работе для проведения вычислений применялся новый подход, названный методом струн и основанный на интегральном описании процесса теплопередачи. Такой подход позволил контролировать погрешность решения в каждой расчетной точке, используя в качестве критерия функцию невязки или рассогласования. Целью выполненных модельных расчетов являлось, кроме получения численных результатов, верификация работоспособности предложенного алгоритма в условиях высоких градиентов температур и сложной геометрической конфигурации изделия, а также тестирование на реальной практической задаче разработанного программного обеспечения.

В работе исследуются возможности расчета тепловых нагрузок на стенки сопловой части ракетного двигателя. Выбор задачи для проведения расчетов связан с актуальностью тепловых расчетов при создании новых как авиационных, так и ракетных двигателей. При конструировании ракетных двигателей остро стоит вопрос об эффективности охлаждения камеры сгорания и сопловой части двигателя с тем, чтобы не допустить прогорания стенок изделия, а также обеспечить необходимый ресурс работоспособности в случае экстремальных внешних тепловых воздействий. В силу указанных выше причин при моделировании таких задач особую важность представляет оценка точности получаемых результатов, именно поэтому для ее решения и был выбран метод струн. В разработанной программной среде, использованной при расчетах, была сформирована геометрическая модель, заданы начальные и граничные условия и проведены расчеты для двух типоразмеров конструкции. Проведенные расчеты показали достаточную эффективность данного подхода для моделирования такого рода тепловых процессов, учитывая возможность анализа точности результатов и относительно высокую скорость сходимости. Созданный программный комплекс может быть использован при разработке программного обеспечения для виртуального проектирования конструкций ракетных и авиационных двигателей.

Ключевые слова: тепловая нагрузка, уравнение теплопроводности, математическое моделирование, интегральные уравнения, программное обеспечение, численные результаты

Введение

В настоящее время имеется большое количество как коммерческих, так и находящихся в свободном доступе программных продуктов, позволяющих проводить тепловые расчеты. Наиболее известным среди них является пакет ANSYS, который часто применяется на промышленных предприятиях. Однако часто основным недостатком таких расчетов является большая погрешность вычислений. Поэтому зачастую численные результаты несут лишь иллюстративную информацию о происходящих тепловых процессах (например, при моделировании работы газотурбинных двигателей точность расчетов считается приемлемой, если полученные результаты лишь качественно согласуются с экспериментом).

В данной работе для проведения вычислений применен новый подход, названный методом струн и основанный на интегральном описании процесса теплопередачи. Такой подход позволил контроли-

ровать погрешность решения в каждой расчетной точке, используя в качестве критерия функцию невязки или рассогласования. Целью выполненных модельных расчетов являлось, кроме получения численных результатов, верификация работоспособности предложенного алгоритма в условиях высоких градиентов температур, а также тестирование на реальной практической задаче разработанного программного обеспечения.

Программный комплекс выполнен в виде отдельного интерактивного программного средства на языке Visual C++ и имеет довольно гибкие возможности по формированию сложных двумерных геометрических объектов, генерации сеточных разбиений, а также заданию граничных условий — включая граничные условия первого, второго и третьего рода (тепловые потоки, описываемые законом Ньютона).

Выбор задачи для проведения расчетов был связан с актуальностью тепловых расчетов при создании

Математическая модель

Численное моделирование прогрева охлаждаемой конструкции основано на решении классического двумерного уравнения теплопроводности с граничными условиями второго рода, когда на границе задаются тепловые потоки:

$$c\rho(T)\frac{\partial T}{\partial t} = \frac{\partial}{\partial x} h_x(T)\frac{\partial T}{\partial x} + \frac{\partial}{\partial y} h_y(T)\frac{\partial T}{\partial y};$$

$$T(x, y, 0) = T_0(x, y); (x, y) \in \Omega; \quad (1)$$

$$q|_{(x,y) \in \Gamma(\Omega)} = q_{\Gamma}(t),$$

где $T = T(x, y, t)$ — распределение температурного поля; $T_0(x, y)$ — начальное распределение температур; $c\rho(T)$ — произведение удельной теплоемкости и плотности материала, зависящее как от температуры, так и от пространственных координат; $h_{x,y}(T)$ — коэффициенты теплопроводности по соответствующим осям, также зависящие от температуры; Ω — область решения задачи; $\Gamma(\Omega)$ — граничная поверхность; $q_{\Gamma} = q_{\Gamma}(t)$ — меняющиеся со временем плотности тепловых потоков на границе области.

Общий геометрический вид и обозначения для охлаждаемой части исследуемой части конструкции приведены на рис. 2, где введены следующие обозначения: e — ширина канала охлаждения; s — толщина ребра; h — высота канала охлаждения; $\delta_{ст}$ — толщина внутренней "горячей" стенки; $\delta_{ст.нар}$ — толщина внешней "холодной" стенки; q_{Γ} — плотность теплового потока от продуктов сгорания; $q_{охл}$ — плотность теплового потока в охлаждающую жидкость.

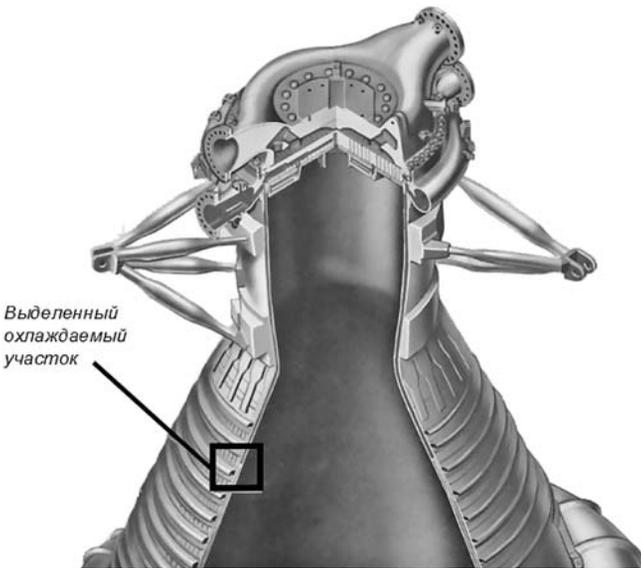


Рис. 1. Сопловая часть ракетного двигателя с системой охлаждения

новых как авиационных, так и ракетных двигателей. При конструировании ракетных двигателей остро стоит вопрос об эффективности охлаждения камеры сгорания и сопловой части двигателя (рис. 1) с тем, чтобы не допустить прогорания стенок изделия, а также обеспечить необходимый ресурс работоспособности в случае экстремальных внешних тепловых воздействий. На практике эта задача часто решается методом проб и ошибок, основываясь на большом опыте инженеров и решениях, используемых в других, созданных ранее конструкциях.

Варьируемыми параметрами при этом являются толщина внутренней стенки сопла, наличие или отсутствие теплозащитного покрытия, скорость протекания и температура охлаждающей жидкости, сечений охлаждающих трубок и др. Проведение экспериментальных исследований для изучения тепловых нагрузок в этих условиях сдерживается большой стоимостью стендового оборудования, необходимостью привлечения высококвалифицированных специалистов, трудоемкостью и большими временными затратами. Методы математического моделирования [1–5] позволяют уже на предварительном этапе разработки двигателя отказаться от неэффективных конструктивных решений, которые не могут обеспечить требуемых рабочих характеристик изделия. В силу указанных выше причин при моделировании таких задач особую важность представляет оценка точности получаемых результатов, именно поэтому для ее решения и был выбран метод струн.

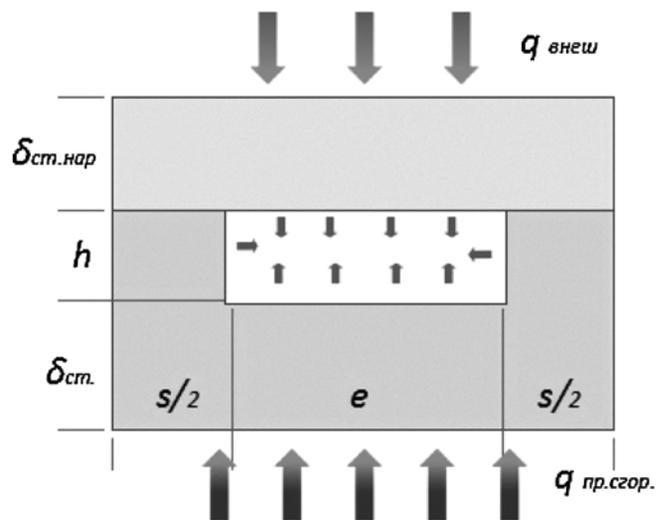


Рис. 2. Расчетная схема участка охлаждаемой части конструкции

Таблица 1

**Размерные и тепловые характеристики
для двух модельных типоразмеров**

Наименование	Типоразмеры	
	I	II
Толщина внутренней стенки, мм ($\delta_{ст}$)	1,0	1,8
Толщина наружной стенки, мм ($\delta_{ст.нар}$)	0,8	0,8
Ширина канала охлаждения, мм	3,51	4,62
Высота канала охлаждения, мм	5,7	4,9
Толщина ребра, мм	1,35	1,35
Коэффициент теплоотдачи от продуктов сгорания в камере, Дж/($m^2 \cdot c \cdot K$) (α_T)	1500,3	895,5
Температура газа в пристеночном слое, К (T_T)	3184	3290
Коэффициент теплоотдачи к охладителю, Дж/($m^2 \cdot c \cdot K$) ($\alpha_{охл}$)	17 793,4	14 421,2
Температура охладителя, К ($T_{охл}$)	343	334
Мощность внешнего излучения, Дж/($m^2 \cdot c$)	$1,26 \cdot 10^6$	$1,26 \cdot 10^6$

Таблица 2

Теплофизические свойства материалов конструкции

Параметры	Основная часть	Внешняя стенка
Коэффициент теплопроводности, Вт/(м · К)	17,8	48
Удельная теплоемкость, кДж/(кг · К)	0,546	0,588
Плотность, кг/м ³	7800	8900

На боковых гранях заданы адиабатические условия, так как в силу симметрии тепловые потоки отсутствуют. Плотности тепловых потоков по абсолютному значению на охлаждаемой стенке и горячей стенке описываются теплообменом по закону Ньютона:

$$q_{охл} = \alpha_{охл}(T - T_{охл});$$

$$q_T = \alpha_T(T_T - T),$$

где коэффициенты теплоотдачи α_T и $\alpha_{охл}$, а также внешние температуры $T_{охл}$ и T_T приведены в табл. 1.

В табл. 2 приведены теплофизические характеристики жаропрочного сплава, из которого состоит конструкция.

Для корректного решения поставленной задачи необходимо также учитывать тепловые потоки со стенок сопловой части, обусловленные собственным излучением стенок, однако, как показывает предварительный анализ, их значение становится существенным при температурах выше 1000 К, да и в этом случае по абсолютному значению они составляют порядка 10^5 Дж/($m^2 \cdot c$), что более чем на

порядок меньше, чем поток от продуктов сгорания в сопле, а также на порядок меньше, чем внешнее тепловое воздействие. Таким образом, влиянием теплового излучения со стенок конструкции в данной постановке можно пренебречь.

Численные результаты

Как уже говорилось выше, для проведения расчетов был использован разработанный программный комплекс [6–9], позволяющий формировать произвольную двумерную геометрическую область для описания исследуемого объекта и задавать начальные и граничные условия. В качестве расчетного метода использован метод "струн" [10], основанный на интегральном представлении процессов теплопередачи. Основным преимуществом этого метода является возможность в каждой точке сеточного разбиения вычислить вектор рассогласования или невязки, что и является действительно надежным способом оценки погрешности численного решения. Все вычисления проводились итерационно до достижения точности 1 % на всей области решения задачи.

Оба исследуемых модельных варианта (два типоразмера) были представлены одинаковой областью решения размером 6×10 мм и сеточным разбиением с $50 \times 90 = 4500$ точками. Временной шаг был выбран равным 0,1 с. В расчетах число временных слоев составляло 500, т.е. рассчитывались температурные поля в течение 50 с работы двигателя. Для получения надежных результатов на каждом временном слое использовалась итерационная процедура, ее сходимость оценивалась по вектору невязки в каждой расчетной точке, условием останова являлось достижение точности более 1 %.

На рис. 3 (см. третью сторону обложки) показана размерная геометрическая модель, сформированная в программе, и расположение расчетных узлов (для типоразмера II). На скрин-шоте, представляющем окно создания модели, видно, что в левой и верхней частях расположены линейки, на которых отображаются размеры конструкции в метрах.

Более ярким цветом выделена внешняя стенка, контактирующая с окружающей средой и имеющая теплофизические характеристики, отличные от характеристик материала основной конструкции.

Целью проводимых расчетов являлось выяснение температуры нагреваемой поверхности в целях недопущения ее прогорания и разрушения, так как температура газа в пристеночном слое имеет достаточно высокие значения (см. табл. 1). Материал основной конструкции представляет собой жаропрочный сплав (см. табл. 2), имеющий температуру плавления около 1000 К.

Проведенные расчеты показали, что уже через 2 с после начала работы двигателя происходит выход

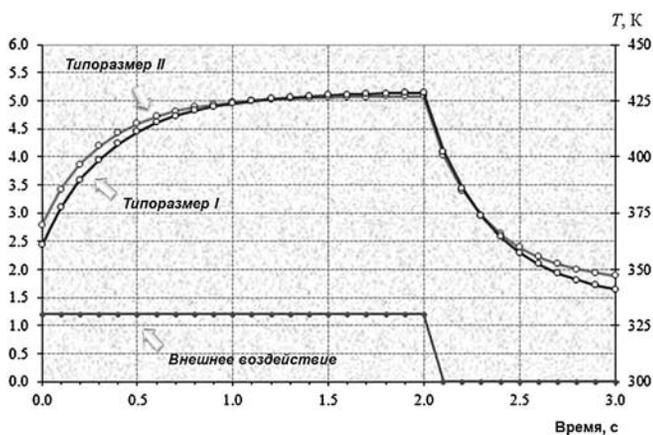


Рис. 5. Изменение температуры внешней поверхности при импульсном тепловом воздействии (для двух типоразмеров)

на стационарный режим прогрева конструкции по толщине, т.е. температурное поле практически не меняется (рис. 4, см. третью сторону обложки).

Для обоих типоразмеров численные результаты, характеризующие распределение температурного поля по толщине конструкции, отличались незначительно и для температуры стенки, контактирующей с продуктами сгорания, составляли ~650...710 К. Внешняя стенка разогревалась до температуры ~335...345 К.

Дополнительная задача, которая была поставлена заказчиком, — это оценка надежности конструкции в условиях интенсивного внешнего теплового воздействия. В качестве модели использовался постоянный тепловой поток заданной мощности (см. табл. 1), действующий в течение заданного промежутка времени. Расчеты проводили для условий, когда внешнее тепловое воздействие начинается на 30-й секунде работы двигателя и продолжается 2 с. На рис. 5 показано изменение температуры внешней поверхности с момента начала воздействия и после его снятия.

Расчеты показывают, что при заданной интенсивности значение температуры на внешней поверхности возрастает незначительно и остается приблизительно постоянным в течение всего времени воздействия. Кроме того, внешний тепловой поток практически не оказывает влияния на температуру в области стенки, контактирующей с продуктами сгорания. Как видно на графике, максимальная температура внешней поверхности достигает для обоих типоразмеров ~426 К.

Заключение

Проведенные вычисления показали хорошую сходимость алгоритма метода струн (менее 50 итераций

на каждый временной шаг для достижения точности менее 1 %) и отсутствие осцилляций при больших градиентах температур. Адекватность математической модели, использованной при моделировании, была подтверждена частичным сравнением полученных числовых данных с результатами экспериментов.

Таким образом, предложенный метод расчета и разработанный программный комплекс может быть рекомендован к использованию при моделировании широкого круга теплофизических задач. В отличие от программного комплекса ANSYS в разработанной среде имеется возможность в интерактивном режиме оценивать погрешность в каждой расчетной точке, что позволяет говорить об адекватности получаемых численных результатов.

Разработанный комплекс можно также использовать как независимый модуль при создании общей программной среды для виртуального проектирования конструкций ракетных и авиационных двигателей.

Автор благодарен начальнику отдела, д-ру техн. наук Стернину Л. Е. и начальнику сектора Кондобе Л. И. ОАО "НПО Энергомаш им. академика В. П. Глушко" за полезные обсуждения и замечания по работе.

Список литературы

1. Зарубин В. С. Математическое моделирование в технике. 2-е изд. М.: Изд-во МГТУ им. Н. Э. Баумана, 2003. 497 с.
2. Patankar S. Numerical Heat Transfer and Fluid Flow. New York, Hemisphere Publishing Corporation, 1980. 199 p.
3. Самарский А. А., Вабишевич П. Н. Численные методы решения задач конвекции-диффузии. М.: УРСС, 2009. 248 с.
4. Галицейский Б. М. Теплопередача в авиационных двигателях. М.: Изд-во МАИ, 1985. 82 с.
5. Дрейцер Г. А. Конвективный теплообмен в каналах: учеб. пособие. М.: Изд-во МАИ, 1984. 77 с.
6. Авдеевский В. С., Галицейский Б. М., Глебов Г. А. и др. Основы теплопередачи в авиационной и ракетно-космической технике. М.: Машиностроение, 1975. 624 с.
7. Светушков Н. Н. Моделирование нестационарных тепловых процессов в неоднородных средах: дис. ... канд. техн. наук, М., ВТИ, 2010.
8. Светушков Н. Н. Метод геометрических интегралов при моделировании процессов теплопередачи в задачах с фазовыми превращениями // Вестник Московского авиационного института. 2012. Т. 19, № 5. С. 182—186.
9. Светушков Н. Н. Программный комплекс по расчету нестационарных тепловых полей в сложных двумерных объектах методом геометрических интегралов. Свидетельство о государственной регистрации программ № 2013615984 от 25 июня 2013 года.
10. Светушков Н. Н. Метод струн в задачах многомерной нестационарной теплопроводности // Информационные технологии. 2014. № 12. С. 14—19.

Integral Strings Method for Modeling Heat Loads in the Cooling Wall of the Rocket Engine Nozzle Part

There are currently a large number of both commercial and freely available, software products that allow for thermal calculations. The most famous among them is the package ANSYS, which is frequently used in industrial plants. Often, however, the main drawback of such calculations is the large computational error, reaching up to 50 percent. So often the thus obtained numerical results are only illustrative information on thermal processes taking place without any real assessment of the accuracy of calculations. In this paper for computing a new approach called the method based on strings and integral description of the process of heat transfer. Such an approach control error of the solution in each design point, using as a criterion function of the residual or error. The purpose of the model calculations was performed, in addition to receiving the numerical results, the verification efficiency of the proposed algorithm in high temperature gradients, and complex geometric configuration of the product, as well as testing on a real practical problem of the developed software. This paper investigates the possibility of calculating the thermal loads on the walls of the nozzle of the rocket engine. Select a task to perform the calculations associated with the relevance of thermal calculations for both the creation of new aircraft and rocket engines. When designing rocket engines is an issue about the effectiveness of the cooling of the combustion chamber and nozzle of the engine in order to prevent burning of the walls of the product and to provide the necessary resource efficiency in the case of extreme external thermal effects. Due to the above reasons for modeling of such problems is particularly important to estimate the accuracy of the results, which is why its solution and was chosen method of strings. The developed software environment used for the calculation was generated geometric model, given initial and boundary conditions and calculations for two type-size design. The calculations have shown sufficient efficacy of this approach for modeling of thermal processes of this kind, given the opportunity to analyze the accuracy of the results and the relatively high rate of convergence. Created software package can be used in the development of software for virtual structural design of missile and aircraft engines.

Keywords: numerical methods, software, numerical simulation, the heat equation, integral equations, the accuracy of calculations

References

1. **Zarubin V. S.** *Matematicheskoe modelirovanie v tekhnike*, 2-e izd., Moscow, MG TU im. N. E. Baumana, 2003. 497 p.
2. **Patankar S.** *Numerical Heat Transfer and fluid flow*. New York, Hemisphere Publishing Corporation, 1980. 199 p.
3. **Samarskiy A. A., Vabishchevich P. N.** *Chislennyye metody resheniya zadach konveksii-diffuzii*, Moscow, URSS, 2009, 248 p.
4. **Galitseyskiy B. M.** *Teplotperedacha v aviatsionnye dvigatelyakh*. Moscow, MAI, 1985, 82 p.
5. **Dreytser G. A.** *Konvektivnyy teploobmen v kanalakh: ucheb. posob.* Moscow, MAI, 1984, 77 p.
6. **Avduevskiy V. S., Galitseyskiy B. M., Glebov G. A.** i dr. *Osnovy teplotperedachi v aviatsionnoy i raketno-kosmicheskoy tekhnike*. Moscow, Mashinostroenie, 1975. 624 s.
7. **Svetushkov N. N.** Modelirovanie nestatsionarnykh teplovykh protsessov v neodnorodnykh sredakh; diss. na soiskanie stepeni kandidata tekhnicheskikh nauk, Moscow, VTI, 2010.
8. **Svetushkov N. N.** Metod geometricheskikh integralov pri modelirovanii protsessov teplotperedachi v zadachakh s fazovymi prevrashcheniyami, *Vestnik Moskovskogo aviatsionnogo instituta*, 2012, vol. 19, no. 5, pp. 182–186.
9. **Svetushkov N. N.** *Programmnyy kompleks po raschetu nestatsionarnykh teplovykh poley v slozhnykh dvumernykh ob'ektakh metodom geometricheskikh integralov*, svidetel'stvo o gosudarstvennoy registratsii programm № 2013615984 ot 25 iyunya 2013 goda.
10. **Svetushkov N. N.** Metod strun v zadachakh mnogomernoy nestatsionarnoy teplotprovodnosti, *Informatsionnye tekhnologii*, 2014, no. 12, pp. 14–19.

Поздравляем Главного редактора журнала академика
Александра Леонидовича Стемпковского
с присуждением Премии Правительства Российской Федерации

Дорогой и уважаемый Александр Леонидович!

Редакционный совет, редакционная коллегия, редакция журнала "Информационные технологии", весь коллектив издательства "Новые технологии" сердечно поздравляют Вас с присуждением Премии Правительства Российской Федерации в области науки и техники за 2015 год за разработку и освоение серийного производства широкой номенклатуры микросхем для смарт-карт и радиочастотной идентификации на основе создания комплексных средств автоматизации проектирования и технологий изготовления современных интегральных схем с встроенной энергонезависимой памятью. Эта работа была выполнена под Вашим руководством в содружестве с ОАО "НИИ молекулярной электроники и завод "Микрон" и ООО "Альфачип" и служит примером эффективного сотрудничества академического института, промышленного и частного секторов.

Желаем Вам и всему коллективу, руководимому Вами, новых творческих успехов!

Ваши друзья и соратники
Журнал "Информационные технологии"
Издательство "Новые технологии"

Диспетчеризация заявок кругового типа в Grid-системах

В дальнейшем развитии среды ресурсных прямоугольников определен квадратичный тип отдельной заявки. Для массивов, состоящих из заявок кругового типа, проведено диспетчеризование и вычислены эвристические меры ресурсных оболочек начально-кольцевого, уровневых по высоте и по протяженности, углового уровневого алгоритмов. Для указанных полиномиальных алгоритмов показано сохранение свойства адаптированности на рассматриваемых массивах заявок.

Ключевые слова: квадратичный тип заявки, заявка кругового типа, Grid-система, диспетчеризование, неэвклидова эвристическая мера, полиномиальная трудоемкость алгоритма, начально-кольцевой алгоритм, уровневый алгоритм по высоте, уровневый алгоритм по протяженности, угловой уровневый алгоритм

1. Постановка задачи

В работах автора [1–5] определена среда ресурсных прямоугольников как основа формального аппарата управления распределением вычислительно-временных ресурсов. Grid-системы с централизованной структурой системы диспетчеризования, состоящие из сайтов, содержащих параллельные системы, и имеющие возможность мультисайтного диспетчеризования (выполнения задачи на нескольких сайтах одновременно) в работах [5, 6] моделированы ресурсным квадрантом. Качество диспетчеризования эвристических алгоритмов оценивают неэвклидовой эвристической мерой, учитывающей наряду с площадью и форму занятой ресурсной области. В работах [1–5] предложена и разработана квадратичная классификация множества заявок, а также предложены и исследованы эвристические алгоритмы распределения ресурсов, имеющие полиномиальную трудоемкость, основанные на введенных операциях над ресурсными прямоугольниками в среде ресурсных прямоугольников и адаптированные под соответствующий квадратичный тип массива заявок.

В настоящей работе впервые ставится вопрос о квадратичной классификации одной заявки и об адаптированности полиномиальных алгоритмов для массивов, состоящих из заявок определенного квадратичного типа.

2. Квадратичный тип заявки на обслуживание в Grid-системе

При обслуживании заявок пользователей в Grid-системе выделяют вычислительные задачи

(*computation-intensive*), требующие преимущественно вычислительных ресурсов, и коммуникативные задачи (*data-intensive*), требующие преимущественно коммуникационных ресурсов и памяти для обработки больших объемов данных [7–10].

В зависимости от того, пользователь или система, и при предоставлении задачи или в процессе выполнения, задает число выделяемых процессоров, различают четыре типа задач [11].

Задача может быть с известным заранее временем решения (*known processing time*) или время решения задачи становится известным после ее завершения (*unknown processing time*) [12].

В настоящей работе рассматриваются вычислительные задачи с заранее известным временем решения, в которых число требуемых процессоров определяет пользователь при подаче в систему жесткие задачи (*rigid job*).

При представлении заявки пользователя для обслуживания диспетчером Grid-системы ресурсным прямоугольником горизонтальное и вертикальное измерения, соответственно, принимаются равными числу единиц ресурса времени и числу процессоров, требуемых для обработки. Символом $a(j) \times b(j)$ обозначается j -я заявка, требующая $a(j)$ единиц времени и $b(j)$ единиц процессоров.

В основу определения квадратичного типа ресурсного прямоугольника положим соотношение между ресурсной мерой $a(j)b(j)$ и асимметрией измерений $(b(j) - a(j))^2$.

При $(b(j) - a(j))^2 \leq a(j)b(j)$ ресурсный прямоугольник имеет кинематический (круговой или гиперболический) тип. В зависимости от того $a(j) \geq b(j)$ или $a(j) < b(j)$ ресурсный прямоуголь-

ник относим к круговому или гиперболическому типу, соответственно.

При $(b(j) - a(j))^2 > a(j)b(j)$ ресурсный прямоугольник имеет кинетический (параболический) тип.

Отметим, что из соотношения $(b(j) - a(j))^2 \leq a(j)b(j)$ следует условие отнесения ресурсного прямоугольника к кинематическому типу:

$$(a(j))^2 + (b(j))^2 \leq 3a(j)b(j)$$

или

$$\frac{a(j)}{b(j)} + \frac{b(j)}{a(j)} \leq 3.$$

Заметим, что при минимальном целом $b(j) = 1$ для кругового типа ресурсного прямоугольника $a(j) \geq b(j)$ целочисленное $a(j)$ может иметь только одно из двух значений: $a(j) = 1$ или $a(j) = 2$. Так, прямоугольник 2×1 (*domino* [13]) относится к круговому типу, как и в общем случае [14] прямоугольники $2j \times j$. Тогда как прямоугольник 3×1 (*tromino* [14]), в силу неравенства $3 + 1/3 > 3$ относится к параболическому типу, как и в общем случае [14] прямоугольники $3j \times j$. Приведем примеры массивов, состоящих из элементов кругового типа:

- последовательные квадраты (*consecutive squares*) от 1×1 до $k \times k$ [5, 15, 16];
- элементы парtridge (*partridge*), содержащие один квадрат 1×1 , два квадрата 2×2 , ..., k квадратов $k \times k$ [17–20];
- элементы квадрирования квадрата, диссекции (*dissection*) [21–23].

Примером массивов, состоящих из элементов гиперболического типа, являются последовательные прямоугольники (*consecutive rectangles* [5, 15, 16]) или почти квадраты (*almost square* [24]) 1×2 , 2×3 , ..., $k \times (k + 1)$ [5, 15, 16, 24].

3. Диспетчеризация массива, состоящего из заявок кругового типа

Вычислим эвристические меры ресурсных оболочек, получаемых при диспетчеризации массивами из ресурсных прямоугольников кругового типа полиномиальными алгоритмами [5]: начально-кольцевым, уровневыми по высоте и по протяженности, угловым уровневым.

Для массива из ресурсных прямоугольников кругового типа $2j \times j$, $j = 1, 2, \dots, k$ при $k = 32$ соответствующие построения начально-кольцевым алгоритмом приведены на рис. 1. В центре прямоугольника указано значение его вертикального измерения.

Эвристические меры ресурсных оболочек начально-кольцевого алгоритма для массива из ресурсных прямоугольников кругового типа приведены в табл. 1.

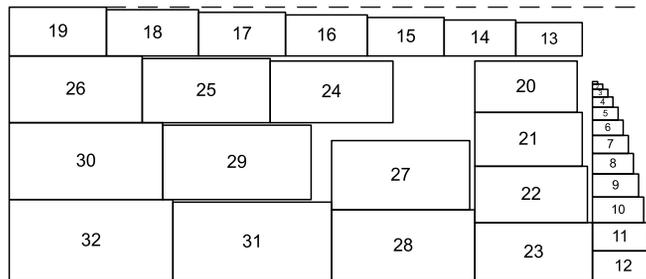


Рис. 1. Укладка начально-кольцевым алгоритмом массива из ресурсных прямоугольников кругового типа

Таблица 1

Эвристические меры ресурсных оболочек начально-кольцевого алгоритма для массива из ресурсных прямоугольников кругового типа

k	Эвристическая мера	k	Эвристическая мера	k	Эвристическая мера
18	1,17	23	1,03	28	1,08
19	1,14	24	1,01	29	1,07
20	1,11	25	0,98	30	1,08
21	1,08	26	1	31	1,06
22	1,06	27	1,07	32	1,05

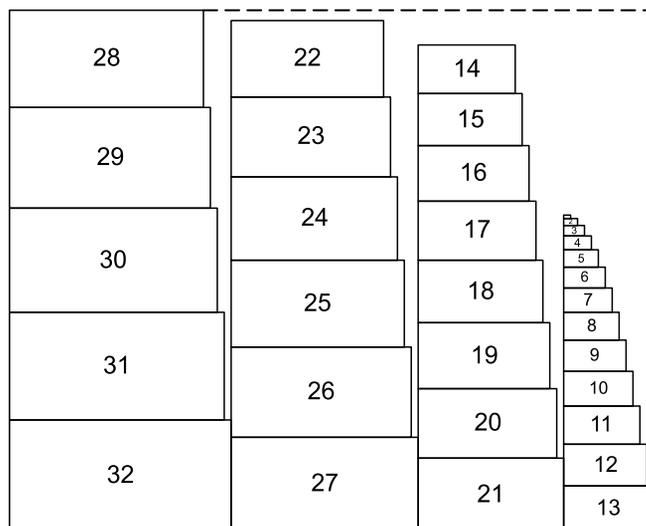


Рис. 2. Укладка уровневым алгоритмом по высоте массива из ресурсных прямоугольников кругового типа

Таблица 2

Эвристические меры ресурсных оболочек уровневого алгоритма по высоте для массива из ресурсных прямоугольников кругового типа

k	Эвристическая мера	k	Эвристическая мера	k	Эвристическая мера
18	0,71	23	0,69	28	0,71
19	0,66	24	0,71	29	0,69
20	0,64	25	0,71	30	0,69
21	0,63	26	0,71	31	0,68
22	0,68	27	0,71	32	0,64

Видим, что эвристические меры ресурсных оболочек начально-кольцевого алгоритма для массива из прямоугольников кругового типа не превосходят значения $0,5 + 0,67$.

Для массива из ресурсных прямоугольников кругового типа $2j_1 \times j_1, j_1 = 1, 2, \dots, k$ при $k = 32$ соответствующие построения уровневый алгоритмом по высоте приведены на рис. 2.

Эвристические меры ресурсных оболочек уровневого алгоритма по высоте для массива из ресурсных прямоугольников кругового типа приведены в табл. 2.

Видим, что эвристические меры ресурсных оболочек уровневого алгоритма по высоте для массива из прямоугольников кругового типа не превосходят значение $0,5 + 0,21$.

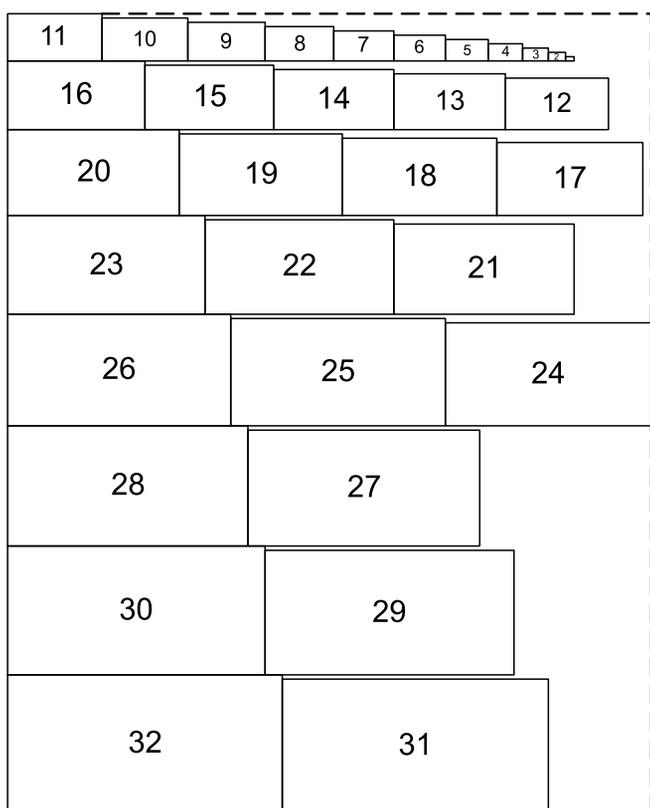


Рис. 3. Укладка уровневый алгоритмом по протяженности массива из ресурсных прямоугольников кругового типа

Таблица 3

Эвристические меры ресурсных оболочек уровневого алгоритма по протяженности для массива из ресурсных прямоугольников кругового типа

k	Эвристическая мера	k	Эвристическая мера	k	Эвристическая мера
18	0,81	23	0,63	28	0,67
19	0,70	24	0,63	29	0,68
20	0,84	25	0,65	30	0,64
21	0,69	26	0,66	31	0,64
22	0,64	27	0,66	32	0,64

Для массива из ресурсных прямоугольников кругового типа $2j_1 \times j_1, j_1 = 1, 2, \dots, k$ при $k = 32$ соответствующие построения уровневый алгоритмом по протяженности приведены на рис. 3.

Эвристические меры ресурсных оболочек уровневого алгоритма по протяженности для массива из ресурсных прямоугольников кругового типа приведены в табл. 3.

Видим, что эвристические меры ресурсных оболочек уровневого алгоритма по протяженности для массива из прямоугольников кругового типа не превосходят значения $0,5 + 0,34$.

Для массива из ресурсных прямоугольников кругового типа $2j_1 \times j_1, j_1 = 1, 2, \dots, k$ при $k = 32$ соответствующие построения угловым уровневый алгоритмом приведены на рис. 4.

Эвристические меры ресурсных оболочек углового уровневого алгоритма для массива из ресурсных прямоугольников кругового типа приведены в табл. 4.

Видим, что эвристические меры ресурсных оболочек углового уровневого алгоритма для массива из прямоугольников кругового типа не превосходят значения $0,5 + 0,22$.

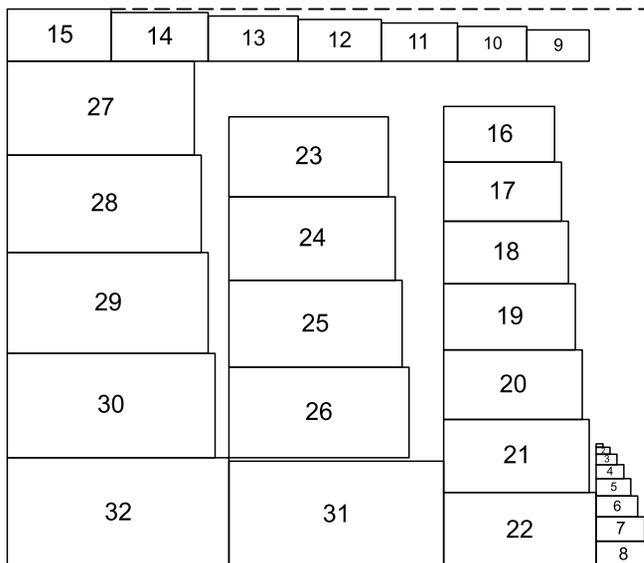


Рис. 4. Укладка угловым уровневый алгоритмом массива из ресурсных прямоугольников кругового типа

Таблица 4

Эвристические меры ресурсных оболочек углового уровневого алгоритма для массива из ресурсных прямоугольников кругового типа

k	Эвристическая мера	k	Эвристическая мера	k	Эвристическая мера
18	0,68	23	0,68	28	0,65
19	0,72	24	0,67	29	0,63
20	0,71	25	0,66	30	0,62
21	0,69	26	0,64	31	0,64
22	0,69	27	0,63	32	0,67

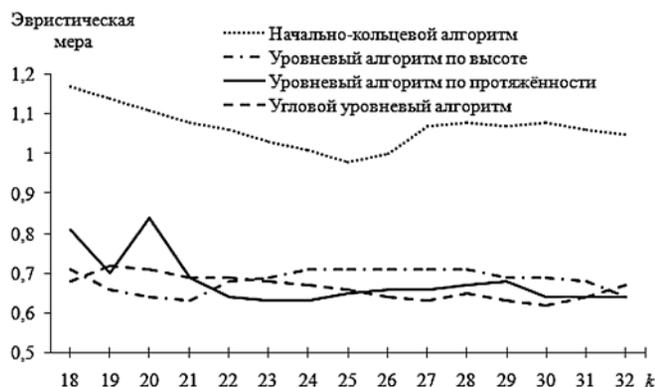


Рис. 5. Эвристические меры ресурсов оболочек полиномиальных алгоритмов диспетчеризации массивами из ресурсов прямоугольников кругового типа

Графики эвристической меры ресурсов оболочек начально-кольцевого, уровневых по высоте и по протяженности, углового уровневого алгоритмов диспетчеризации массива из ресурсов прямоугольников кругового типа показаны на рис. 5.

Видим, что начально-кольцевой алгоритм имеет большую эвристическую меру, чем другие исследуемые алгоритмы. Сравнение результатов эвристических мер ресурсов оболочек углового уровневого, уровневых по высоте и по протяженности алгоритмов не позволяет отдать предпочтение какому-либо из них. Целесообразность использования предложенных алгоритмов при диспетчеризации массивами ресурсов прямоугольников кругового типа подтверждается приемлемым значением отклонения от минимально возможного значения эвристической меры, равного $1/2$, достигаемого при беспустотной укладке в квадрат.

Заключение

В развитии формального аппарата неевклидовой среды ресурсов прямоугольников, как основы теории полиномиальной диспетчеризации, дается определение квадратичного типа одной заявки. Результаты исследования показали адаптированность ранее предложенных автором полиномиальных алгоритмов уровневых по высоте и по протяженности, углового уровневого для массивов, состоящих из заявок кругового типа. Таким образом, показана целесообразность применения данных алгоритмов в Grid-системах с централизованной структурой системы диспетчерирования и мультисайтным режимом обслуживания заявки в случаях, когда известен квадратичный тип каждой заявки.

Список литературы

1. Саак А. Э. Локально-оптимальные ресурсные распределения // Информационные технологии. 2011. № 2. С. 28–34.

2. Саак А. Э. Алгоритмы диспетчеризации в Grid-системах на основе квадратичной типизации массивов заявок // Информационные технологии. 2011. № 11. С. 9–13.

3. Саак А. Э. Диспетчеризация в Grid-системах на основе однородной квадратичной типизации массивов заявок пользователей // Информационные технологии. 2012. № 4. С. 32–36.

4. Саак А. Э. Сравнительный анализ полиномиальных алгоритмов диспетчеризации в Grid-системах // Информационные технологии. 2012. № 9. С. 28–32.

5. Саак А. Э. Полиномиальные алгоритмы распределения ресурсов в Grid-системах на основе квадратичной типизации массивов заявок // Информационные технологии. 2013. № 7. Приложение. 32 с.

6. Саак А. Э. Управление ресурсами и заявками пользователей в Grid-системах с централизованной архитектурой // Труды XII Всероссийского совещания по проблемам управления ВСПУ-2014. Москва, 16–19 июня 2014 г. М.: Институт проблем управления им. В. А. Трапезникова РАН, 2014. С. 7489–7498.

7. Foster I., Kesselman C. The Grid in a nutshell. In: Nabrzyski J., Schopf J., Weglarz J. (eds.) Grid Resource Management: state of the art and future trends. Kluwer, 2003.

8. Jacob B., Brown M., Fukui K., Trivedi N. Introduction to grid computing. IBM Corp., USA, 2005.

9. Christodoulopoulos K., Sourlas V., Mpakolas I., Varvarigos E. A comparison of centralized and distributed meta-scheduling architectures for computation and communication tasks in Grid networks // Computer Communications. 2009. N. 32. P. 1172–1184.

10. Rahman M., Ranjan R., Buyya R., Benatallah B. A taxonomy and survey on autonomic management of applications in grid computing environments // Concurrency Computat: Pract. Exper. 2011. N. 23. P. 1990–2019.

11. Feitelson D., Rudolph L. Toward convergence in job schedulers for parallel supercomputers // Lecture Notes in Computer Science. 1996. Vol. 1162. P. 1–26.

12. Ye D., Zhang G. On-Line Scheduling of Parallel Jobs // LNCS. Vol. 3104. 2004. P. 279–290.

13. Klarner D., Pollack J. Domino tilings of rectangles with fixed width // Discrete Mathematics. 1980. N. 32. P. 45–52.

14. <http://www.mathpuzzle.com/4Dec2001.htm>

15. Korf R., Moffitt M., Pollack M. Optimal rectangle packing // Annals of Operations Research. 2010. Vol. 179, N. 1. P. 261–295.

16. Huang E., Korf R. Optimal rectangle packing: an absolute placement approach // Journal of Artificial Intelligence Research. 2012. Vol. 46. P. 47–87.

17. Ahearn S., Jepsen C. On tiling an $m \times m$ square with m squares // Crux Mathematicorum. 1993. Vol. 19, N. 7. P. 189–191.

18. Simonis H., O’Sullivan B. Search strategies for rectangle packing. 2008. <http://4c.ucc.ie/~hsimonis/cp2008slides.pdf>

19. Ågren M., Beldiceanu N., Carlsson M., Sbihi M., Truchet C., Zampelli S. Six ways of integrating symmetries within non-overlapping constraints // LNCS. 2009. Vol. 5547. P. 11–25.

20. Hougardy S. A Scale Invariant Exact Algorithm for Dense Rectangle Packing Problems. Research Institute for Discrete Mathematics, University of Bonn, Report No 101020, 2012.

21. Duijvestijn A. J. W. Simple perfect squared square of lowest order. J. Combin. Theory. 1978. Ser. B 25. P. 240–243.

22. Duijvestijn A. J. W., Federico P. J., Leeuw P. Compound perfect squares // Amer. Math. Monthly. 1982. Vol. 89. P. 15–32.

23. Gambini I. A method for cutting squares into distinct squares. Discrete Applied Math. 1999. Vol. 98. P. 65–80.

24. Simonis H., O’Sullivan B. Almost square packing // LNCS. 2011. Vol. 6697. P. 196–209.

Circular-Typed Multiprocessor Tasks Scheduling in Grid Systems

Developing formal instrumentation of the Non-Euclidean resource rectangles environment, as a base for the polynomial scheduling theory, we denote a quadratic type of a user task. In the paper it is raised the question of the adaptedness characteristic of polynomial algorithms for the arrays which consist of user tasks of definite quadratic type. The model of a task is a resource rectangle. It is given the quadratic classification of the following resource rectangles which are used as the test examples: rectangles $2j \times j$ (domino) are circular-typed, rectangles $1 \times 2, 2 \times 3, \dots, k \times (k + 1)$ (almost square) relates to the hyperbolic type, rectangles $3j \times j$ (tromino) are ones of parabolic type. For a model of the set of multiprocessor tasks in the form of extensive linear polyhedral of resource rectangles and for a model of the Grid system with centralized architecture which supports multisite task performance in the form of resource quadrant, author solves the problem of allocation with minimization of a resource enclosure heuristic measure. The Non-Euclidean heuristic measure which takes into consideration both area and the form of an occupied resource region, reaches its minimum value at in-square packing without emptiness. It is considered an array which consists of circular-typed resource rectangles. Then scheduling is performed with the use of an initial ring algorithm, level algorithm with the height parameter and the length one, angular-level algorithm and all of them have polynomial complexity. This result in the respective resource enclosures for each scheduling with the heuristic measures calculated. The study of scheduling by the arrays of circular-typed resource rectangles shows that the initial ring algorithm has bigger heuristic measure in comparison with the others. The comparison of the heuristic measures of the resource enclosures derived from the use of the level algorithm with the height "parameter and the length one and angular-level algorithm shows that we can't prefer one to another. Expedience of the use of suggested algorithms in scheduling by arrays of circular-typed resource rectangles is confirmed by the acceptable deviation value from the theoretical minimum of the heuristic measure.

Keywords: user task of quadratic type, circular-typed user task, Grid-system, scheduling, resource management, Non-Euclidean heuristic measure, algorithm of polynomial complexity, initial ring algorithm, level algorithm with the height parameter, level algorithm with the length parameter, angular-level algorithm

References

1. Saak A. Eh. Lokalno-optimalnye resursnye raspredeleniya [Locally optimal resource allocation], *Informatsionniye tekhnologii* [Information Technologies], 2011, no. 2, pp. 28–34.
2. Saak A. Eh. Algoritmy dispatcherizatsii v Grid-sistemakh na osnove kvadrachnoy tipizatsii massivov zayavok [Algorithms scheduling in Grid-based systems for quadratic typing, arrays applications], *Informatsionnye tekhnologii* [Information Technologies], 2011, no. 11, pp. 9–13.
3. Saak A. Eh. Dispatcherizatsiya v GRID-sistemakh na osnove odnorodnoy kvadrachnoy tipizatsii massivov zayavok polzovatelye [Scheduling in GRID-systems on the basis of homogeneous quadratic typing, arrays of user requests], *Informatsionnye tekhnologii* [Information Technologies], 2012, no. 4, pp. 32–36.
4. Saak A. Eh. Sravnitelniy analiz polinomialnykh algoritmov dispatcherizatsii v GRID-sistemakh [Comparative analysis of polynomial algorithms for scheduling in GRID-systems], *Informatsionnye tekhnologii* [Information Technologies], 2012, no. 9, pp. 28–32.
5. Saak A. Eh. Polinomialnye algoritmy raspredeleniya resursov v Grid-sistemakh na osnove kvadrachnoy tipizatsii massivov zayavok [Polynomial algorithms for resource allocation in Grid-based systems for quadratic typing, arrays applications], *Informacionnye tekhnologii* [Information Technologies], Prilozhenie. 2013, no. 7. 32 p.
6. Saak A. Eh. Upravleniye resursami i zayavkami polzovatelye v Grid-sistemakh s centralizovannoy arkhitekturoy [Resource and multi processor task management in Grid system of centralized architecture], *Trudy XII Vserossiyskogo soveshanya po problem upravleniya VSPU-2014. Moskva, 16 iyunya — 19 inmya 2014 g.* [Proceedings of XII all-Russian conference "Control problems" RCCP'2014. Moscow, 16 June — 19 June 2014]. Moscow: Institut problem upravleniya im. V. A. Trapeznikova RAN, 2014, pp. 7489–7498.
7. Foster I., Kesselman C. The Grid in a nutshell. In: Nabrzyski J., Schopf J., Weglarz J. (eds.) *Grid Resource Management: state of the art and future trends*. Kluwer, 2003.
8. Jacob B., Brown M., Fukui K., Trivedi N. *Introduction to grid computing*. IBM Corp., USA, 2005.
9. Christodouloupoulos K., Sourlas V., Mpakolas I., Varvarigos E. A comparison of centralized and distributed meta-scheduling architectures for computation and communication tasks in Grid networks, *Computer Communications*, 2009, no. 32, pp. 1172–1184.
10. Rahman M., Ranjan R., Buyya R., Benatallah B. A taxonomy and survey on autonomic management of applications in grid computing environments, *Concurrency Computat. Pract. Exper.*, 2011, no. 23, pp. 1990–2019.
11. Feitelson D., Rudolph L. Toward convergence in job schedulers for parallel supercomputers, *Lecture Notes in Computer Science*, 1996, vol. 1162, pp. 1–26.
12. Ye D., Zhang G. On-Line Scheduling of Parallel Jobs. *LNCSS*, 2004, vol. 3104, pp. 279–290.
13. Klarner D., Pollack J. Domino tilings of rectangles with fixed width, *Discrete Mathematics*, 1980, no. 32, pp. 45–52.
14. <http://www.mathpuzzle.com/4Dec2001.htm>
15. Korf R., Moffitt M., Pollack M. Optimal rectangle packing, *Annals of Operations Research*, 2010, vol. 179, no. 1, pp. 261–295.
16. Huang E., Korf R. Optimal rectangle packing: an absolute placement approach, *Journal of Artificial Intelligence Research*, 2012, no. 46, pp. 47–87.
17. Ahearn S., Jepsen C. On tiling an $m \times m$ square with m squares, *Crux Mathematicorum*, 1993, vol. 19, no. 7, pp. 189–191.
18. Simonis H., O'Sullivan B. Search strategies for rectangle packing. URL: <http://4c.ucc.ie/~hsimonis/cp2008slides.pdf>
19. Ågren M., Beldiceanu N., Carlsson M., Sbihi M., Truchet C., Zampelli S. Six ways of integrating symmetries within non-overlapping constraints. *LNCSS*, 2009, vol. 5547, pp. 11–25.
20. Hougardy S. *A Scale Invariant Exact Algorithm for Dense Rectangle Packing Problems*, Research Institute for Discrete Mathematics, University of Bonn, Report No: 101020, 2012.
21. Duijvestijn A. J. W. Simple perfect squared square of lowest order, *J. Combin. Theory*, 1978. Ser. B 25, pp. 240–243.
22. Duijvestijn A. J. W., Federico P. J., Leeuw P. Compound perfect squares, *Amer. Math. Monthly*, 1982, vol. 89, pp. 15–32.
23. Gambini I. A method for cutting squares into distinct squares, *Discrete Applied Math*, 1999, vol. 98, pp. 65–80.
24. Simonis H., O'Sullivan B. Almost square packing. *LNCSS*, 2011, vol. 6697, pp. 196–209.

С. В. Поршнев, д-р техн. наук, проф., e-mail: sergey_porshnev@mail.ru,
Д. А. Божалкин, аспирант, e-mail: bozhalkin@mail.ru,
А. С. Копосов, аспирант, e-mail: alexkopas@gmail.com,
 ФГАОУ ВПО "Уральский федеральный университет
 имени первого Президента России Б. Н. Ельцина", г. Екатеринбург

Опыт использования суперкомпьютера для обработки дампов сетевого трафика магистрального интернет-канала

Проведен сравнительный обзор общедоступных кластеров MATLAB и суперкомпьютеров. Выбран вычислительный кластер для обработки дампа сетевого трафика. Выполнена настройка супервычислителя для использования собственного программного обеспечения. Описана технология анализа дампов трафика, передаваемого в магистральном интернет-канале. Представлены результаты, подтверждающие эффективность технологии.

Ключевые слова: суперкомпьютер, кластер, компьютерные сети, потоки трафика, дамп сетевого трафика, извлечение данных, обработка данных, параллельные вычисления в MATLAB, параметры распределения

Введение

Дамп сетевого трафика можно рассматривать как "снимок данных", прошедших через сеть или сетевое устройство за определенный момент времени. Сегодня стандартом записи потока сетевого трафика де-факто является формат *.pcap* (*Packet Capture*). *Pcap*-файл имеет заголовок, содержащий в себе глобальную информацию, за которым записывается информация, содержащаяся в каждом захваченном пакете, или ноль, если в данный временной интервал не было получено ни одного пакета. Формат записи захваченных пакетов зависит от протокола и используемой технологии передачи данных. Общая структура *pcap*-файла представлена на рис. 1.



Рис. 1. Структура *pcap*-файла

Для просмотра содержимого *pcap*-файлов можно использовать, например, программу Wireshark [1], типичное состояние рабочего окна которой в ходе анализа дампа сетевого трафика представлено на рис. 2.

Из рис. 2 видно, что файл дампа сетевого трафика содержит в себе исчерпывающую информацию (как текстовую, так и численную) о данных, передаваемых в сети Интернет. Следовательно, для проведения количественного анализа данных дампа интернет-трафика необходим предварительный полный семантический анализ (парсинг) *pcap*-файлов, призванный извлечь из них необходимую количественную информацию, используемую далее для получения количественных оценок характеристик трафика.

Отметим, что здесь помимо проблемы парсинга *pcap*-файлов существует проблема передачи извлекаемой количественной информации в программные инструменты, используемые для обработки количественной информации. Технологию, основанную на записи/чтении в/из файла, в связи с большими объемами извлекаемой из дампа информации вследствие низкой скорости ее работы использовать не удалось. При этом понятно, что обмен данными между семантическим анализатором и программой обработки будет более эффективным, если реализовать прямую передачу количественной информации непосредственно в пространство данных программного средства для количественного анализа экспериментальной информации. Технология, в которой реализован данный подход, а в качестве программы обработки используется пакет MATLAB, описана в работе [2].

Однако опыт анализа ежедневных пятнадцатиминутных дампов трафика, передаваемого в магистральном канале между США и Японией (средний объем пятнадцатиминутного дампа архива MAWI (Measurement and Analysis on the WIDE Internet) [3] составляет около 10 Гбайт) показал, что такой анализ требует значительных временных затрат. Здесь исходный *pcap*-файл делится на файлы размером по 100 Мбайт. Далее независимо проводится семантический анализ каждого из файлов меньшего объема. Выбранные в ходе парсера количественные показатели передаются в рабочее пространство MATLAB для дальнейшей обработки на одноядерном процессоре с тактовой частотой 3 ГГц. Среднее время обработки пяти дампов размером 100 Мбайт составляет около 9 ч. Таким образом, примерное время обработки всего пятнадцатиминутного дампа на данном компьютере составит примерно 960 ч.

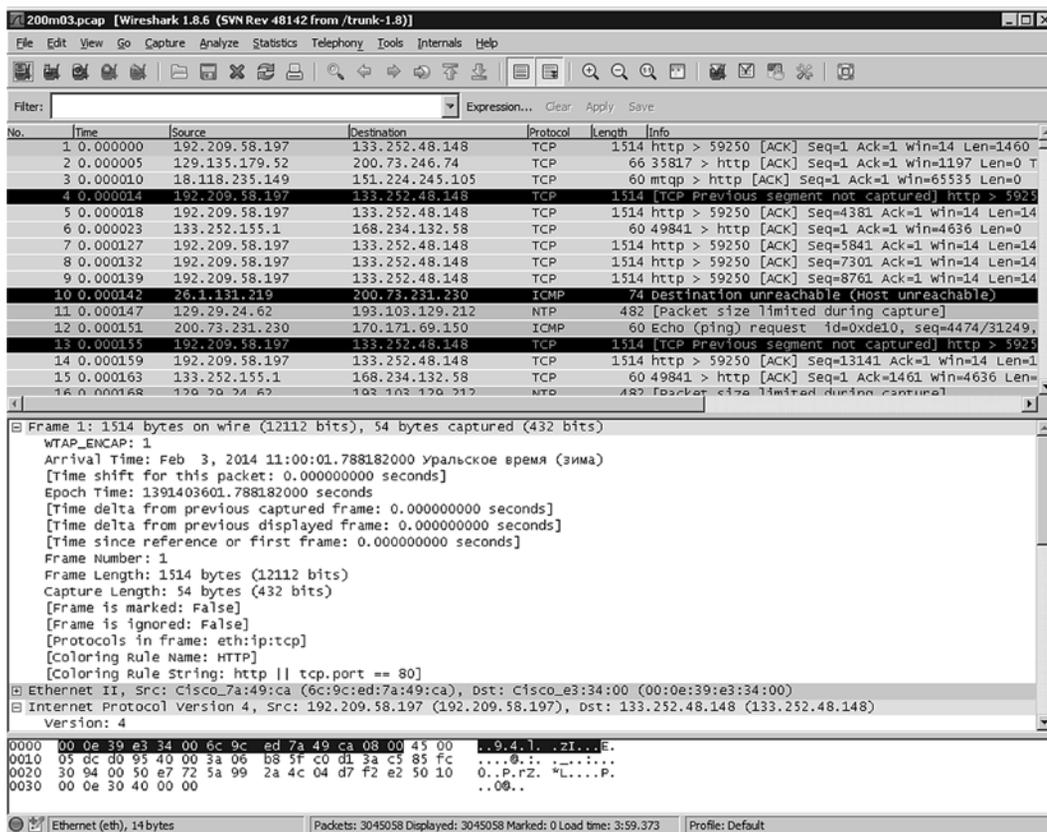


Рис. 2. Состояние рабочего окна программы Wireshark в ходе анализа дампа сетевого трафика

В связи с этим задача переноса описанной выше технологии обработки дампа сетевого трафика на суперкомпьютер является актуальной. Действительно, априори, можно ожидать, что использование многоядерных процессоров и кластеров в сочетании с технологиями параллельных вычислений¹, в том числе и суперкомпьютеров, в зависимости от размера обрабатываемых данных, числа доступных ядер и программной реализации обработчика позволит сократить время обработки в десятки и сотни раз.

В данной статье описан опыт использования суперкомпьютера "Уран", находящегося в Институте математики и механики Уральского отделения Российской академии наук, для обработки дампов сетевого трафика, передаваемого в магистральных интернет-каналах.

1. Анализ механизмов параллельных вычислений MATLAB

Пакет MATLAB предоставляет пользователю следующие средства для проведения параллельных вычислений на нескольких уровнях.

1. Функции, реализующие параллельные алгоритмы для решения задач линейной алгебры, вы-

¹ Под технологией параллельных вычислений мы понимаем совокупность программных и аппаратных средств, позволяющих организовать набор взаимодействующих вычислительных процессов, работающих одновременно.

числения преобразования Фурье, операции с матрицами и др., для выполнения которых одновременно используют несколько центральных процессоров [4]. Данные функции реализуют режим параллельных вычислений автоматически без участия пользователя, используя для этого локальные ресурсы систем с общей памятью.

2. Пакет расширения MATLAB — Parallel Computing Toolbox, обеспечивающий пользователя средствами настройки и контроля процесса распараллеливания вычислений. Данный инструмент предназначен для написания параллельных алгоритмов (циклов, распределенных массивов) и организации распределенных вычислений. При совместном использовании с пакетом Distributed Computing Server Toolbox позволяет задействовать не только локальные, но и распределенные вычислительные ресурсы, работает на серверах с общей памятью и графических ускорителях.

3. Пакет Distributed Computing Server является серверной частью системы организации распределенных вычислений, он может работать как на выделенном кластере MATLAB, так и на разделяемом кластере под управлением системы запуска задач. Он используется для запуска задач MATLAB на кластере с распределенной памятью. На узлах кластера запускается один или несколько процессов MATLAB, которые обмениваются данными между

собой с помощью библиотеки функций MPI (Message passing interface).

Отметим, что возможны различные способы организации одновременного взаимодействия между разными уровнями параллельных вычислений в пакете MATLAB. Например, в кластере можно одновременно выполнять несколько вычислительных процессов MATLAB, использующих многопоточные функции. При использовании встроенной многопоточности MATLAB или средств Parallel Computing Toolbox можно задействовать только локальные ресурсы компьютера. При этом существует программное ограничение на максимальное число процессоров/ядер, применяемых в параллельных вычислениях — не более 12.

Отметим, что увеличение числа процессоров/ядер, как правило, увеличивает скорость выполнения задачи в арифметической прогрессии. При использовании одной ЭВМ с 12 ядрами, тактовая частота которых равна 3,0 ГГц, время обработки дампа составит около 80 ч. Однако при необходимости обработки большого числа дампов сетевого трафика данный результат также следует признать неудовлетворительным. В то же время обойти ограничение на число центральных процессоров можно, используя вычислительный кластер, на котором установлен пакет Distributed Computing Server.

2. Обзор кластеров MATLAB

В связи с высокой стоимостью программного и аппаратного обеспечения, необходимого для организации Distributed Computing Server, и текущей политикой его продаж (только для юридических лиц) необходимо рассмотреть другие проекты, позволяющие получить доступ к аналогичным программным и аппаратным ресурсам с меньшими финансовыми затратами.

1. Облако Mathworks Cloud, разработанное компанией *Mathworks* [5]. Данный сервис предназначен, в первую очередь, для мобильных устройств и использовать его можно только из специализированной версии MATLAB Mobile. Данный сервис имеет существенные ограничения: отсутствует возможность применения редактора MATLAB, объем передаваемых данных не может превышать 500 Мбайт и др. В связи с этим данное средство не подходит для обработки дампов.

2. Вычислительное облако Amazon EC2 является совместным проектом *Mathworks* и *Amazon* [6]. Для его применения на сервере EC2 должен быть запущен кластер Distributed Computing Server, а на клиентском компьютере может быть использован Parallel Computing Toolbox. Однако в данный момент проект находится на начальной стадии развития и в России недоступен.

3. Облако Red Cloud на основе использования Distributed Computing Server, доступ к которому предоставляет Корнельский университет [7]. К не-

достаткам данного сервиса следует отнести обязательное наличие клиента MATLAB с пакетом расширения Parallel Computing Toolbox на компьютере пользователя для выполнения задач на кластере.

4. Российские суперкомпьютерные вычислительные системы с терафлопсной производительностью, например: "Ломоносов" МГУ им. Ломоносова, "МВС-10П" межведомственного Суперкомпьютерного Центра Российской академии наук, "Лобачевский" НГУ им. Лобачевского и др. [8]. Однако только единицы из них имеют программное обеспечение MATLAB, предоставляемое в аренду вместе с вычислительными ресурсами, в том числе суперкомпьютеры "Торнадо ЮУрГУ" Южно-Уральского государственного университета и "Уран" Института математики и механики Уральского отделения РАН.

Оба этих супервычислителя входят в первую десятку [8] рейтинга суперкомпьютеров России, обладают достаточными ресурсами для обеспечения удовлетворительного времени (не более одного рабочего дня) обработки дампов, предоставляют комфортные условия взаимодействия, не требующие наличия программного обеспечения MATLAB на клиентской машине. Поскольку серверная лицензия "Торнадо ЮУрГУ" позволяет использовать не более 32 ядер [9], авторами было принято решение использовать вычислитель "Уран", лицензия которого позволяет использовать сотни ядер [10].

3. Облачный кластер MATLAB суперкомпьютера "Уран"

Суперкомпьютер "Уран" ИММ УрО РАН собран на базе серверов Blade фирмы *Hewlett-Packard*. Он состоит из 204 вычислительных узлов, установленных в модулях с высокой плотностью упаковки. Вычислительные узлы оснащены процессорами Intel Xeon, работающими на частотах 2,2...3 ГГц, 16...200 Гбайт оперативной памяти, графическими ускорителями NVIDIA Tesla и предоставляет пользователям для работы следующие вычислительные средства:

- число ядер — 1864 CPU Xeon (3,0 ГГц) и 352 GPU Tesla;
- оперативная память — 6976 Гбайт;
- объем памяти системы хранения — 10 Тбайт;
- коммуникативная среда — Infiniband, GiEthernet.

Они обеспечивают пиковую производительность 216,56 Tflor/s и 105,36 Tflor/s на тесте Linpack.

Базовое программное обеспечение вычислительного кластера включает в себя:

- операционную систему Linux;
- систему запуска задач Slurm;
- языки программирования C, C++, Fortran;
- компиляторы Intel, GNU, PGI;
- библиотеку Math Kernel Library (MKL) Intel;
- реализации MPI — OpenMPI и MVAPICH2;
- пакеты MATLAB, ANSYS CFX Academic Research.

Способы взаимодействия пользователя с супер-ЭВМ представлены ниже.

Тип доступа	Командная строка	Графический интерфейс	Обмен файлами
Программное обеспечение	PuTTY	WinSCP	Nomachine NX MobaXterm

3.1. Особенности подготовки кластера для запуска программ

Для подключения к вычислительному кластеру каждому пользователю создается учетная запись (логин и пароль). В зависимости от заявки пользователя задается время для выполнения задач, режим выделения ресурсов (динамический или монопольный), а также выделяются вычислительные ресурсы: число доступных процессорных ядер, размер оперативной памяти и дискового пространства.

В процессе запуска программы на кластере были обнаружены проблемы, связанные с использованием дополнительных компонентов: *mex*-функции [2] и библиотеки *libwireshark.so*. Это потребовало выполнения дополнительных настроек вычислительного кластера, позволивших MATLAB "увидеть" необходимые компоненты: в домашнем каталоге была создана директория (*home/username/lib*) для размещения дополнительных библиотек и функций. Для указания каталога, в котором находятся

библиотеки, потребовалось скопировать стартовый *shell*-скрипт *MATLAB.matlab7rc.sh* в домашнюю директорию и добавить в него переменную *LDPATH_PREFIX="home/username/lib"*. Это обеспечило автоматическое добавление каталога с необходимыми библиотеками при запуске MATLAB для каждого нового пользователя.

3.2. Запуск программы на кластере

Графический интерфейс, предоставляемый программой MobaXterm, является наиболее удобным способом запуска программ MATLAB на кластере. MobaXterm эмулирует программу MATLAB так, как будто она запущена на клиентском компьютере.

Для запуска программы в многопоточном режиме оказывается необходимым модифицировать ее код, применяя специальные конструкции языка — *parfor* и *spmd* [11]. Использование данных конструкций предполагает открытие пула MATLAB, равного необходимому числу процессоров (*MATLAB workers* или *labs*) на кластере. При этом пул можно задавать в явном и неявном виде.

Для разгрузки клиентской ЭВМ и обеспечения возможности интерактивной работы на ней использовалась команда *batch*, позволяющая запускать программы асинхронно, в пакетном режиме, разгружая клиента MATLAB. Команда запуска

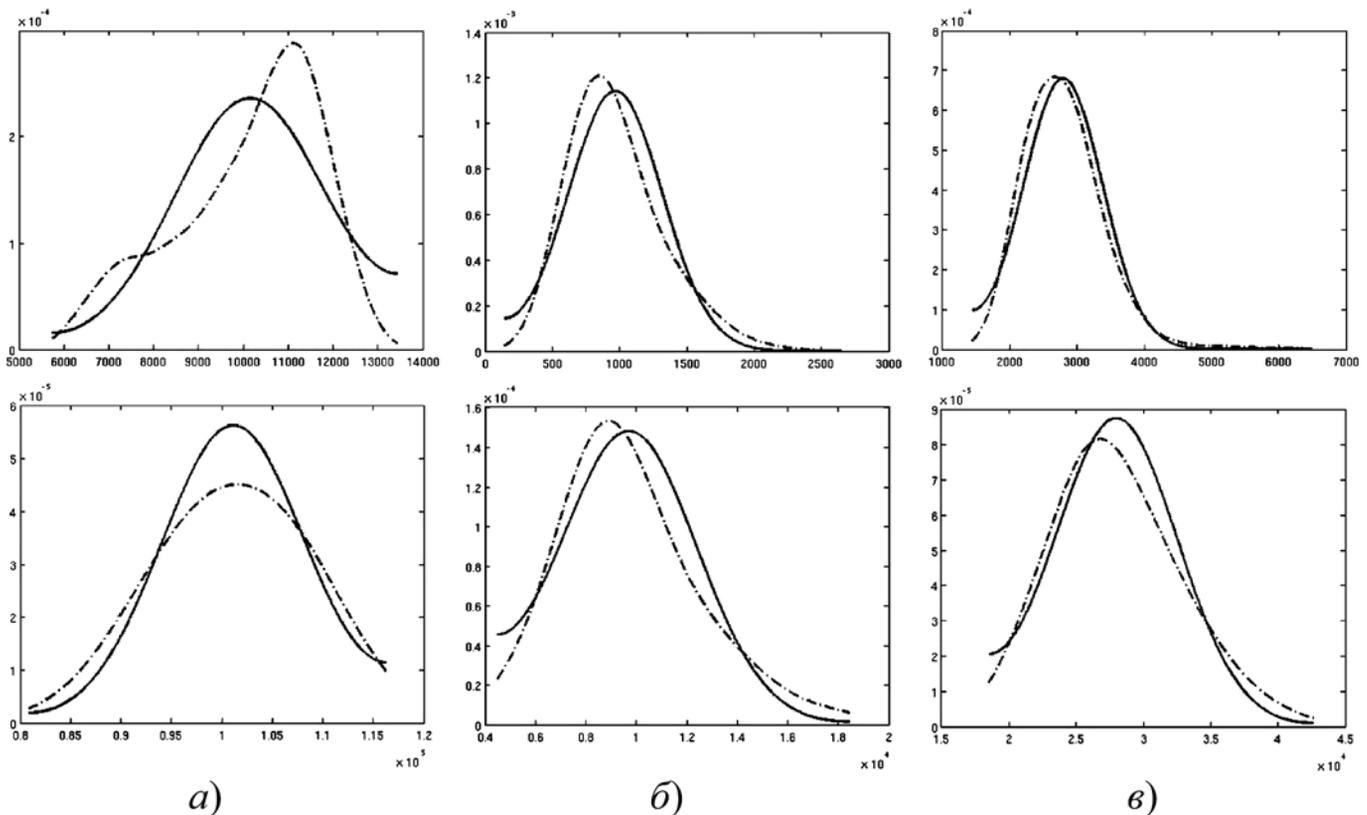


Рис. 3. Функции плотности распределений потоков сетевого трафика [3] различных классов на интервале 1 с (вверху) и 0,1 с (внизу): а — "Мыши"; б — "Мулы"; в — "Слоны"

программы вида `program_name.m` выглядят следующим образом:

— `job = batch('program_name')` (запуск программы в однопроцессорном варианте).

— `job = batch('program_name', 'matlabpool', 100)` (запуск программы в режиме использования 101 процессора: 1 — для программы `program_name.m` и 100 — для пула).

Для выполнения программ создается объект `job` (работа). Максимальное время и ресурсы для выполнения программы, указанной в объекте `job`, определяет планировщик на кластере "Уран" (тип планировщика `generic`). По завершении работы результаты сохраняются в домашней папке пользователя с указанием номера работы.

В проведенном исследовании была использована трехкомпонентная модель [12] классификации источников трафика.

1. Класс "Слоны": размер данных, передаваемых потоком, больше либо равен 10 Мбайт (P2P, скачивание больших файлов). Время жизни таких потоков велико.

2. Класс "Мулы": размер данных, передаваемых потоком, составляет от 0,3 до 10 Мбайт (просмотр видеороликов, прослушивание музыки, скачивание файлов небольших размеров).

3. Класс "Мыши": размер данных, передаваемых потоком, составляет менее 0,3 Мбайт (ICQ-сообщения, просмотр web-страниц и т.д.).

Типичные результаты обработки дампа на суперкомпьютере "Уран" (рис. 3) — функции плотностей распределений потоков (метод генетических алгоритмов — непрерывная линия и Розенблатта—Парзена — штрихпунктирная линия), создаваемых различными классами пользователей.

Заключение

Результаты проведенной проверки технологии анализа дампа [2] показали, что среднее время обработки одного дампа размером около 10 Гбайт составляет 7–8 ч, что примерно в 120–140 раз быстрее обработки дампа на одном процессоре. В течение двух суток, выделенных в ИММ УрО РАН для использования суперкомпьютера (120 процессорных ядер, 2 Гбайта оперативной памяти на каждое ядро), было обработано шесть 15-минутных дампов интернет-трафика, передаваемого в магистральном интернет-канале. Это подтверждает работоспособность и эффективность использованной технологии обработки данных дампов сетевого трафика, которая позволяет значительно ускорить процесс проведения экспериментов по изучению свойств трафика и дает возможность проведения детальных исследований особенностей информационных потоков в магистральных интернет-каналах.

Получены результаты анализа дампов трафика (подробное обсуждение которых является темой отдельной публикации), свидетельствующие об отсутствии подобия статистических свойств трафика при изменении размера окна агрегации. Данные результаты позволяют, в частности, поставить под сомнение адекватность модели самоподобного трафика, широко используемой в настоящее время (см, например, работы [13–15] и др.).

Список литературы

1. **Wireshark** — go deep: [Электронный ресурс]. URL: <https://www.wireshark.org/> (дата обращения: 01.04.2015).
2. **Поршнев С. В., Божалкин Д. А.** Технология семантического анализа дампа трафика информационных потоков в компьютерных сетях // Информационные технологии. 2014. № 11. С. 12–19.
3. **MAWI Working Group Traffic Archive**: [Электронный ресурс]. URL: <http://mawi.nyu.edu/> (дата обращения: 01.02.2015).
4. **MATLAB** и многоядерность. Работа в MATLAB на многоядерных и многопроцессорных компьютерах [Электронный ресурс] // MATLAB и Simulink центр компетенций компании Mathworks. URL: <http://matlab.ru/solutions/tech-calc/parallel-computing/matlab-multicore> (дата обращения: 10.02.2015).
5. **Connecting to the Cloud** [Электронный ресурс] // Mathworks — MATLAB and Simulink for technical computing. URL: <http://www.mathworks.com/mobile/connect-to-cloud.html> (дата обращения: 10.02.2015).
6. **Register for Access to MATLAB Distributed Computing Server on the Cloud** [Электронный ресурс] // Mathworks — MATLAB and Simulink for technical computing. URL: <http://www.mathworks.com/programs/mdcs-cloud.html> (дата обращения: 03.02.2015).
7. **Red Cloud** [Электронный ресурс] // Cornell University Center for Advanced Computing. URL: <https://www.cac.cornell.edu/RedCloud/default.aspx> (дата обращения: 10.02.2015).
8. **TOP 50**. Текущий рейтинг. 21-я редакция от 23.09.2014 [Электронный ресурс] Суперкомпьютеры. URL: <http://top50.supercomputers.ru/?page=rating> (дата обращения: 10.02.2015).
9. **Пакет Math Works MATLAB** [Электронный ресурс] // Лаборатория суперкомпьютерного моделирования ЮУрГУ. URL: <http://supercomputer.susu.ac.ru/users/simulation/matlab/> (дата обращения: 10.02.2015).
10. **Параллельный MATLAB**. Общая информация [Электронный ресурс] // Параллельные вычисления в УрО РАН. URL: <http://parallel.uran.ru/node/119#com-lic> (дата обращения: 10.02.2015).
11. **Functions in Parallel Computing Toolbox** [Электронный ресурс] // Mathworks — MATLAB and Simulink for technical computing. URL: <http://www.mathworks.com/help/distcomp/function-list.html> (дата обращения: 10.02.2015).
12. **Pellicer-Lostao C., Morato D., Popez-Ruiz R.** Modelling user's activity in a real-world complex network // International Journal of Computer Mathematics. Bristol: Taylor & Francis. 2008. Vol. 85. P. 1287–1298.
13. **Ромасевич П. В.** Оценка необходимой канальной емкости телекоммуникационной системы с ограниченной буферной памятью в условиях самоподобного трафика // Современные информационные технологии и ИТ-образование. 2014. № 10. С. 456–460.
14. **Линец Г. И., Говорова С. В.** Использование функциональных преобразований самоподобного потока пакетов для повышения производительности транспортных сетей // Информационные технологии. 2014. Т. 12, № 1. С. 29–41.
15. **Рыжаков В. В.** Применение метода операторной оценки функций плотности вероятности для решения задач управления в телекоммуникационных сетях // Т-КОММ: Телекоммуникации и транспорт. 2014. Т. 8, № 11. С. 80–87.

S. V. Porshnev, PhD, Professor, e-mail: sergey_porshnev@mail.ru,
D. A. Bozhalkin, Postgraduate Student, e-mail: bozhalkin@mail.ru,
A. S. Koposov, Postgraduate Student, e-mail: alexkopas@gmail.com

Federal State Autonomous Educational Institution of Higher Professional Education
"Ural Federal University named after the first President of Russia B. N. Yeltsin", Ekaterinburg

Experience of Using the Supercomputer for Processing of Traffic Dumps of Backbone Internet Channel

A comparative review of public MATLAB clusters and supercomputers was made. Computing cluster was selected to perform processing of network traffic dumps. Customization of supercomputer was made to use proprietary software. The analysis technology of the network traffic dumps of the backbone Internet channel was described. The results confirming effectiveness of this technology was shown.

Keywords: supercomputer, cluster, computer networks, traffic flows, traffic dump, data extraction, data processing, parallel computing in MATLAB, distribution parameters

References

1. **Wireshark go deep**, available at: <https://www.wireshark.org/> (date of access: 01.04.2015).
2. **Porshnev S. V., Bozhalkin D. A.** Tehnologija semanticheskogo analiza dampa trafika informacionnyh potokov v komp'yuternyh setjah, *Informacionnye tehnologii*, 2014, no. 11, pp. 12–19.
3. **MAWI** Working Group Traffic Archive, available at: <http://mawi.nyu.edu/> (date of access: 01.02.2015).
4. **MATLAB i mnogojadernost'**. Rabota v MATLAB na mnogojadernyh i mnogoprocessornyh komp'yuterah, *MATLAB i Simulink centr kompetencij kompanii Mathworks*, available at: <http://matlab.ru/solutions/tech-calc/parallel-computing/matlab-multicore> (date of access: 10.02.2015).
5. **Connecting to the Cloud**, *Mathworks — MATLAB and Simulink for technical computing*, available at <http://www.mathworks.com/mobile/connect-to-cloud.html> (date of access: 10.02.2015).
6. **Register for Access to MATLAB Distributed Computing Server on the Cloud**, *Mathworks — MATLAB and Simulink for technical computing*, available at: <http://www.mathworks.com/programs/mdcs-cloud.html> (date of access: 03.02.2015).
7. **Red Cloud**, Cornell University Center for Advanced Computing, available at <https://www.cac.cornell.edu/RedCloud/default.aspx> (date of access: 10.02.2015).
8. **TOP 50 Superkomp'yutery**. Tekushij rejting. 21-aja redakcija ot 23.09.2014, available at <http://top50.supercomputers.ru/?page=rating> (date of access: 10.02.2015).
9. **Paket Math Works MATLAB**, *Laboratorija superkomp'yuternogo modelirovaniya JuUrGU*, available at <http://supercomputer.susu.ac.ru/users/simulation/matlab/> (date of access: 10.02.2015).
10. **Parallel'nyj MATLAB**. Obshaja informacija [Elektronnyj resurs], *Parallel'nye vychisleniya v UrO RAN*, available at <http://parallel.uran.ru/node/119#com-lic> (date of access: 10.02.2015).
11. **Functions** in Parallel Computing Toolbox, *Mathworks — MATLAB and Simulink for technical computing*, available at <http://www.mathworks.com/help/distcomp/functionlist.html> (date of access: 10.02.2015).
12. **Pellicer-Lostao C., Morato D., Popez-Ruiz R.** Modelling user's activity in a real-world complex network, *International Journal of Computer Mathematics*, Bristol, Taylor & Francis, 2008, vol. 85, pp. 1287–1298.
13. **Romasevich P. V.** Ocenka neobhodimoy kanal'noj emkosti telekommunikacionnoj sistemy s ogranichennoj bufernoj pamjat'ju v uslovijah samopodobnogo trafika, *Sovremennye informacionnye tehnologii i IT-obrazovanie*, 2014, no. 10, pp. 456–460.
14. **Linec G. I., Govorova S. V.** Ispol'zovanie funkcional'nyh preobrazovaniy samopodobnogo potoka paketov dlja povyshenija proizvoditel'nosti transportnyh setej, *Infokommunikacionnye tehnologii*, 2014, vol. 12, no. 1, pp. 29–41.
15. **Ryzhakoy V. V.** Primenenie metoda operatornoj ocenki funkcion plotnosti verojatnosti dlja reshenija zadach upravleniya v telekommunikacionnyh setjah, *T-COMM: Telekommunikacii i transport*, 2014, vol. 8, no. 11, pp. 80–87.

ИНФОРМАЦИЯ

11–14 марта 2016 г.

на базе Алтайского государственного университета состоится

VI Международная научно-практическая конференция

"МНОГОЯДЕРНЫЕ ПРОЦЕССОРЫ, ПАРАЛЛЕЛЬНОЕ ПРОГРАММИРОВАНИЕ, ПЛИС, СИСТЕМЫ ОБРАБОТКИ СИГНАЛОВ" (МППОС-2016)

Тематические направления конференции

Многоядерные процессоры, ПЛИС и обработка сигналов

Параллельное программирование и моделирование явлений и процессов с использованием параллельных вычислений
Робототехника и автоматизация технологических процессов и научного эксперимента

КОНТАКТНЫЕ АДРЕСА И ТЕЛЕФОНЫ:

Тел.: каф. ВТиЭ АлтГУ: 8-(3852) 38-07-51

Иордан Владимир Иванович, e-mail: jordan@phys.asu.ru, тел.: + 7-960-937-89-00

Калачев Александр Викторович, e-mail: kalachev@phys.asu.ru, тел.: + 7-913-027-84-06

Белозерских Василий Вениаминович, e-mail: bww@phys.asu.ru, тел.: + 7-903-947-71-15

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ИЗОБРАЖЕНИЙ DIGITAL PROCESSING OF SIGNALS AND IMAGES

УДК 004.9

А. С. Наумов, инженер-программист,
ООО "Электра-Н", Великий Новгород, e-mail: alex.naumov53@mail.ru

Метод автоматизированной реконструкции фрагментированных изображений на основе совокупности признаков фрагментов

Рассмотрена проблема использования совокупности признаков фрагментов при автоматизированной реконструкции фрагментированных изображений в приложении к восстановлению разрушенных материалов в криминалистике, археологии, реставрации и других областях. Предложен универсальный трехэтапный метод автоматизированного поиска стыков фрагментов на основе совокупности признаков, использующий принцип конвейерной обработки. Рассмотрен алгоритм синтеза изображения на основе найденных стыков. Показаны их особенности и преимущества.

Ключевые слова: стыковка фрагментов, реконструкция изображения, классификация, поиск, верификация, граф, дескриптор фрагмента, поиск в ширину, конвейерная обработка

Введение

В различных прикладных областях, таких как реставрация, криминалистика, музейное и архивное дело, возникает задача восстановления механически фрагментированных объектов: фресок, керамических изделий, бумажных документов, фотографий и других материалов. Традиционный подход к ее решению — это ручной подбор стыкующихся между собой фрагментов для последующего воссоздания композиции, соответствующей исходному объекту. Такой процесс предполагает выполнение большого объема малоэффективной ручной работы, которая приводит к необратимому разрушению материалов в процессе физического кон-



Рис. 1. Изображение фресковой композиции

такта с ними. Актуальным направлением исследований является внедрение в этот процесс перспективных информационных технологий, в частности, создание методов его автоматизации с использованием компьютерной техники и соответствующего программного обеспечения, что позволит перенести процесс реконструкции в виртуальную среду и многократно повысить его эффективность.

Рассмотрим класс задач, предполагающих реконструкцию материалов с плоской поверхностью, которая может быть сведена к задаче синтеза изображения из фрагментов произвольной формы. Для этого используют двумерные цифровые изображения поверхности фрагментов, полученные путем сканирования или фотографирования, а также дополнительную информацию об их свойствах, задаваемую вручную или автоматически. На рис. 1 показан пример фрагментированного объекта.

Процесс автоматизированной реконструкции изображения объекта в общем виде включает следующие этапы (рис. 2):

- оцифровка материалов;
- формализация и описание характеристик (признаков) фрагментов с использованием соответствующих математических моделей;
- перебор пар фрагментов в целях сравнения признаков и выявления стыков;
- синтез изображения композиции на основе стыкующихся пар фрагментов.

В зарубежных источниках рассматриваемой проблеме посвящено значительное число публикаций. В качестве примера можно привести работы [1] и [2] (описана реконструкция керамических изделий),

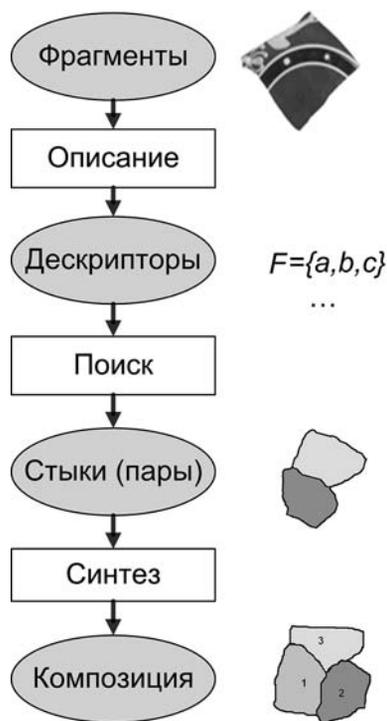


Рис. 2. Схема процесса реконструкции

[3] и [4] (реконструкция бумажных документов), [5] и [6] (реконструкция фотографий). Анализ литературы показывает, что практически все предложенные подходы базируются на конкретных примерах и задачах, следовательно, имеют ограниченную сферу применения и невысокую практическую ценность. Рассматриваются материалы какого-либо одного конкретного типа, учитываются их одиночные признаки. Таким образом, проблема в определенной степени разработана на уровне отдельных характеристик фрагментов, но не рассматривается в комплексе, на более общем концептуальном уровне с точки зрения обеспечения универсальности методов и их применения. В открытых источниках не удалось найти публикаций, затрагивающих подобные аспекты, также отсутствует единая терминология в рассматриваемой области.

Возникает задача создания обобщенного метода реконструкции, позволяющего использовать различные (существующие и перспективные) способы описания и сравнения отдельных признаков фрагментов, однако обеспечивающего их совместное применение и универсальный подход, не привязанный к специфике конкретной задачи. Такой метод предложен автором и описан в данной статье. Совместное применение совокупности признаков призвано повысить эффективность и результативность процесса реконструкции за счет максимального полного использования доступной информации о свойствах фрагментов. Общая структура метода является универсальной, а адаптация к конкретным задачам происходит за счет изменения набора и процедур сравнения отдельных признаков.

1. Признаки фрагментов и их дескрипторы

Поиск стыков фрагментов осуществляется путем сравнения их характеристик (признаков). Типичными признаками являются геометрический и цветовой контур, текстура, цветовая палитра, особенности материала, семантические свойства изображения и т.д. Их обзор и классификация представлены в работах [7], [8]. Способы описания и сравнения отдельных признаков в данной статье не приведены, поскольку являются темой отдельных исследований. Достаточно рассмотреть их абстрактно с учетом формальных особенностей.

С точки зрения представления выделяют следующие группы признаков.

1. *Элементарные*. Принимают конечное число значений из заранее определенного множества. К ним относится, например, тип бумаги, тип краски, информация о принадлежности какой-либо части композиции и т.п. Данные признаки сравниваются по критерию точного совпадения.

2. *Скалярные*. Представляются вещественным числом и допускают тривиальное количественное сравнение с помощью порогового или статистического критерия. Таким признаком может являться, например, толщина фрагмента.

3. *Векторные*. Представляются набором значений, организованных в структуру данных. К ним относятся, например, контур фрагмента, цветовая палитра, дескриптор текстуры и т.п. Векторные признаки обладают наибольшей информативностью и требуют наибольших вычислительных затрат для их сравнения, которое может выполняться с использованием пороговых, статистических, логических, семантических или иных критериев. В большинстве задач основным векторным признаком является контур, так как он позволяет идентифицировать и локализовать область стыка [7].

С точки зрения сравнения признаки можно также разделить на три группы.

1. *Предварительные*. Позволяют разделить исходное множество фрагментов на отдельные классы, но не позволяют судить о возможности стыка между ними. В частности, не имеет смысла искать стык между фрагментами рукописного документа и фотографии или сравнивать фрагменты существенно разной толщины.

2. *Основные*. Служат для установления наличия стыка и локализации его области.

3. *Верификационные*. Их можно сравнить только после расчета взаимного положения фрагментов в соответствии с предполагаемым стыком. Например, направление строк текста в документе на соседних фрагментах должно совпадать с точностью до значения допустимой ошибки, однако проверить этот факт можно только после выполнения совмещения.

Для каждого признака фрагмента формируется дескриптор — структура данных, описывающая его либо явно (для элементарных и скалярных призна-

ков), либо с использованием заданной математической модели (для векторных). Совокупность дескрипторов отдельных признаков составляет дескриптор фрагмента.

Пусть дескриптор фрагмента F формально представлен совокупностью признаков $D(F) = \{T_1(F), T_2(F), \dots, T_n(F)\}$, где $T_i(F)$ — i -й признак фрагмента; n — общее их число. Пусть имеются два фрагмента F_1 и F_2 . Тогда совпадение признака T_i у этих фрагментов с точки зрения выбранного критерия обозначим равенством $T_i(F_1) = T_i(F_2)$, несовпадение — неравенством $T_i(F_1) \neq T_i(F_2)$. Возможна ситуация, когда значение признака не задано и считается нулевым, $T_i(F) = null$.

Формирование дескрипторов осуществляется при предварительной обработке исходных данных (изображений фрагментов).

2. Процедура поиска стыков

Процедура поиска стыков в качестве входных данных использует множество дескрипторов фрагментов, а в качестве выходных данных формирует список потенциально возможных стыков с описанием параметров каждого из них. Задача перебора пар фрагментов и сравнения признаков имеет квадратичную вычислительную сложность, что является критичным с точки зрения производительности, поскольку в реальных задачах число фрагментов исчисляется тысячами. При росте общего числа потенциально возможных пар фрагментов растет вероятность обнаружения ложных стыков, что также осложняет процесс реконструкции.

В качестве оптимального решения по организации процедуры перебора и поиска стыков предложена схема с разделением ее на три этапа (рис. 3), на каждом из которых происходит проверка отдельной группы признаков. Этапы включают в себя классификацию фрагментов, поиск стыков и их верификацию.

На этапе *классификации* на основе группы предварительных признаков выделяются непересекающиеся классы фрагментов, внутри которых выполняется дальнейший поиск. Это позволяет сократить множество, подлежащее перебору и распараллелить процесс обработки независимых классов.

На этапе *поиска* происходит перебор всех возможных пар фрагментов внутри каждого класса, проверяются основные признаки, и при нахождении стыка выполняется расчет взаимного положения фрагментов. Результатом данного этапа является ранжированный по весу список возможных стыков. В качестве веса выступает величина, учитывающая длину предполагаемой области стыка и число совпадающих признаков.

На этапе *верификации* проверяются признаки, которые невозможно проверить до пробного совмещения, и исключаются пары, не совпадающие с точки зрения этих признаков. В случае реализации интерактивного режима реконструкции на данном этапе реализуется процедура ручного утверждения стыков.

Трехэтапная процедура позволяет учесть неравнозначность признаков и обеспечить широкие возможности модификации.

В соответствии с предложенной схемой и особенностями признаков их общая совокупность делится на три группы: предварительные $TP = \{T_1, T_2, \dots, T_k\}$; основные $TS = \{T_{k+1}, \dots, T_m\}$; верификационные $TV = \{T_{m+1}, \dots, T_n\}$, где n — общее число признаков. Пусть имеются два фрагмента — F_1 и F_2 .

Алгоритм классификации строят на основе следующих правил.

1. Если $T_i(F_1) = T_i(F_2) \neq null$ для всех признаков $T_i \in TP$, то фрагменты относятся к одному классу $F_1 \in C_q, F_2 \in C_q, q$ — номер класса.

2. Если существует хотя бы один признак $T_i \in TP$, такой что $T_i(F_1) \neq T_i(F_2), T_i(F_1) \neq null, T_i(F_2) \neq null$, то фрагменты относятся к разным классам: $F_1 \in C_p, F_2 \in C_q, p \neq q$.

3. Если существуют такие признаки $T_i \in TP$, значения которых хотя бы для одного из фрагментов не заданы ($T_i(F_1) = null$ или $T_i(F_2) = null$), то этот признак на классификацию не влияет, и она проводится по оставшимся признакам.

Требуется провести перебор всех фрагментов и для каждого текущего фрагмента выполнить сравнение значений дескрипторов его признаков со значениями, определенными для каждого класса. На основании этого выполняется либо отнесение фрагмента к одному из существующих классов, либо

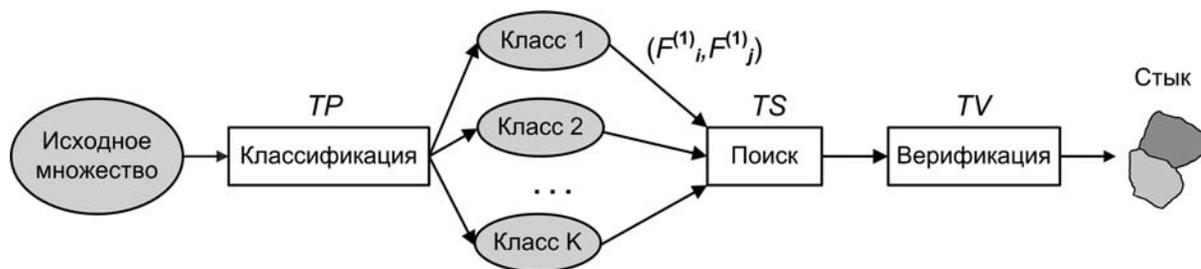


Рис. 3. Трехэтапная процедура поиска стыков: TP, TS, TV — соответствующие группы признаков (предварительные, основные, верификационные); $(F_i^{(1)}, F_j^{(1)})$ — текущая пара фрагментов из класса 1

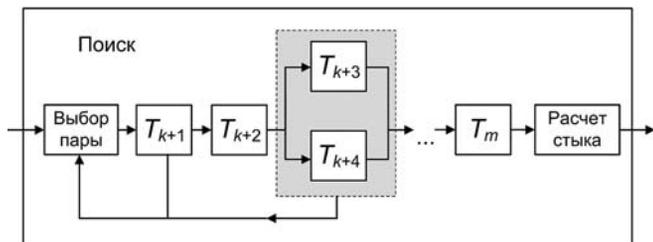


Рис. 4. Пример конвейера поиска. Признак T_{k+1} является обязательным, T_{k+3} и T_{k+4} образуют группу альтернативного выбора, прочие признаки являются вспомогательными (необязательными)

формирование нового, если класса с таким набором значений еще не существует.

Для поиска стыков на множестве из n фрагментов потребуется перебор $n(n-1)/2$ пар. При формировании двух независимых классов по $n/2$ фрагментов общее число пар будет уменьшено более чем в 2 раза. Таким образом, эффективная классификация существенно уменьшает число ненужных вычислительных операций.

Этап поиска предлагается реализовать в виде конвейера (рис. 4), который организуется в соответствии с протоколом сравнения — формальным описанием порядка проверки признаков группы TS и значимости каждого из них. Протокол сравнения индивидуален для каждой задачи и формируется экспертом. В случае совпадения некоторого признака процедура переходит к сравнению следующего и т. д. Ключевой особенностью является учет неравнозначности признаков: некоторые из них могут быть обязательными, а некоторые — вспомогательными.

Если признак является обязательным, то в случае несовпадения текущая пара отвергается, и алгоритм переходит к выбору новой пары (рис. 4). Если же он является вспомогательным, то факт совпадения или несовпадения фиксируется, а пара продолжает дальнейшее движение по конвейеру. Возможен вариант альтернативного выбора, когда обязательно требуется совпадение хотя бы одного из двух или более признаков.

Каждый найденный стык фрагментов характеризуется парой значений (l, w) , где l — длина стыка, w — вес, рассчитываемый как сумма весов вспомогательных признаков $w = r_{k+1}w_{k+1} + \dots + r_m w_m$. При этом $(m-k)$ — общее число признаков группы TS ; w_i — вес соответствующего признака $T_i \in TS$, а $r_i = 1$, если $T_i(F_1) = T_i(F_2)$ и $r_i = 0$, если $T_i(F_1) \neq T_i(F_2)$. Двухмерная метрика позволяет более гибко выбирать варианты при синтезе результирующей композиции.

Для оптимизации производительности протокол сравнения по возможности организуется так, чтобы отсеять больше ложных вариантов на ранних стадиях конвейера.

Процесс верификации организуется по принципу конвейера, аналогично процессу поиска. В слу-

чае реализации интерактивного режима реконструкции на данном этапе (как последней стадии конвейера) осуществляется процедура ручного выбора и утверждения стыков оператором-экспертом, который при необходимости может выполнить коррекцию взаимной привязки фрагментов для более качественной визуализации.

Результатом работы трехэтапной процедуры является список стыков с параметрами каждого из них (координаты границ области стыка, вес и другая необходимая информация). Процедура обеспечивает гибкость временных рамок, поскольку, например, классификацию удобно реализовать еще на этапе описания фрагментов, а верификация может быть выполнена независимо от поиска, что удобно при реализации интерактивного режима с оператором-экспертом. Обработка каждой потенциальной пары фрагментов является независимой, таким образом, задача обладает массовым параллелизмом, что должно быть учтено при разработке аппаратной и программной архитектуры комплекса для реконструкции.

3. Синтез композиции

На основе списка найденных стыков выполняется процедура синтеза композиции, позволяющая непосредственно реконструировать изображение исходного объекта. Модель композиции представляется неориентированным взвешенным графом, в котором каждая вершина соответствует фрагменту, ребро — стыку между фрагментами, а вес ребра является характеристикой стыка. Таким образом, на основе списка идентифицированных стыков необходимо построить набор графов, каждый из которых соответствует части исходного изображения, или же изображению целиком. При этом в исходном множестве могут содержаться фрагменты, как единственной композиции, так и множества отдельных композиций.

Пусть $G_{\text{общ}} = (V, E, w)$ — взвешенный граф связности, сформированный по результатам работы алгоритма поиска стыков, соответствующий множеству фрагментов, принадлежащих одному из непересекающихся классов. При этом V — множество вершин графа; E — множество ребер; w — выбранная весовая функция. Моделью реконструированной композиции (или ее части) является граф $G_{\text{комп}}$, являющийся покрывающим деревом для некоторого подграфа $G' \subset G_{\text{общ}}$ [9]. Выбор древовидной структуры обусловлен удобством хранения информации и обработки при последующей визуализации изображений. Задача синтеза композиции сводится к преобразованию общего графа в набор деревьев, соответствующих частям композиции, или в единственное дерево, соответствующее всей композиции, если таковое возможно.

Пусть L — общий список стыкующихся пар. Алгоритм синтеза композиции основан на поиске в

ширину и включает в себя следующие действия (рис. 5).

1. Выбор начального фрагмента с наибольшей длиной контура, который соответствует корневой вершине дерева с уровнем $d = 0$.
2. Формирование списка возможных стыков $L_d \subset L$ для фрагментов уровня d .

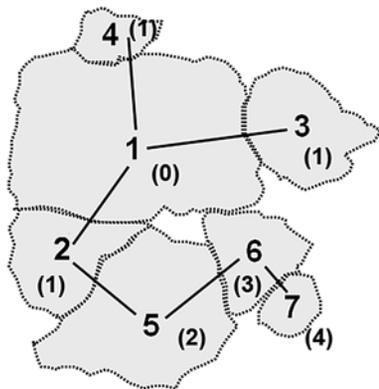


Рис. 5. Построение дерева композиции для некоторого изображения. Указан порядок добавления элементов и уровень вершин графа (в скобках)

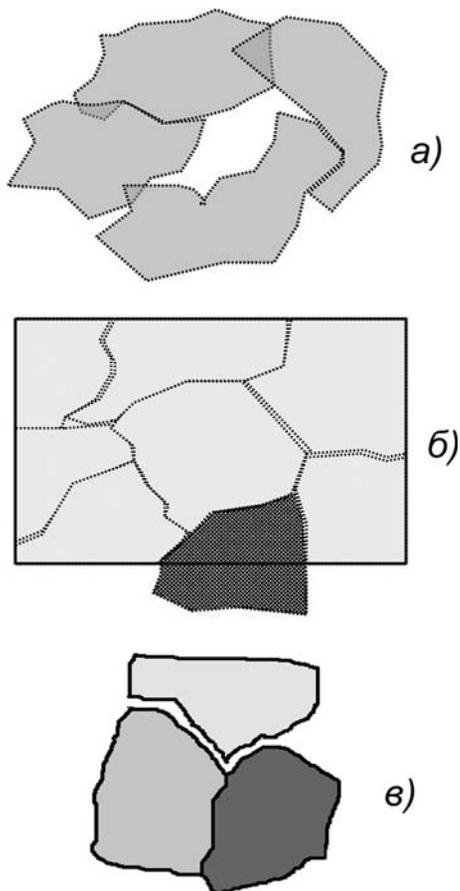


Рис. 6. Совместимость фрагментов: *a* — перекрытия недопустимы; *b* — размеры и форма композиции должны удовлетворять ограничениям; *в* — фрагменты, образующие общий угол, являются приоритетными

3. Выбор текущего стыка. Из списка L_d выбирается стык с наибольшей длиной контура l среди всех стыков с максимальным весом w . Возможны другие варианты выбора, определяемые спецификой задачи и экспериментом.

4. Проверка совместимости выбранного фрагмента с уже имеющейся частью композиции. Если обнаружен конфликт, то стык исключается, и происходит переход на шаг 3.

5. Стыковка выбранного фрагмента к текущей композиции. Визуализация. При этом добавленный фрагмент имеет уровень $d + 1$.

6. Повторение алгоритма с шага 3, до тех пор, пока есть доступные стыки для любого из фрагментов уровня d . Если таких кандидатов нет, то выполняется переход на следующий уровень $d = d + 1$, и алгоритм повторяется с шага 2 для всех фрагментов следующего уровня.

Если существуют независимые кластеры фрагментов, соответствующие несвязанным частям композиции, то алгоритм повторяется для каждого из них.

Совместимость фрагментов на локальном уровне (уровне отдельных стыков) не гарантирует глобальную совместимость на уровне всей композиции, поскольку возможно появление ложных или эквивалентных вариантов. Чтобы их избежать, на шаге 3 проверяются эвристические правила, гарантирующие, что фрагмент успешно встраивается в текущую композицию. Такими правилами являются, например, отсутствие перекрытий с уже состыкованными фрагментами, соответствие композиции заданным размерам (рис. 6) и т.п. Фрагменты, образующие общий угол, могут быть приоритетными для выбора, так как имеют стык с двумя фрагментами сразу. Эвристические правила могут быть достаточно разнообразны и их адекватный подбор повышает достоверность реконструкции.

Заключение

Данная статья является результатом теоретического анализа проблемы реконструкции. В настоящее время автором ведется экспериментальная отработка предложенных решений. Программные средства, включающие представленные методы, планируется внедрить на базе Новгородского музея-заповедника, где остро стоит проблема сохранения, реставрации и реконструкции объектов и предметов культурного наследия, однако такая работа является длительной и затратной.

Представленный подход предлагает решение проблемы комплексного использования совокупности отдельных признаков при решении задачи реконструкции изображения из фрагментов. Он может быть полезен при разработке как экспериментальных, так и полномасштабных прикладных программно-аппаратных комплексов для восстановления фрагментированных материалов в криминалистике, реставрации, археологии и других областях.

Принципиальной особенностью метода является поэтапное многоступенчатое отсеивание неподходящих вариантов, которое направлено на повышение достоверности конечного результата, и использование максимального количества доступной информации о свойствах фрагментов. Эти решения позволяют повысить производительность и результативность процесса реконструкции и труда специалистов. Преимуществом является универсальность и гибкость с точки зрения адаптации к различным задачам при сохранении общей логики процесса. Для достижения наилучших результатов требуется учесть как можно больше отдельных признаков фрагментов и применить адекватные методы описания и сравнения каждого из них.

Список литературы

1. **Leitao H. C., Stolfi J.** Multiscale method for the reassembly of two-dimensional fragmented objects // *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 2002. V. 24. N. 9. P. 1239—1251.
2. **McBride J. C., Kimia B. B.** Archaeological fragment reconstruction using curve-matching // *Proc. of Conference on computer vision and pattern recognition CVPRW '03*. 2003. N. 1. P. 3—11.
3. **Deever A., Gallagher A.** Semi-automatic assembly of real cross-cut shredded documents // *Proc. 19th IEEE International Conference on Image Processing, ICIP 2012*, 2012. P. 233—236.
4. **Justino E., Oliveira L. S., Freitas C.** Reconstructing shredded documents through feature matching // *Forensic Science International*. 2006. V. 160. N. 2. P. 140—147.
5. **Amigoni F., Gazzani S., Podico S.** A Method for reassembling fragments in image reconstruction // *International conference on image processing ICIP 2003. Proceedings*. V. 2. P. 581—584.
6. **Shengjiao C., Hairong L., Shuicheng Y.** Automated assembly of shredded pieces from multiple photos // *IEEE International Conference on Multimedia & Expo (ICME)*. Proceedings. 2010.
7. **Наумов А. С.** Анализ методов описания контуров фрагментов при решении задачи синтеза изображения; НовГУ им. Ярослава Мудрого. Великий Новгород, 2010. — Деп. в ВИНТИ РАН 24.01.11, № 15-V2010.
8. **Наумов А. С.** Критерии сравнения характеристик фрагментов в задаче синтеза композиции; НовГУ им. Ярослава Мудрого. Великий Новгород, 2011. Деп. в ВИНТИ РАН 01.07.2011, № 315-V2011.
9. **Емеличев В. А., Мельников О. И., Сарванов В. И., Тышкевич Р. И.** Лекции по теории графов. М.: Наука, 1990. 384 с.

A. S. Naumov, Software Engineer, OOO Electra-N, Velikiy Novgorod, alex.naumov53@mail.ru

Method for Automated Reconstruction of the Fragmented Images Based on Multiple Fragment Features

Article deals with a problem of the automated fragmented images reassembling in application to the object reconstruction in forensics, archaeology, restoration, archive science and other fields. The basic idea involves using a two-dimensional digital images and additional information to automatically search matching fragment pairs in order to reconstruct entire image. The key feature of the presented approach is concerning multiple fragment features to use all available fragment information. It allows building effective tools for reconstruction process that minimizes manual job in compare to traditional methods. Article presents universal match searching method and image synthesis algorithm. Match searching includes 3 stages: classification, search and verification. Classification stage divides initial fragment set into several independent classes. Search stage performs feature comparison between all possible fragment pairs in each class and calculates junctions if they exist. Verification stage compares features that are not possible to deal before junctions are calculated.

Pipeline processing on the search stage allows to build flexible solutions for particular applied problems by variation of the set of used features, criterions, order of comparison, significance etc., while general scheme remains the same. Image synthesis algorithm is based on a breadth-first search graph algorithm and uses pairs of fragments to form entire composition image. Article shows advantages and significance of the developed methods. Presented results could be useful in a program tools development for actual applications.

Keywords: fragment matching, image reconstruction, classification, search, verification, graph, fragment descriptor, breadth-first search, pipeline processing

References

1. **Leitao H. C., Stolfi J.** Multiscale method for the reassembly of two-dimensional fragmented objects, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002, vol. 24, no. 9, pp. 1239—1251.
2. **McBride J. C., Kimia B. B.** Archaeological fragment reconstruction using curve-matching, *Proc. of Conference on computer vision and pattern recognition CVPRW'03*, 2003, no. 1, pp. 3—11.
3. **Deever A., Gallagher A.** Semi-automatic assembly of real cross-cut shredded documents, *Proceedings 19th IEEE International Conference on Image Processing, ICIP 2012*, pp. 233—236.
4. **Justino E., Oliveira L. S., Freitas C.** Reconstructing shredded documents through feature matching, *Forensic Science International*, 2006, vol. 160, no. 2, pp. 140—147.
5. **Amigoni F., Gazzani S., Podico S.** A Method for reassembling fragments in image reconstruction, *International conference on image processing ICIP 2003. Proceedings*, vol. 2, pp. 581—584.
6. **Shengjiao C., Hairong L., Shuicheng Y.** Automated assembly of shredded pieces from multiple photos, *IEEE International Conference on Multimedia & Expo (ICME), 2010, Proceedings*.
7. **Naumov A. S.** Analiz metodov opisanija konturov fragmentov pri reshenii zadachi sinteza izobrazhenija, NovGU im. Jaroslava Mudrogo, Velikiy Novgorod; 2010. Dep. v VINITI 24.01.11, № 15-V2010.
8. **Naumov A. S.** Kriterii sravnenija harakteristik fragmentov v zadache sinteza kompozicii, NovGU im. Jaroslava Mudrogo, Velikiy Novgorod, 2011. Dep. v VINITI RAN 01.07.2011, № 315-V2011.
9. **Emelichev V. A., Mel'nikov O. I., Sarvanov V. I., Tyshkevich R. I.** Lekcii po teorii grafov, Moscow, Nauka, 1990.

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ CRYPTOSAFETY INFORMATION

УДК 004.056.53

К. А. Щеглов, аспирант, **А. Ю. Щеглов**, д-р техн. наук, профессор,
Исследовательский университет информационных технологий,
механики и оптики, Санкт-Петербург, e-mail: info@npp-itb.spb.ru

Интерпретация и моделирование угрозы атаки на информационную систему.

Часть 2. Моделирование угрозы атаки

Построены марковские модели угрозы атаки на информационную систему, как систему с отказами и восстановлениями характеристики безопасности, а также с фатальным отказом, основанные на рассмотрении угрозы уязвимости в качестве простейшего элемента информационной безопасности при интерпретации угрозы атаки схемой последовательного резервирования угроз уязвимостей. Разработана модель нарушителя на основе введенной характеристики сложности реализации угрозы атаки, интерпретируемой, как вероятностная мера количества информации, которой должен обладать нарушитель в отношении угроз уязвимостей, создающих угрозу атаки, для ее реализации. Построенные модели позволяют рассчитывать параметры и характеристики угрозы атаки без необходимости получения каких-либо экспертных оценок, с использованием исключительно стохастических параметров уязвимостей, в отношении которых существует и непрерывно ведется соответствующая статистика.

Ключевые слова: безопасность, информационная система, угроза атаки, модель, проектирование, потенциальный нарушитель, система защиты информации

Введение

В работе [1] в качестве простейшего элемента безопасности информационной системы при построении математических моделей угрозы атаки на информационную систему предложено рассматривать угрозу уязвимости, что позволило использовать существующую и непрерывно ведущуюся статистику в отношении выявляемых и устраняемых уязвимостей при задании входных параметров разрабатываемых моделей. Существенным преимуществом предложенного подхода является то, что для задания входных параметров математических моделей не требуется использования каких-либо экспертных оценок, применение которых ставит под сомнение адекватность получаемых результатов моделирования. Кроме того, в работе [1] исследовали вопросы моделирования и оценки актуальности угроз уязвимостей, введена их классификация, разработана модель угрозы уязвимости, как системы с отказами и восстановлениями характеристики безопасности, предложена интерпретация угрозы атаки схемой параллельного резервирования угроз уязвимостей. Определена ключевая задача современных систем защиты информационных систем, состоящая в нивелировании актуальных угроз безусловных и условных технологических уязвимостей.

Вместе с тем при проектировании системы защиты информационной системы ключевым элементом моделирования становится угроза атаки, поскольку именно от актуальных угроз атак реализуется защита, посредством нивелирования системой защиты актуальных угроз уязвимостей, создающих угрозу атаки, как следствие, необходимо разработать модели угрозы атаки в целях получения количественной оценки ее актуальности. При моделировании угрозы атаки практический интерес уже представляет не только построение модели с отказами и восстановлениями характеристики безопасности, причем с учетом того, что угроза атаки создается не одной, а в общем случае некоторой совокупностью разнородных угроз уязвимостей, позволяющей получать количественные оценки характеристик возникновения и устранения в системе реальной угрозы атаки, но и построение модели с фатальным отказом, предполагающей уже непосредственно реализацию нарушителем реальной угрозы атаки на конкретную информационную систему (реализацию несанкционированного доступа), для которой проектируется система защиты. С этой целью уже необходимо построить модель нарушителя, позволяющую оценить (опять же без применения экспертных оценок для задания входных параметров для модели нарушителя) готовность реализации атаки

определенной сложности нарушителем на конкретную информационную систему, характеризующую меру его заинтересованности в осуществлении подобной атаки, что в том числе предполагает количественную оценку сложности реализации атаки.

1. Марковские модели угрозы атаки

В работе [1] было дано обоснование корректности использования при моделировании угрозы уязвимости аппарата марковских случайных процессов при допущениях о пуассоновском характере потока заявок и о показательном распределении времени обслуживания и был сделан вывод о том, что при моделировании угрозы безопасности информационной системы, в том числе и при моделировании угроз атак, можно использовать марковские модели, позволяющие определять граничные (худшие) значения характеристик безопасности, которые и необходимы при проектировании систем защиты информационных систем.

1.1. Марковская модель угрозы атаки как системы с отказами и восстановлениями характеристики безопасности. Информационную систему, как в отношении возникновения и устранения угрозы уязвимости, так и в отношении возникновения и устранения угрозы атаки в целом, создаваемой соответствующей совокупностью выявленных уязвимостей, можно рассматривать как систему с отказами и восстановлениями, в нашем случае, характеристики безопасности.

В работах [1, 2] предложено представлять угрозу атаки последовательностью используемых при реализации атаки уязвимостей — орграфом, вершины которого взвешены значениями P_{0yr} , $r = 1, \dots, R$, — значениями вероятности отсутствия в системе r -й уязвимости (информационная система готова к безопасной эксплуатации в отношении угрозы r -й уязвимости) — одной из R угроз уязвимостей последовательно (дуги графа определяют последовательность использования выявленных уязвимостей при реализации атаки) используемых атакой на информационную систему (рис. 1).

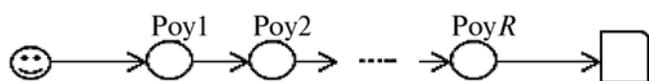


Рис. 1. Орграф угрозы атаки

В работе [1] также предложена классификация угроз уязвимостей, которые подразделены на технологические уязвимости (уязвимости корректности реализации функций защиты) — условные и безусловные, и угрозы уязвимости реализации (в частности, это ошибки программирования системных средств и приложений), возникновение которых в

системе позволяет реализовать условные (возникают при условии возникновения соответствующей уязвимости реализации) технологические уязвимости. Поскольку угрозы технологических уязвимостей характеризуются $P_{0yr} = 1$, при построении марковской модели угрозы атаки их вершины необходимо исключать из орграфа угрозы атаки (см. рис. 1), так как модель угрозы атаки строится для приведенного подобным образом орграфа, включающего в себя только размеченные вершины угроз уязвимостей реализации.

Построим марковскую модель, описывающую процесс возникновения и устранения реальной угрозы атаки в информационной системе. Под реальной угрозой атаки понимаем возникновение в системе условий возможности ее реализации нарушителем [1], при которых все уязвимости, угрозы которых создают угрозу атаки (см. рис. 1), выявлены и не устранены.

Отметим, что подобные условия будут, как возникать в системе, что можно интерпретировать, как отказ характеристики безопасности, под которой понимаем свойство системы находиться в безопасном состоянии, так и устраняться, для чего достаточно устранить по крайней мере одну выявленную в системе уязвимость, необходимую для реализации атаки, что можно интерпретировать, как восстановление характеристики безопасности. Таким образом, данной моделью описываются исключительно свойства безопасности системы, собственно реализация атаки на информационную систему нарушителем не моделируется.

Построим марковскую модель и рассмотрим математическое описание марковского процесса с дискретными состояниями и непрерывным временем на примере приведенного орграфа угрозы атаки (см. рис. 1), содержащего (для простоты представления) две взвешенные вершины угроз уязвимостей реализации, — угроза атаки создается двумя уязвимостями, с соответствующими их параметрами — интенсивностями выявления и устранения уязвимостей (аналогичным образом можно построить модель для приведенного орграфа угрозы атаки любой сложности). Граф системы состояний случайного процесса (марковского процесса) приведен на рис. 2, а. На графе представлены четыре возможных состояния: S_0 — исходное состояние системы; S_1 — в системе выявлена и не устранена первая уязвимость; S_2 — в системе выявлена и не устранена вторая уязвимость; S_{12} — в системе выявлены и не устранены обе уязвимости — создается реальная угроза атаки. Естественно полагаем, что все переходы системы из одного состояния в другое происходят под воздействием простейших потоков событий с соответствующими интенсивностями выявления или устранения уязвимостей.

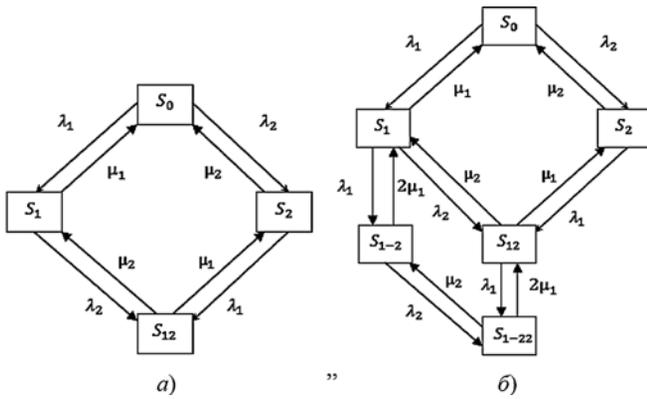


Рис. 2. Графы системы состояний случайного процесса для угрозы атаки:
 а — при условии $\rho \leq 0,2$; б — при условии $\rho > 0,2$

Система дифференциальных уравнений Колмогорова вероятностей состояний для данного графа будет иметь следующий вид:

$$\begin{cases} P'_0 = \mu_1 P_1 + \mu_2 P_2 - (\lambda_1 + \lambda_2) P_0; \\ P'_1 = \lambda_1 P_0 + \mu_2 P_3 - (\lambda_2 + \mu_1) P_1; \\ P'_2 = \lambda_2 P_0 + \mu_1 P_3 - (\lambda_1 + \mu_2) P_2; \\ P'_{12} = \lambda_2 P_1 + \lambda_1 P_2 - (\mu_1 + \mu_2) P_{12}. \end{cases}$$

Заменяя в уравнениях Колмогорова их производные нулевыми значениями, получим систему линейных алгебраических уравнений, описывающих стационарный режим. Решая эту систему с учетом полной группы событий, т.е., используя условие

$$P_0 + P_1 + P_2 + P_{12} = 1,$$

находим искомые предельные (или финальные) вероятности состояний.

Применительно к рассматриваемой задаче моделирования интерес представляет состояние S_{12} — в системе выявлены обе уязвимости. Характеризуемое вероятностью P_{12} это состояние, в котором создаются условия для осуществления атаки (угроза атаки реальна), т.е. выявлены и не устранены все уязвимости, необходимые для осуществления атаки.

Таким образом, эту характеристику можем далее рассматривать в качестве вероятности возникновения угрозы атаки ($P_{y.a} = P_{12}$). Соответственно вероятность готовности к безопасной эксплуатации системы в отношении угрозы атаки P_{0a} (или стационарный коэффициент готовности K_r системы к безопасной эксплуатации) определяется следующим образом:

$$P_{0a} = K_r = P_0 + P_1 + P_2.$$

Замечание. Для графа, представленного на рис. 2, а, P_{0a} можно рассчитывать по следующей формуле:

$$P_{0a} = \frac{\mu_1 \mu_2 + \lambda_1 \mu_2 + \lambda_2 \mu_1}{(\lambda_1 + \mu_1)(\lambda_2 + \mu_2)}. \quad (1)$$

Модель угрозы атаки (см. рис. 2, а) построена в предположении, что для обеих угроз уязвимостей выполняется условие $\rho = \lambda/\mu \leq 0,2$. В работе [1] проведено соответствующее исследование, в результате которого сделан вывод о том, что при условии $\rho \leq 0,2$ для моделирования угрозы уязвимости можно использовать одноканальную схему "гибели и размножения" (одновременно можно устранять только одну выявленную в системе уязвимость), при условии же $\rho > 0,2$ необходимо использовать двухканальную схему (в системе одновременно можно устранять две выявленные уязвимости).

Включение в модель (см. рис. 2, а) двухканальной схемы "гибели и размножения", в предположении, что $\lambda_1/\mu_1 > 0,2$, проиллюстрировано на рис. 2, б.

Для иллюстрации применения рассмотренных моделей на практике, вновь обратимся к примеру угрозы атаки на повышение привилегий, рассмотренному в работе [1]. Подобная угроза атаки предполагает внедрение на компьютер вредоносной программы, что можно рассматривать как безусловную технологическую уязвимость. Использование выявленной уязвимости реализации (выявленной программной ошибки) в компоненте (программе) ядра ОС, запущенного с системными правами, для исполнения внедренного на компьютер в процессе работы вредоносного файла можно рассматривать уже в качестве условной технологической уязвимости системы (возможно при выявлении соответствующей ошибки в системном средстве), с системными правами. Имеем оргграф угрозы атаки, предполагающей последовательное использование нарушителем данных трех уязвимостей (оргграф угрозы атаки содержит в своем составе три взвешенных вершины). Интересующий нас приведенный оргграф угрозы атаки будет содержать одну взвешенную вершину — вершину соответствующей угрозы уязвимости реализации. Для данной угрозы на основании существующей статистики уязвимостей определяем значения соответствующих параметров безопасности — интенсивность выявления рассматриваемой уязвимости составляет 3 в год, устранения — 12 в год, при этом вероятность готовности к безопасной эксплуатации системы в отношении данной угрозы уязвимости составляет $P_{0y} = 0,75$ [1].

Система защиты включается в оргграф угрозы атаки в виде отдельной взвешенной вершины с параметрами безопасности $\lambda_{CЗИ}$ и $\mu_{CЗИ}$ уже собственно системы защиты (это параметры угроз уязвимостей системы защиты, куда включены угрозы уязвимости реализации, безусловные и условные угрозы

технологических уязвимостей) [1]. В результате включения вершины системы защиты приведенный оргграф угрозы атаки с системой защиты от этой атаки будет содержать две взвешенных вершины.

Замечание. Для упрощения расчетов используем в качестве модели такой системы модель, приведенную на рис. 2, а (получим нижнюю — худшую границу характеристики безопасности, так как в данном случае для угрозы уязвимости реализации имеем $\rho = 0,25$, что предполагает включение в модель для этой угрозы двухканальной схемы "гибели и размножения" (см. рис. 2, б).

Оценим, используя выражение (1), изменение вероятности готовности к безопасной эксплуатации защищенной информационной системы в отношении угрозы рассматриваемой атаки P_{0a} при изменении параметров безопасности системы защиты $\lambda_{СЗИ}$ и $\mu_{СЗИ}$. Для этого рассмотрим следующие случаи: интенсивность выявления уязвимостей в системе защиты составляет 5 в год при интенсивности их устранения — 20 в год (вероятность готовности к безопасной эксплуатации системы защиты $P_{0yСЗИ} = 0,75$), интенсивность выявления уязвимостей в системе защиты — 4 в год при интенсивности их устранения 20 в год ($P_{0yСЗИ} = 0,8$), интенсивность выявления уязвимостей в системе защиты — 2 в год при интенсивности их устранения 20 в год ($P_{0yСЗИ} = 0,9$), интенсивность выявления уязвимостей в системе защиты — 1 в год при интенсивности их устранения — 20 в год ($P_{0yСЗИ} = 0,95$). Результаты расчетов приведены в таблице.

Влияние параметров системы защиты на вероятность готовности информационной системы к безопасной эксплуатации

$P_{0yСЗИ}$	0,75	0,8	0,9	0,95
P_{0a}	0,96	0,97	0,98	0,99

Как видим из таблицы, высокий уровень защищенности информационной системы от актуальной угрозы атаки может быть обеспечен при достаточно невысоких (вполне достижимых на практике) требованиях к значениям параметров безопасности системы защиты.

Отметим, что представленные модели отображают следующее важное свойство угрозы атаки: характеристики угрозы атаки зависят только от набора угроз уязвимостей, создающих угрозу атаки, и их параметров безопасности (набора взвешенных вершин в оргграфе атаки) и не зависят от очередности использования выявленных уязвимостей при реализации атаки (не зависят от последовательности переходов между взвешенными вершинами в оргграфе атаки). Как следствие, две угрозы атаки, имеющие одинаковый набор взвешенных вершин

в оргграфе угрозы атаки, эквивалентны — имеют одинаковые характеристики.

Как видим, для расчета значений характеристик угрозы атаки не требуется использования каких-либо экспертных оценок. Адекватность подобной модели угрозы атаки обуславливается использованием объективных значений требуемых для проведения расчетов параметров угроз уязвимостей, получаемых на основании существующей их статистики.

1.2. Укрупненная марковская модель угрозы атаки как системы с отказами и восстановлениями характеристики безопасности. Построение укрупненной модели угрозы атаки необходимо для расчета следующих важнейших характеристик угрозы атаки: интенсивность возникновения λ_a ; интенсивность устранения μ_a реальной угрозы атаки; среднее время наработки на отказ (восстанавливаемая система) характеристики безопасности T_{0ya} , определяющее средний интервал времени между отказами характеристики безопасности — возникновениями реальной угрозы атаки. Эти важнейшие параметры и характеристику угрозы атаки, необходимые для проектирования системы защиты информационной системы, должны рассчитываться исходя из того, что в качестве простейшего элемента безопасности информационной системы при моделировании рассматривается угроза уязвимости — именно для угроз уязвимости с учетом существующей статистики определяются параметры безопасности [1].

Основой построения укрупненной модели является использование параметра потока отказов. В марковских моделях надежности параметр потока отказов ω определяется (для стационарного участка) следующим образом:

$$\omega = \sum_{i \in Q_+} P_i \sum_{j \in Q_-} \lambda_{ij}$$

где Q_+ — множество состояний работоспособности системы; Q_- — множество состояний отказа системы; λ_{ij} — интенсивность перехода из i -го работоспособного состояния, вероятность нахождения в котором системы — P_i , в j -е неработоспособное состояние [3].

Параметр потока отказов, характеризующий частоту возникновения событий отказа в восстанавливаемых системах, обратно пропорционален среднему времени между отказами $T_{моа}$, в западной литературе используется аббревиатура МТВФ (*Mean Time Between Failures*), строгое доказательство этого отношения приведено в теории восстановления:

$$T_{моа} = \frac{1}{\omega} = T_{0y.a} + T_B$$

где T_B — среднее время восстановления.

Исходя из того, что $K_r = \frac{T_{0y.a}}{T_{0y.a} + T_B}$, имеем $T_{0y.a} = \omega K_r$.

Для построения укрупненной модели угрозы атаки вновь обратимся к модели, представленной на рис. 2, а (опять же для наглядности рассматриваем простейший пример), и определимся с тем, как формируется поток отказов характеристики безопасности и каким образом определить его эффективность. Как видим, угроза атаки создается в двух случаях: при переходе из состояния S_1 , в котором система находится с вероятностью P_1 (в марковской модели вероятность состояния интерпретируется как относительная доля времени нахождения системы в этом состоянии), в состояние S_{12} (это состояние реальной угрозы атаки), переходы осуществляются с интенсивностью λ_2 (с учетом же соответствующей доли времени нахождения в состоянии S_1 — с интенсивностью $P_1\lambda_2$), и при переходе из состояния S_2 , в котором система находится с вероятностью P_2 , в состояние S_{12} , переходы осуществляются с интенсивностью λ_1 (с учетом же соответствующей доли времени нахождения в состоянии S_2 — с интенсивностью $P_2\lambda_1$). В нашем случае определяемый подобным образом поток отказов может интерпретироваться как поток возникновения реальной угрозы с интенсивностью λ_a :

$$\lambda_a = \omega = P_1\lambda_2 + P_2\lambda_1.$$

В общем случае значение параметра безопасности угрозы атаки λ_a можно рассчитать по следующей формуле:

$$\lambda_a = \omega = \sum_{i \notin S_{R-1}, i=1, \dots, R} P_{S_{R-1}} \lambda_i$$

где S_{R-1} — множество состояний системы, число которых R , характеризуемых $R-1$ из R возможных (за исключением i -й) выявленных и не устраненных уязвимостей. В каждом из состояний система находится с вероятностью $P_{S_{R-1}}$, из которых осуществляется переход в состояние S_R (все уязвимости выявлены и не устранены — угроза атаки реальна) с интенсивностью λ_i .

Остальные искомые характеристики угрозы атаки рассчитывают по следующим формулам:

$$T_{0y.a} = 1/\lambda_a;$$

$$\mu_a = \frac{\lambda_a P_{0a}}{1 - P_{0a}}.$$

Отметим, что граф системы состояний случайного процесса укрупненной модели будет иметь два состояния: исходное состояние и состояние, характеризующее выявление и неустранение в системе всех уязвимостей — состояние реальной угрозы атаки с соответствующими интенсивностями переходов между ними λ_a и μ_a .

Приведенные марковские модели можно применять для общей оценки свойств безопасности

отдельных средств, в том числе и систем защиты, системных средств, приложений и информационных систем в целом. При проектировании же системы защиты для конкретной информационной системы, используемой для обработки определенной (определенного типа) информации, необходимо учитывать готовность реализации создаваемой в системе реальной угрозы атаки нарушителем, что во многом обуславливается субъективными факторами, определяющими заинтересованность нарушителя в реализации соответствующей атаки на соответствующую информационную систему, используемую для обработки определенной информации. Именно это можно отнести к кардинальным отличиям задачи моделирования в области информационной безопасности от соответствующей задачи моделирования в теории надежности. Таким образом, в разработанные модели необходимо включить состояние фатального отказа, под которым будем понимать успешную реализацию нарушителем атаки на информационную систему. В результате подобного отказа нарушителем осуществляется несанкционированный доступ к информации (например, информация будет похищена), как следствие, в отношении фатального отказа характеристики безопасности система может рассматриваться как невосстанавливаемая.

1.3. Марковские модели угрозы атаки как системы с отказами, восстановлениями и фатальным отказом характеристики безопасности. Состояние фатального отказа в марковской модели угрозы атаки может быть учтено с использованием поглощающей вершины (вершины, не имеющей выхода). Введем понятие коэффициента готовности реализации нарушителем реальной угрозы атаки, обозначим его через $K_{г.а}$, который имеет физический смысл вероятности того, что создаваемая в системе реальная угроза атаки будет реализована нарушителем. Граф системы состояний случайного процесса (марковского процесса), соответствующий системе, граф которой представлен на рис. 2, а, но уже с фатальным отказом, представлен на рис. 3.

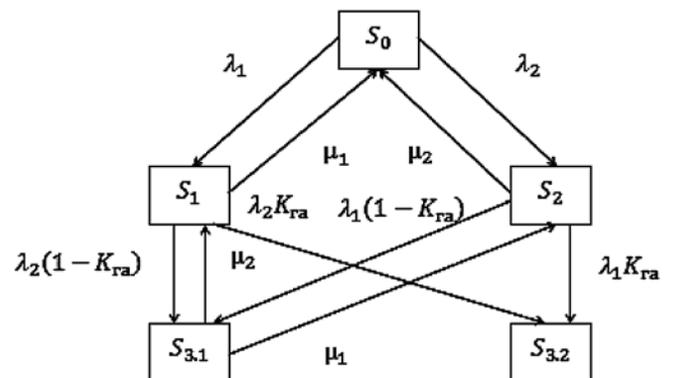


Рис. 3. Граф системы состояний случайного процесса для угрозы атаки с фатальным отказом

При возникновении условия реализации атаки ($P_{y.a} = 1$, соответственно при вероятности готовности к безопасной эксплуатации $P_{0a} = 0$) атака будет реализована потенциальным нарушителем с вероятностью $K_{г.а}$, с вероятностью же $1 - K_{г.а}$ атаки не произойдет. Это учитывается включением в граф системы состояний случайного процесса, представленного на рис. 2, а, вместо вершины S_{12} двух вершин $S_{3.1}$ и $S_{3.2}$ (см. рис. 3). Переход в вершину $S_{3.1}$ предполагает неготовность совершения атаки нарушителем при возникновении ее реальной угрозы (поэтому для этой вершины присутствуют переходы в вершины S_1 и S_2). Переход в вершину $S_{3.2}$ — поглощающую вершину, характеризует реализацию атаки нарушителем на информационную систему.

Как ранее отмечали, значение вероятности P_i состояния (как предельной вероятности) показывает среднее относительное время пребывания системы в i -м состоянии. В данном случае эти вероятности рассчитываются так же, как было описано ранее (с учетом того, что из поглощающей вершины нет выхода). Отличие состоит в интерпретации вероятности $P_{y.a}$ ($P_{y.a} = P_{3.2}$). В данном случае это вероятность реализации успешной атаки на информационную систему. Для вычисления среднего абсолютного времени пребывания системы в каждом i -м состоянии и в системе уравнений Колмогорова нужно положить нулю все производные P'_i ($P'_i = 0$), кроме P'_0 , если считать, что в начальный момент вероятность первого состояния $P_0 = 1$. Тогда на основании теоремы о дифференцировании изображений в преобразовании Лапласа правая часть первого уравнения будет равна -1 . В правых частях уравнений вместо P_i подставляются T_i , и относительно них решается система алгебраических уравнений.

С учетом сказанного, для рассматриваемого примера применительно к графу, приведенному на рис. 3, например, для случая $K_{г.а} = 1$ (отсутствует вершина $S_{3.1}$, вершину же $S_{3.2}$ обозначим как S_3) получаем

$$\begin{cases} -1 = \mu_1 T_1 + \mu_2 T_2 - (\lambda_1 + \lambda_2) T_0, \\ 0 = \lambda_1 T_0 + \mu_2 T_3 - (\lambda_2 + \mu_1) T_1, \\ 0 = \lambda_2 T_0 + \mu_1 T_3 - (\lambda_1 + \mu_2) T_2, \\ 0 = \lambda_2 T_1 + \lambda_1 T_2 - (\mu_1 + \mu_2) T_3. \end{cases}$$

Рассчитав же значения T_i и просуммировав их для состояний, не являющихся поглощающими, можем вычислить важнейшую характеристику — среднее время наработки системы до отказа характеристики безопасности (система с фатальным отказом — невозстанавливаемая), до реализации на нее успешной атаки — реализации угрозы атаки нарушителем $T_{доу.а}$. Например, для системы, описываемой графом, представленным на рис. 3, $T_{доу.а}$ определяется следующим образом:

$$T_{доу.а} = T_0 + T_1 + T_2 + T_{3.1}.$$

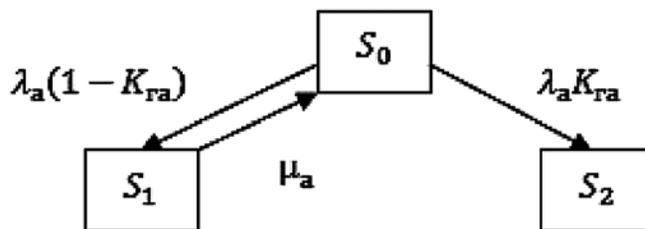


Рис. 4. Граф системы состояний случайного процесса для угрозы атаки с фатальным отказом для укрупненной марковской модели

Граф системы состояний случайного процесса (марковского процесса) укрупненной марковской модели угрозы атаки как системы с отказами, восстановлениями и фатальным отказом характеристики безопасности представлен на рис. 4.

Применение укрупненной марковской модели угрозы атаки как системы с отказами, восстановлениями и фатальным отказом существенно упрощает задачу проектирования системы защиты информационной системы в том случае, когда для используемых в ней средств построены укрупненные марковские модели угрозы атаки как системы с отказами и восстановлениями характеристики безопасности, т.е. определены соответствующие характеристики потенциально возможных для информационной системы угроз атак — интенсивности возникновения λ_a и интенсивности устранения μ_a реальных угроз атак.

Ключевым вопросом возможности и обоснованности практического применения приведенных выше моделей угрозы атаки с фатальным отказом является возможность и обоснованность задания характеристики $K_{г.а}$ — вероятности (коэффициента готовности) осуществить атаку (реализовать угрозу атаки) потенциальным нарушителем — реализовать создавшуюся в информационной системе реальную угрозу атаки. При этом для возможности использования в модели коэффициент $K_{г.а}$ необходимо задавать количественно, причем этот коэффициент для возможности моделирования должен быть универсальным для разнородных угроз атак, создаваемых разнородными угрозами уязвимостей. Естественно, что коэффициент $K_{г.а}$ нужно определять применительно к конкретной информационной системе, обрабатывающей конкретную информацию, которой в конечном счете и определяется заинтересованность и возможность нарушителя в реализации угрозы атаки той или иной сложности на эту систему. Данная задача решается путем построения математической модели потенциального нарушителя безопасности конкретной информационной системы. Опять же для возможности получения адекватных и обоснованных результатов моделирования требованием к построению модели нарушителя является определение значения характеристики $K_{г.а}$ без использования каких-либо экспертных оценок.

2. Математическая модель потенциального нарушителя

Как отмечали, риск реализации атаки на информационную систему невозможно оценить без построения модели потенциального нарушителя, без подобной модели можно оценить лишь риск отказа безопасности информационной системы. Естественно, что данная модель должна учитывать заинтересованность нарушителя в реализации атаки на конкретную информационную систему и его потенциальные возможности (очевидно, что эти характеристики взаимосвязаны).

В настоящее время модель потенциального нарушителя безопасности формируется как набор предположений о возможном нарушителе безопасности, его квалификации, технических и материальных возможностях и т. д. При этом строится неформальная модель нарушителя, отражающая причины и мотивы действий, априорные знания, преследуемые цели, их приоритетность для нарушителя, основные пути достижения поставленных целей: способы реализации исходящих от него угроз, место и характер действия, возможная тактика и т. п. В конечном счете подобная модель используется в целях выявления совокупности актуальных угроз атак для конкретной информационной системы, для которой проектируется система защиты информации — именно актуальных, поскольку потенциально возможные угрозы атак на информационную систему определяются возможностью их технической реализации (архитектура, используемые программные и аппаратные средства и т. д.).

Математическое же моделирование нарушителя сводится к моделированию воздействия нарушителя на защищаемую систему и представляет собой формализованное описание сценариев в виде логико-алгоритмической последовательности действий нарушителей, количественных значений, характеризующих результаты действий, и функциональных (аналитических, численных или алгоритмических) зависимостей, описывающих протекающие процессы взаимодействия нарушителей с элементами защищаемого объекта [4].

Однако подобный подход к моделированию не позволяет количественно оценить актуальность угроз атак, учесть эту важнейшую характеристику безопасности при проектировании системы защиты для конкретной информационной системы.

Для построения математической модели нарушителя в первую очередь введем количественную меру сложности атаки (сложности реализации угрозы атаки), поскольку в общем случае следует говорить о том, готов ли (заинтересован ли и может ли) нарушитель реализовать атаку определенной сложности. При этом необходимо учитывать, что как угрозы уязвимостей, создающие угрозу атаки, так и собственно угрозы атаки по своей сути разнородны, количественная же мера должна быть единой.

Обратимся к основам теории информации, понимая, что для осуществления успешной атаки на отдельно взятую уязвимость, нарушитель должен обладать соответствующей информацией в отношении этой уязвимости — информацией о том, что такая уязвимость выявлена и не устранена, т. е. неким количеством информации в отношении угрозы уязвимости. Так как нас интересует исключительно вероятность того, что уязвимость присутствует в информационной системе, при этом возможны два исхода события: уязвимость присутствует, либо нет. Количество информации в отношении угрозы уязвимости в данном случае следует рассматривать как вероятностную меру.

Замечание. Сложность технической реализации атак на те уязвимости, которые требуют разработки соответствующих программных средств (эксплоитов) для их эксплуатации при реализации атаки, учитывается при задании соответствующего параметра угрозы уязвимости — интенсивности возникновения угрозы уязвимости λ [1]. При задании этого параметра безопасности для угрозы уязвимости на основании анализа соответствующей статистики уязвимостей должна рассматриваться только та часть уязвимостей, для которых за анализируемый период времени подобные exploits были разработаны и использованы [1].

Вероятностная мера количества информации I (в рассматриваемом случае — в одном сообщении) определяется по формуле [5]

$$I = -\log_2 P_i,$$

где P_i — вероятность i -го исхода.

В нашем случае неопределенность можно рассматривать в отношении любой угрозы уязвимости, которую может использовать нарушитель при осуществлении атаки, вероятность ее присутствия (реальная угроза) в системе определяется как $1 - P_{0y}$. Нарушитель для осуществления успешной атаки должен иметь соответствующую информацию, в отношении присутствия уязвимости в системе, т. е. получить сведения, уменьшающие неопределенность в отношении данной угрозы уязвимости. Очевидно, что чем выше для угрозы уязвимости значение P_{0y} (в общем случае уязвимость реже возникает и за меньшее время устраняется), тем сложнее нарушителю осуществить соответствующую атаку.

С учетом сказанного, сложность реализации угрозы уязвимости, обозначим ее S_y , можно интерпретировать как вероятностную меру количества информации $I(P_{0y})$, которой должен обладать злоумышленник для реализации этой угрозы уязвимости, как следствие, может быть определена следующим образом [2]:

$$S_y = I(P_{0y}) = -\log_2(1 - P_{0y}).$$

Корректность применения данной метрики для оценки сложности реализации угрозы уязвимости

обосновывается использованием логарифмической функции (в нашем случае по основанию 2, поскольку у события возможны два исхода), позволяющей соответствующим образом учесть нелинейность функции изменения сложности реализации нарушителем угрозы уязвимости от изменения значения вероятности P_{0y} : $S_y = f(P_{0y})$.

Формулы для расчета характеристики P_{0y} угроз уязвимостей приведены в работе [1].

Проиллюстрируем сказанное примером, для чего сравним сложности реализации двух угроз уязвимостей. Пусть для одной из них значение характеристики P_{0y} составляет 0,7, а для другой — 0,99. Видим, что в первом случае $S_{y1} = 1,74$, во втором случае $S_{y2} = 6,64$, т.е. реализация угрозы второй уязвимости для нарушителя в 3,82 раза сложнее, чем реализация первой уязвимости (ему понадобится в 3,82 раза больше количества информации об угрозе уязвимости в целях снятия неопределенности в отношении наличия в системе этой уязвимости — создания в системе реальной угрозы) для осуществления успешной атаки на угрозу второй уязвимости, чем на угрозу первой уязвимости.

Замечание. Единица сложности реализации угрозы уязвимости $S_y = I(P_{0y}) = 1$ задается условием $P_{0y} = 0,5$, определяющим то, что уязвимость с равной вероятностью присутствует в системе (реальная угроза), либо нет.

Поскольку угрозу атаки создает соответствующая совокупность угроз уязвимостей, сложность атаки для нарушителя в общем случае определяется совокупной сложностью атак на каждую создающую угрозу атаки уязвимость. Если рассмотреть атаку как последовательность использования нарушителем выявленных и не устраненных в системе уязвимостей, имеющих характеристики P_{0yr} и S_{yr} , $r = 1, \dots, R$, можно ввести количественную характеристику сложности атаки $I(P_{0a})$, обозначим ее S_a , где $S_a = I(P_{0a})$, определяемую количеством информации, которым должен обладать нарушитель для осуществления успешной атаки, угрозу которой создают R угроз уязвимостей (с учетом того, что события возникновения (выявления) угроз уязвимостей являются независимыми, а условием реализации нарушителем угрозы атаки является наличие в системе одновременно всех уязвимостей, создающих угрозу атаки):

$$S_a = I(P_{0a}) = -\log_2(1 - P_{0a}) = -\log_2 \prod_{r=1}^R (1 - P_{0yr}),$$

где $P_{0a} = 1 - \prod_{r=1}^R (1 - P_{0yr})$ — вероятность того, что в любой момент времени угроза атаки реальна.

Используя же соответствующее свойство логарифмов, можем записать:

$$S_a = I(P_{0a}) = \sum_{r=1}^R I(P_{0yr}) = \sum_{r=1}^R S_{yr}.$$

При этом информация в отношении угроз уязвимостей, получаемая нарушителем, рассматривается с точки зрения ее полезности (ценности) для достижения потребителем информации поставленной практической цели, в нашем случае для осуществления злоумышленником успешной атаки на информационную систему.

Замечание. Использование в системе системы защиты, призванной увеличить сложность реализации угрозы атаки, увеличивает значение сложности реализации соответствующей угрозы атаки на информационную систему на значение сложности реализации угрозы атаки на систему защиты информации ΔS_a .

Отметим, что характеристика ΔS_a может рассматриваться в качестве так называемой в теории информации прагматической меры количества информации, определяемой в данном случае по формуле

$$\begin{aligned} \Delta S_a &= \log_2(1 - P_{0a.исх}) - \log_2(1 - P_{0a.защ}) = \\ &= \log_2 \frac{(1 - P_{0a.исх})}{(1 - P_{0a.защ})}, \end{aligned}$$

где $P_{0a.исх}$ и $P_{0a.защ}$ — соответственно вероятности готовности к безопасной эксплуатации исходной и защищенной (при использовании системы защиты) информационных систем в отношении угрозы атаки.

Прагматика данной оценки состоит в выявлении условий, при которых необходима реализация соответствующих мер защиты для информационной системы.

Универсальность данной метрики обуславливается тем, что она позволяет сравнивать между собой сложности реализации разнородных атак, основанных на различных принципах реализации, в общем случае использующих совершенно различные угрозы уязвимостей.

Как отмечалось, коэффициент готовности нарушителя осуществить атаку $K_{г.а}$ требуется определять применительно к конкретной информационной системе при проектировании для нее системы защиты. На практике при решении задачи проектирования можно рассматривать (и как правило, рассматривается) некую подобную информационную систему (аналог), характеризуемую обработкой аналогичной информации — именно характеристики обрабатываемой в системе информации определяют заинтересованность и возможности нарушителя. В отношении аналога, как правило, существует соответствующая статистика реализованных (в том числе и отраженных) на информационную систему атак в процессе ее эксплуатации.

С учетом сказанного математическая модель нарушителя (количественная интегральная оценка заинтересованности и возможности реализации нарушителем атаки на конкретную информационную систему) может быть представлена следующим образом:

$$S_{\text{ан}} = \max\{S_{\text{ан}m}, m = 1, \dots, M\},$$

где $S_{\text{ан}}$ — максимальная сложность реализованных (с учетом и отраженных) в подобной (аналогичной) информационной системе атак, характеризующихся $P_{0\text{ан}}$, определяемая на множестве выявленных совершенных атак на подобную информационную систему (аналог) в процессе ее эксплуатации $S_{\text{ан}m}$, $m = 1, \dots, M$.

Рассчитав же значение характеристики $S_{\text{а}}$ — характеристики сложности реализации атаки на информационную систему, для которой проектируется система защиты, применительно к исследуемой угрозе атаки, и значение характеристики $S_{\text{ан}}$ — характеристики максимальной сложности реализованных (в том числе и отраженных) в подобной информационной системе атак, можно определить искомую характеристику коэффициента готовности нарушителя осуществить атаку сложности $S_{\text{а}}$ (реализовать реальную угрозу атаки, в отношении которой проводится исследование) на конкретную информационную систему (для которой проектируется система защиты) $K_{\text{г.а}}$:

$$K_{\text{г.а}} = \begin{cases} \frac{S_{\text{ан}}}{S_{\text{а}}}, & \text{если } S_{\text{ан}} < S_{\text{а}}, \\ 1, & \text{если } S_{\text{ан}} \geq S_{\text{а}}. \end{cases}$$

Заметим, что, исходя из того, что

$$K_{\text{г.а}} = \frac{S_{\text{ан}}}{S_{\text{а}}} = \frac{\log_2(1 - P_{0\text{ан}})}{\log_2(1 - P_{0\text{а}})} = \log_{1 - P_{0\text{а}}}(1 - P_{0\text{ан}}),$$

коэффициент $K_{\text{г.а}}$ может интерпретироваться, как значение степени, в которую надо возвести значение вероятности осуществления атаки на информационную систему $(1 - P_{0\text{а}})$, для получения значения вероятности атаки, которую может успешно реализовать нарушитель $(1 - P_{0\text{ан}})$.

Как видим, для расчета значений искомой характеристики $K_{\text{г.а}}$ при применении рассмотренной модели нарушителя не требуется каких-либо экспертных оценок, опять же используются только параметры безопасности угроз уязвимостей и статистика в отношении безопасности эксплуатации аналогичных информационных систем.

3. Моделирование угрозы атаки с использованием аппроксимирующей функции

На практике при проектировании системы защиты при формировании требований к характеристикам безопасности защищаемой информацион-

ной системы необходима оценка экономической целесообразности реализации системы защиты той или иной сложности, соответственно, той или иной стоимости, включая стоимость ее эксплуатации. Для этого необходима оценка изменения вероятности фатального отказа характеристики безопасности (соответственно, вероятности готовности к безопасной эксплуатации) в процессе эксплуатации информационной системы [6]. Проиллюстрируем сказанное.

Пусть потери от реализации успешной атаки на информационную систему — несанкционированного доступа к информации (в результате нарушения ее конфиденциальности, целостности или доступности) составляют $C_{\text{инф}}$. Тогда риск потерь применительно к угрозе атаки (характеристика угрозы атаки $P_{0\text{а}}$, соответственно $P_{\text{у.а}} = 1 - P_{0\text{а}}$) можно оценить следующим образом:

$$R_{C_{\text{у.инф}}} = C_{\text{инф}}(1 - P_{0\text{а}}).$$

Если использовать при проектировании системы защиты соответствующую марковскую модель, то определив среднее время наработки информационной системы до реализации на нее успешной атаки (фатальный отказ), определим тем самым средний интервал времени эксплуатации системы, через который потери составят $C_{\text{инф}}$. Данный подход к моделированию не дает возможности ответить на вопрос, а каков будет риск потерь на некотором интервале времени эксплуатации системы, меньшем среднего времени наработки информационной системы до отказа характеристики безопасности, и как риск потерь распределен во времени эксплуатации системы. Важность подобной оценки обуславливается тем, что кроме потенциальных потерь, связанных с несанкционированным доступом к обрабатываемой информации, при внедрении системы защиты присутствуют еще и реальные потери, определяемые стоимостью внедряемой системы защиты $C_{\text{СЗИ}}$ и удельной стоимостью (стоимостью в единицу времени) ее эксплуатации $C_{\text{у.эСЗИ}}(t)$. В первом приближении можно рассматривать линейную зависимость изменения стоимости эксплуатации системы защиты во времени. При этом возникает оптимизационная задача задания требуемого значения характеристики защищаемой информационной системы $P_{0\text{а}}$ при проектировании системы защиты, с учетом того, что потенциальные потери от несанкционированного доступа к обрабатываемой информации при условии $t \rightarrow \infty$ стремятся к $C_{\text{инф}}$, в то время как потери, связанные с эксплуатацией системы защиты, при тех же условиях стремятся к ∞ (т.е. задание значения $P_{0\text{а}}$ из условия "чем больше, тем лучше", естественно менее единицы, в общем случае, с учетом сказанного, не корректно).

Задача моделирования состоит в следующем. Как ранее отмечали, в процессе эксплуатации ин-

формационной системы реальная угроза атаки средней продолжительностью $1/\mu_{y.a}$, в случае, если она не будет реализована нарушителем, в среднем через интервалы времени $T_{0y.a}$ будет многократно повторяться (характеристика $T_{0y.a}$ определяется с использованием марковской модели угрозы атаки как системы с отказами и восстановлениями характеристики безопасности). При этом каждую возникающую реальную угрозу атаки злоумышленник может использовать, реализовав атаку с вероятностью $K_{г.а}$. Иллюстрация сказанного приведена на рис. 5.

Рассчитать значение характеристики $P_{y.a}$, достигаемое при эксплуатации системы в некоторый момент времени t , кратный $T_{0y.a}$ при условии $t \geq T_{0y.a}$, обозначив $P_{y.a}(t \geq T_{0y.a})$, можно следующим образом [6]:

$$P_{y.a}(t \geq T_{0y.a}) = \sum_{i=1}^{\lceil t/T_{0y.a} \rceil} K_{г.а} (1 - K_{г.а})^{i-1},$$

где через $\lceil d \rceil$ обозначено меньшее целое числа d .

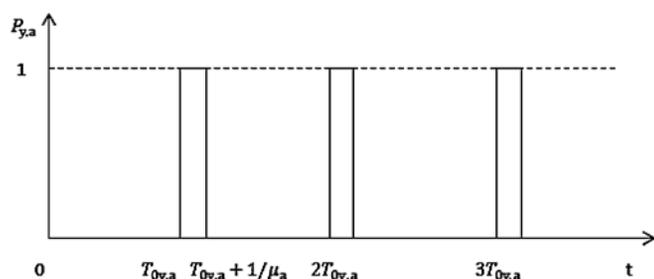


Рис. 5. Иллюстрация появления реальной угрозы атаки в процессе эксплуатации информационной системы

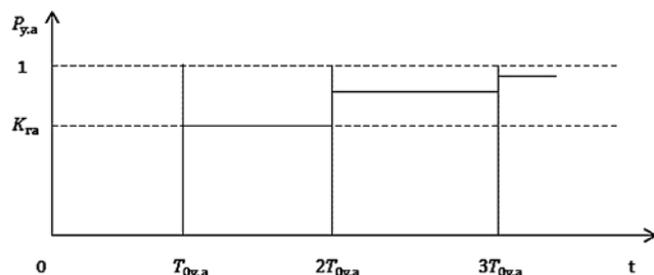


Рис. 6. Иллюстрация изменения характеристики $P_{y.a}(t \geq T_{0y.a})$ в процессе эксплуатации информационной системы

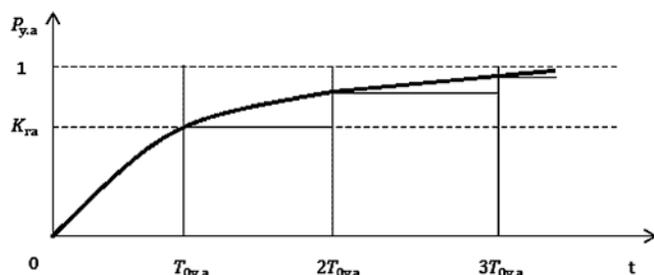


Рис. 7. Иллюстрация аппроксимирующей функции $P_{Ay.a}(t)$

Изменение характеристики $P_{y.a}$ во времени проиллюстрировано рис. 6.

Для расчета значения $P_{y.a}(t \geq T_{0y.a})$ в любой момент времени t эксплуатации информационной системы можно построить и использовать соответствующую аппроксимирующую функцию. Основное правило аппроксимации при этом состоит в том, что значение аппроксимирующей функции, обозначим ее $P_{Ay.a}(t)$, для любого момента времени $t \geq T_{0y.a}$ должно быть не меньше значения функции $P_{y.a}(t \geq T_{0y.a})$ в соответствующий момент времени — аппроксимирующая функция должна предоставлять возможность получения соответствующей граничной оценки, что требуется при проектировании системы защиты (рис. 7).

Таким образом, с использованием построенной подобным образом аппроксимирующей функции в отношении угрозы атаки можно определить вероятность возникновения реальной угрозы атаки $P_{y.a}$ на конкретную информационную систему с учетом готовности реализации этой атаки нарушителем в любой момент времени t эксплуатации информационной системы — вероятность фатального отказа, $P_{Ay.a}(t)$, как следствие и значение потенциальных потерь $R_{C_{y.инф}}(t)$:

$$R_{C_{y.инф}}(t) = C_{инф} P_{Ay.a}(t).$$

В общем случае искомая аппроксимирующая функция имеет следующий вид:

$$P_{Ay.a}(t) = (((1/(1 - K_{г.а}))^{t/T_{0y.a}} - 1)(1 - K_{г.а}))^{t/T_{0y.a}}.$$

Заключение

Представленные в работе математические модели позволяют рассчитывать параметры и характеристики угрозы атаки на информационную систему без необходимости получения каких-либо экспертных оценок, с использованием исключительно стохастических параметров безопасности угроз уязвимостей, в том числе и при построении модели потенциального нарушителя. Данные модели можно применять для формирования требований к значениям параметров безопасности системы защиты при ее проектировании для реализации защиты конкретной информационной системы в целях обеспечения требуемого уровня безопасности информационной системы в отношении угрозы атаки. Не рассмотренными в данной работе остались следующие два ключевых вопроса: как оценить угрозу безопасности информационной системы в целом с учетом подверженности ее множеству угроз атак, причем в общем случае зависимых по угрозам уязвимостей, и как оптимальным образом определить те угрозы технологических уязвимостей, которые необходимо нивелировать системой защиты. Эти вопросы авторы предполагают рассмотреть в следующей работе.

Список литературы

1. Шеглов К. А., Шеглов А. Ю. Интерпретация и моделирование угрозы атаки на информационную систему. Часть 1. Моделирование угрозы уязвимости и интерпретация угрозы атаки // Информационные технологии. 2015. № 12. С. 930–940.
2. Шеглов К. А., Шеглов А. Ю. Математические модели эксплуатационной информационной безопасности // Вопросы защиты информации. 2014. Вып. 106, № 3. С. 52–65.
3. Половко А. М., Гуров С. В. Основы теории надежности. СПб.: БХВ-Петербург, 2006. 704 с.
4. Белов Е. Б., Лось В. П., Мещеряков Р. В., Шелупанов А. А. Основы информационной безопасности. М.: Горячая линия — Телеком, 2006.
5. Шеннон К. Е. Математическая теория связи. Работы по теории информации и кибернетике: пер. с англ. 1963. С. 243–332.
6. Шеглов К. А., Шеглов А. Ю. Эксплуатационные характеристики риска нарушений безопасности информационной системы // Научно-технический вестник информационных технологий, механики и оптики. 2014. № 1 (89). С. 129–139.

К. А. Shcheglov, A. Yu. Shcheglov, Professor,
ITMO University, St. Petersburg, Russian Federation, e-mail: info@npp-itb.spb.ru

Informational System Attack Threat Modeling and Interpretation. Part 2. Attack Threat Modelling

We build Markov's models of attack threat on informational system representing it as a system with failures and recoveries of security characteristics including fatal failures, based on viewing threat as basic element of informational security while interpreting attack threat with a scheme of serial exploit threats reservation. We build the violator model basing on defined attack threat complexity characteristic and interpret it as probabilistic measure of information amount (which must be obtained by violator towards exploit threats creating attack threat). Built models allow to calculate parameters and characteristics of attack threat without need for any experimental estimates with using of exclusively stochastic exploits parameters towards which common statistics exist and being updated.

Keywords: security, informational system, attack threat, model, design, potential violator, informational security system

References

1. Scheglov K. A., Scheglov A. Yu. Interpretaciya i modelirovanie ugrozy ataki na informacionnuyu sistemu. Chast 1. Modelirovanie ugrozy uyazvimosti i interpretaciya ugrozy ataki, *Informacionnye tehnologii*, 2015, vol. 21, no. 12, pp. 930–940.
2. Scheglov K. A., Scheglov A. Yu. Matematicheskie modeli ekspluatacionnoy informacionnoy bezopasnosti, *Voprosy zaschity informacii*, 2014, vol. 106, no. 3, pp. 52–65.
3. Polovko A. M., Gurov S. V. *Osnovy teorii nadezhnosti*, SPb.: BHV-Peterburg, 2006, 704 p.
4. Belov E. B., Los V. P., Mescheryakov R. V., Shelupanov A. A. *Osnovy informacionnoy bezopasnosti*, Moscow, Goryachaya liniya — Telekom, 2006.
5. Shannon K. E. *Matematicheskaya teoriya svyazi. Raboty po teorii informacii i kibernetike*, per. s angl., 1963, pp. 243–332.
6. Scheglov K. A., Scheglov A. Yu. Ekspluatacionnye harakteristiki riska narusheniy bezopasnosti informacionnoy sistemy, *Nauchno-tehnicheskij vestnik informacionnyh tehnologiy, mehaniki i optiki*, 2014, no. 1 (89), pp. 129–139.

ИНФОРМАЦИЯ

*Продолжается подписка на журнал
"ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ" на первое полугодие 2016 г.*

Оформить подписку можно через подписные агентства
или непосредственно в редакции журнала.

Подписные индексы по каталогам:

Роспечать — 72656; Пресса России — 94033

Адрес редакции: 107076, Москва, Стромьинский пер., д. 4,

Издательство "Новые технологии",

редакция журнала "Информационные технологии"

Тел.: (499) 269-55-10. E-mail: it@novtex.ru

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В БИМЕДИЦИНСКИХ СИСТЕМАХ INFORMATION TECHNOLOGIES IN BIOMEDICAL SYSTEMS

УДК 621.317; УДК 616.127-07

В. Г. Полосин, канд. техн. наук, доц. кафедры "Физика", e-mail: polosin-vitalij@yandex.ru,
О. Н. Бодин, д-р техн. наук, проф. кафедры "Информационно-измерительной техники и метрологии",
e-mail: bodin_o@inbox.ru
Пензенский государственный университет, Пенза

Энтропийно-параметрический критерий проверки адекватности модели распределения ионных токов миокарда

Рассмотрено описание различных распределений в едином пространстве централизованных оценок коэффициента энтропии и контрэксцесса, приведенных к их средним квадратическим отклонениям. Предложен метод установления справедливости гипотезы распределения на основе энтропийно-параметрического критерия, рассчитанного по выборочным значениям случайных отсчетов потенциала эпикарда. Показано, что в пространстве приведенных оценок энтропийного коэффициента и контрэксцесса возможно установление справедливости распределения на основе контроля энтропийно-параметрического критерия адекватности модели распределения. Получено выражение зависимости значения энтропийно-параметрического критерия от уровня значимости критерия. Данное выражение позволяет оценить адекватность модели распределения ионных токов миокарда для выборки значений ионных токов в выделенной области пространства приведенных оценок энтропийного коэффициента и контрэксцесса.

Ключевые слова: электрокардиосигнал, модель распределения ионных токов миокарда, энтропийно-параметрический критерий, уровень значимости критерия

Введение

В современных медицинских исследованиях находят широкое применение статистические модели благодаря своей лучшей адаптации к изменениям условий внешней среды за счет оптимальной перестройки физиологических параметров и возможности учета случайных процессов. Среди наиболее распространенных и доступных методов инструментального медицинского исследования следует выделить электрокардиографию сердечно-сосудистой системы, основанную на регистрации биопотенциалов электрического поля с помощью электродов на торсе пациента [1]. В целях повышения разрешающей способности при выявлении места повреждения миокарда разработана компьютерная диагностическая система (КДС) "Кардиовид", которая позволяет на основе анализа электрокардиосигнала (ЭКС) исследовать электрическую активность 816 областей эпикардиальной поверхности [2]. В КДС "Кардиовид" для анализа ЭКС используются:

- значения отсчетов ЭКС, которые содержат важную информацию и отражают состояние эпикарда;
- модель распределения ионных токов миокарда, которая отражает ионные процессы в миокарде;

— энтропийно-параметрический критерий, предназначенный для установления справедливости гипотезы распределения отсчетов ЭКС [3, 4] и проверки адекватности модели распределения ионных токов миокарда.

Цель настоящей статьи — обоснование энтропийно-параметрического критерия для проверки адекватности модели распределения ионных токов миокарда.

Модель распределения ионных токов миокарда

Важные показатели состояния миокарда, такие как проводимость и способность к восстановлению, связаны с характером протекания ионных токов в интервалы времени абсолютной и относительной рефрактерных фаз. Для исследования ионных токов по известной активности опорных точек миокарда разработаны статистические модели, основанные на изменении деполяризующего мембрану заряда [5].

Для описания трансмембранного потенциала действия вводится распределение $F_I(t)$ информационно-измерительных квантов. Это распределение связано как с распределением во времени избы-

Энтропийно-параметрический критерий

точного поляризующего заряда эпикарда $q(t)$, так и с изменением во времени потенциала действия $u(t)$ на его поверхности. Выражение для распределения $F_f(t)$ имеет следующий вид:

$$F_f(t) = 1 - \frac{u(t)}{\Delta B_u} = 1 - \frac{q(t)}{\Delta B_q}, \quad (1)$$

где ΔB_u и ΔB_q — коэффициенты приведения, равные максимальным изменениям потенциала действия и заряда деполяризации.

Распределение $F_f(t)$ отражает изменение деполяризующего заряда эпикарда во времени и содержит информацию о распространении усредненного ионного тока через его поверхность в период рефрактерной фазы. Усредненный ионный ток, протекающий через эпикардальную поверхность, связан с функцией плотности $f_f(t)$ распределения дискретных изменений потенциала с помощью выражения [5]

$$I = \frac{d}{dt} q(t) = \Delta B_q (1 - f_f(t)), \quad (2)$$

где $f_f(t)$ — функция плотности дискретных изменений потенциала на эпикардальной поверхности,

$$f_f(t) = \frac{d}{dt} F_f(t). \quad (3)$$

Для диагностической системы "Кардиовид" число отличимых квантов дискретизации колеблется в пределах от 450 до 1000, что позволяет сделать оценку числа интервалов группирования данных как корень квадратный от числа дискретных величин. Эти значения при использовании методов статистики в системе "Кардиовид" лежат в пределах от 20 до 35.

Для построения статистической модели ионных токов через эпикардальную поверхность предлагается использовать смесь распределений следующего вида:

$$f_f(t, c_1, a_1, c_2, a_2) = K_1 \cdot f_1(t, c_1, a_1) + (1 - K_1) \cdot f_2(t, c_2, a_2), \quad (4)$$

где K_1 — весовой коэффициент смеси; $f_1(t, c_1, a_1)$ и $f_2(t, c_2, a_2)$ — составляющие смеси распределений; c_1, c_2, a_1, a_2 — параметры составляющих смеси распределений.

Задача установления составляющих смеси распределения наблюдаемой случайной величины состоит в выборе таких моделей закона распределения вероятностей, которые наилучшим образом соответствуют результатам экспериментальных наблюдений [6]. Справедливость выбора закона распределения устанавливается на основе статистических критериев, которые обеспечивают математическое обоснование принятия истинной и отклонение ложной гипотез.

В современной литературе широко представлены параметрические критерии установления справедливости гипотез, направленные на проверку параметра математического ожидания или среднего квадратического отклонения [7]. Информационные свойства статистики используются для построения критериев *только* для ограниченного набора функций: критерия для нормального распределения [8]; критерия для распределения Лапласа [9]; критерия для распределения Вейбулла—Гнеденко [10], и *не позволяют* установить справедливость гипотезы в случае смеси распределения наблюдаемой случайной величины. Установление справедливости гипотезы распределения с помощью критерия сводится к расчету по выборочным значениям случайной величины статистического критерия при допущении, что распределение вероятности статистики критерия известно.

Среди известных методов для анализа формы распределения находит применение метод топографических диаграмм Новицкого [11], согласно которому результаты эксперимента отображают в едином пространстве энтропийного коэффициента и контрэксцесса, после чего оценивают совпадение этих параметров с соответствующими значениями нулевой гипотезы. Расчет оценок энтропийного коэффициента и контрэксцесса проводится непосредственно по выборочным данным по формулам [11]

$$K_3 = \frac{\Delta y N}{2\sigma} \exp\left(-\frac{1}{N} \sum_{j=1}^m n_j \cdot \ln(n_j)\right);$$
$$\kappa = \frac{\sigma^4}{\mu_4}, \quad (5)$$

где Δy — интервал группирования данных; N — число отсчетов в выборке; m — число интервалов группирования; n_j — число отсчетов в j -м интервале группирования; σ и μ_4 — среднее квадратическое отклонение и четвертый центральный момент распределения значений отсчетов в выборке данных.

Среди недостатков метода топографических диаграмм необходимо отметить, что несмотря на возможности отражения различных типов распределений в едином пространстве коэффициента энтропии и контрэксцесса, затруднителен выбор предпочтительной реализации на основе анализа формы распределения при использовании различных областей топографической диаграммы.

Для анализа параметров распределения с различной формой предлагается метод установления справедливости гипотезы распределения, который сводится к расчету по выборочным значениям энтропийно-параметрического критерия, учитываю-

шего неупорядоченность и статистические параметры выборки данных.

Для построения единого критерия оценки справедливости гипотез рассмотрим распределение оценок в пространстве энтропийного коэффициента и параметра статистического распределения контрэксцесса вблизи положения нулевой гипотезы H_0 . В пространстве, ограниченном средними квадратическими отклонениями (СКО) $S(k_3)$ и $S(\kappa)$ оценок коэффициента энтропии k_3 и контрэксцесса κ , содержится 47 % всех возможных выборок распределений нулевой гипотезы (коэффициент $t_p = 1$, при вероятности $P = 0,684$). На рис. 1 дана иллюстрация области возможных положений для выборок распределения нулевых гипотез в форме прямоугольника 1 со сторонами, равными удвоенному значению СКО оценок. Формально этой области соответствует область в форме эллипса 2 из условия, что в ней содержится такое же число возможных выборок распределений нулевой гипотезы.

Уравнение кривой эллипса имеет следующий вид:

$$\left(\frac{k_3 - k_{30}}{S(k_3)}\right)^2 + \left(\frac{\kappa - \kappa_0}{S(\kappa)}\right)^2 = \frac{r^2}{b^2}, \quad (6)$$

где k_{30} и κ_0 — коэффициент энтропии и контрэксцесс для нулевой гипотезы H_0 ; $S(k_3)$ и $S(\kappa)$ — средние квадратические отклонения оценок коэффициента энтропии и контрэксцесса, равные интервалам Δk_3 и $\Delta \kappa$ оценок коэффициента энтропии и контрэксцесса при вероятности $P = 0,684$ и коэффициенте $t_p = 1$; b — коэффициент пропорциональности.

Для широкого класса известных симметричных распределений используется единая аналитическая модель вида [11]

$$f(x) = \frac{\alpha}{2\lambda\sigma\Gamma(1/\alpha)} \exp\left(-\left|\frac{x - X_{II}}{\lambda\sigma}\right|\right), \quad (7)$$

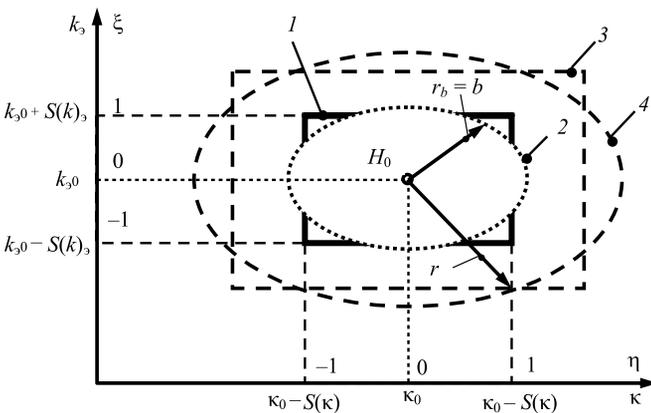


Рис. 1. Соответствие пространств энтропийно-параметрических и приведенных оценок

где $\Gamma(z)$ — гамма-функция; α — показатель степени распределения; σ — среднее квадратическое отклонение; λ — постоянный параметр,

$$\lambda = \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}.$$

Для расчета интервалов оценок коэффициента энтропии и контрэксцесса распределения (5) используются выражения вида [11]

$$\Delta \kappa = t_p \kappa \frac{\sqrt[3]{(\varepsilon^2 - 1)^3}}{\sqrt{29n}}; \quad (8)$$

$$\Delta k_3 = t_p 0,9(\kappa k_3 \sqrt{k_3 n})^{-1}.$$

Авторами работы для проведения сравнения гипотетических распределений различного типа предложено отображать все реализации распределений в едином пространстве, приведенных относительно средних квадратических отклонений $S(\xi)$ и $S(\eta)$, централизованных оценок ξ и η коэффициента энтропии k_3 и контрэксцесса κ , рассчитываемых с помощью выражений вида

$$\xi = \frac{k_3 - k_{30}}{S(k_3)}; \quad (9)$$

$$\eta = \frac{\kappa - \kappa_0}{S(\kappa)}.$$

Область пространства, ограниченная значениями приведенных централизованных оценок ξ и η коэффициента энтропии k_3 и контрэксцесса κ , рассчитанных из выражения (9), в пределах $(-1) \leq \xi_p \leq 1$ и $(-1) \leq \eta_p \leq 1$, содержит 47 % всех возможных выборок распределений нулевой гипотезы.

На рис. 1 показано соответствие пространства приведенных централизованных оценок ξ — η пространству коэффициента энтропии и контрэксцесса, из которого видно, что границам прямоугольной области соответствуют единичные значения приведенных оценок. Уравнение границы соответствующей эллиптической области в пространстве приведенных централизованных оценок ξ — η получено подстановкой выражений (9) для централизованных оценок ξ — η коэффициента энтропии и контрэксцесса в уравнение эллипса (6):

$$r = b \sqrt{\xi^2 + \eta^2}. \quad (10)$$

Из уравнения (10) следует, что граница области, соответствующая выделенной в пространстве приведенных централизованных оценок ξ — η с помощью СКО $S(k_3)$ и $S(\kappa)$ прямоугольной области, задана окружностью радиуса $\sqrt{\pi}$. При увеличении вероятности P попадания выборки для нулевой гипотезы в выделенную область происходит увеличение границ. При этом область круговой границы при заданной вероятности P полностью задана с помо-

шью радиуса r_Γ , связанного с границами центрированных оценок ξ_Γ и η_Γ прямоугольной области с помощью выражения

$$r_\Gamma = \frac{2b}{\sqrt{\pi}} \sqrt{\xi_\Gamma \eta_\Gamma}. \quad (11)$$

Из выражения (11) следует, что параметр r_Γ связан с границами центрированных оценок ξ_Γ и η_Γ , которые определены их средними квадратическими отклонениями:

$$\begin{aligned} \xi_\Gamma &= t_P S(k_3); \\ \eta_\Gamma &= t_P S(\kappa). \end{aligned} \quad (12)$$

Здесь t_P — коэффициент распределения оценок, заданный вероятностью P . Выражение для расчета коэффициента пропорциональности b при заданном радиусе r_Γ границы области допустимых значений и коэффициенте распределения t_P имеет следующий вид:

$$b = \frac{\pi r_\Gamma^2}{2 t_P \sqrt{S(k_3) S(\kappa)}}, \quad (13)$$

где r_Γ — радиус границы области допустимых значений,

$$r_\Gamma = (-2) \ln(1,25 + 0,5 \Phi_0(t_P S(k_3)) + 0,5 \Phi_0(t_P S(\kappa)) + \Phi_0(t_P S(k_3)) \Phi_0(t_P S(\kappa))).$$

Коэффициент пропорциональности определяется из условия равенства вероятностей попадания гипотезы в области пространства, ограниченные оценками границы r_Γ и границами приведенных оценок ξ_Γ и η_Γ :

$$F_r(r_\Gamma) = (0,5 + \Phi_0(\xi_\Gamma))(0,5 + \Phi_0(\eta_\Gamma)).$$

Тогда взаимосвязь радиуса границы области допустимых значений и коэффициента распределений t_P имеет вид

$$r_\Gamma = \sqrt{(-2) \ln(1,25 + 0,5 \Phi_0(\xi_\Gamma) + 0,5 \Phi_0(\eta_\Gamma) + \Phi_0(\xi_\Gamma) \Phi_0(\eta_\Gamma))}.$$

Так как вероятность совершения ошибки первого рода определена коэффициентом распределения t_P и однозначно связана с радиусом границы области допустимых значений в пространстве приведенных оценок, то радиус окружности r_Γ , рассчитанный в приведенной центрированной системе координат, может быть использован в качестве критерия проверки статистических гипотез.

Таким образом, предлагаемый энтропийно-параметрический критерий (см. выражение (10)) позволяет на основе анализа энтропии и статистических параметров выборки отсчетов оценить справедливость принятой гипотезы путем оценки значимости критерия для различных форм симметричных распределений.

Проверка адекватности модели распределения ионных токов миокарда

Построение модели ионных токов миокарда в виде выражения (4) основано на формировании искомой функции с помощью смеси распределений. Для этого все значения нормируются на площадь под кривой. Тогда число значений отсчетов характеризует появление отдельного отсчета с помощью вероятности его наблюдения. Для аппроксимации таких кривых удобно использовать математический аппарат статистики. Отношение числа квантов, попадающих в отдельные интервалы группирования данных, к временному интервалу группирования позволяет построить гистограмму распределения частоты наблюдения квантов в зависимости от временных интервалов наблюдений. Формирование модели распределения ионных токов (4) с помощью смеси распределений иллюстрирует рис. 2, где обозначено 1 — гистограмма распределения квантов для трансмембранного потенциала; 2 — смесь плотностей экспоненциального $f_1(t_i, c_1, a_1)$ и логистического $f_2(t_i, c_2, a_2)$ распределений; 3 — смеси плотностей экспоненциального $f_1(t_i, c_1, a_1)$ и нормального $f_2(t_i, c_2, a_2)$ распределений; 4 — смеси плотностей двухпараметрического $f_1(t_i, c_1, a_1)$ и трехпараметрического $f_2(t_i - t_0, c_2, a_2)$ распределений Вейбулла—Гнеденко.

Выбор наиболее оптимальной смеси распределения связан с задачей проверки адекватности модели, которая сводится к установлению правомерности выбора распределения при описании полученной выборки отсчетов. Для этих целей используется энтропийно-параметрический критерий, позволяющий на основе анализа энтропии и статистических параметров выборки отсчетов оценить справедли-

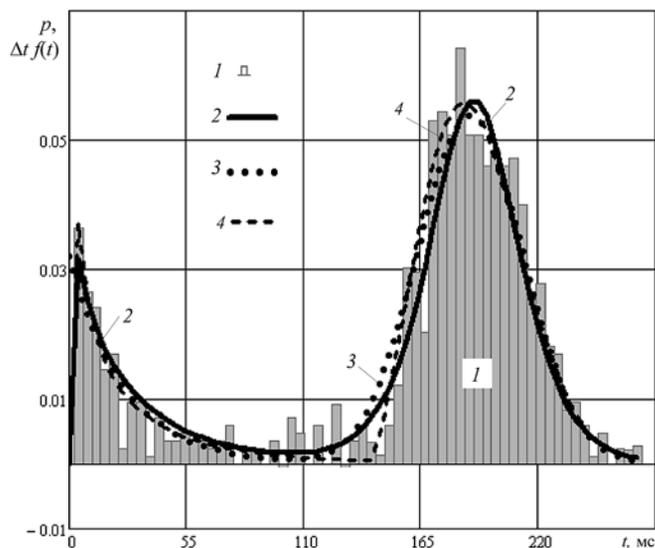


Рис. 2. Гистограмма распределения квантов по временным интервалам и аппроксимации распределения для трансмембранного потенциала

вость принятой гипотезы путем оценки значимости критерия для различных форм симметричных распределений.

Установление соответствия модели реальным физическим процессам, происходящим в системе, состоит в проверке адекватности модели распределения ионных токов и содержит три основных этапа.

На первом этапе проводится разделение смеси распределения на части, соответствующие составляющим модели. Суть разделения состоит в том, что каждому значению тока, протекающего через эпикардиальную поверхность, соответствуют коэффициенты статистической значимости, которые определяют вклад отдельных составляющих смеси распределений в распределение отсчетов. Коэффициенты статистической значимости q_{1i} и q_{2i} для отсчета тока в i -й момент времени можно найти как отношение значений составляющих плотности распределений $f_1(t_i, c_1, a_1)$ и $f_2(t_i, c_2, a_2)$ информационных квантов к значению плотности смеси распределения $f_I(t_i, c_1, a_1, c_2, a_2)$, найденных при известных параметрах проверяемой гипотезы распределения. Формулы для расчета коэффициентов значимости имеют вид:

$$\begin{aligned} q_{1i} &= \frac{f_1(t_i, c_1, a_1)}{f_I(t_i, c_1, a_1, c_2, a_2)}; \\ q_{2i} &= \frac{f_2(t_i, c_2, a_2)}{f_I(t_i, c_1, a_1, c_2, a_2)}. \end{aligned} \quad (14)$$

Расчет коэффициентов значимости проводится при различных значениях весовых коэффициентов K и $(1 - K)$ составляющих смеси распределений, используемой в модели ионных токов. Для выбора наиболее предпочтительного варианта необходимо установить соответствие аппроксимаций выборке экспериментальных данных.

Второй этап проверки адекватности модели состоит в преобразовании полученных данных к симметричной форме. В работе [4] показано, что целый ряд несимметричных распределений допускает их преобразование к симметричным распределениям. Соответственно, если несимметричное распределение является описанием выборки данных, то после преобразования выборки в соответствии с преобразованием несимметричного распределения к симметричной форме, получим выборку, удовлетворяющую симметричному распределению. Тогда проверка адекватности модели распределения ионных токов сводится к расчету энтропийно-параметрического критерия и оценки его уровня значимости для каждой из составляющих модели. Если признается справедливой нулевая гипотеза для обоих распределений при заданном значении уровня значимости критерия, то признается соответствие модели исходной выборке значений.

На третьем этапе рассчитывается энтропийно-параметрический критерий для выборок симмет-

ричных распределений и оценивается значимость критерия. Для оценки значимости критерия рассмотрим распределение реализаций в пространстве приведенных центрированных оценок ξ и η .

При большом числе отсчетов выборки N ($N > 100$) эпикардиального потенциала оценки коэффициента энтропии k_ξ и контрэксцесса κ на два порядка больше их разбросов Δk_ξ и $\Delta \kappa$, что позволяет для приближенной аппроксимации распределения оценок k_ξ и κ использовать нормальное распределение Гаусса. Плотности распределения приведенных центрированных оценок ξ и η в пространстве приведенных оценок равны

$$\begin{aligned} f_1(\xi) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}}; \\ f_2(\xi) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{\eta^2}{2}}. \end{aligned} \quad (15)$$

Плотности распределения (15) приведенных центрированных оценок ξ и η позволяют получить выражения для расчета вероятности возникновения ошибки первого рода в результате принятия неверной нулевой гипотезы, равной уровню значимости α энтропийного критерия r .

Элементарная вероятность $\Delta P(\xi, \eta)$ попадания оценок ξ и η для выборки результатов нулевой гипотезы в элементарную область, ограниченную приращениями $\Delta \xi$ и $\Delta \eta$ приведенных оценок, имеет вид

$$\Delta P(\xi, \eta) = f_1(\xi) \Delta \xi f_2(\eta) \Delta \eta. \quad (16)$$

Сделав подстановку плотности распределений (15) оценок ξ и η в выражение (16) и заменив элемент площади приращения $\Delta \xi$ и $\Delta \eta$ площадью кольца радиуса r элементарной ширины Δr , получим плотность распределения $f_r(r)$ случайной величины модуля расстояния r между нулевой гипотезой и приведенными оценками выборки результатов:

$$f_r(r) = r e^{-\frac{1}{2}r^2}. \quad (17)$$

Плотность распределения (17) для модуля расстояния r представляет собой плотность вероятности классического двухпараметрического распределения Вейбулла—Гнеденко с параметрами масштаба и формы, равными $\sqrt{2}$ и 2, соответственно. Тогда аналитическое выражение для распределения модуля расстояния r имеет вид

$$F_r(r) = 1 - e^{-\frac{1}{2}r^2}. \quad (18)$$

Из анализа распределения (18) следует, что вероятность $F_r(r)$ однозначно определена значением случайной величины r и не содержит каких-либо других параметров, влияющих на распределение модуля расстояния r . Такую случайную величину удобно использовать в качестве критерия для проверки гипотез различных типов симметричных распределений. Для уровня значимости $\alpha(r)$ критерия r будет справедливо выражение вида

$$\alpha(r) = 1 - F_r(r) = e^{-\frac{1}{2}r^2}. \quad (19)$$

Тогда для расчета критерия при различных уровнях значимости справедливо выражение вида

$$r(\alpha) = \sqrt{\ln(\alpha^{-2})}. \quad (20)$$

В зависимости от требуемой надежности результата при решении задач пределы уровня значимости α выбирают от 20 до 0,1 %, что определяет границы задания критерия r от 1,794 до 3,717.

В таблице даны значения параметров смесей и значения энтропийно-параметрического критерия.

Параметры статистической модели ионных токов

Состав смеси распределения для статистической модели ионных токов	Параметры составляющих смесей распределений		Весовой коэффициент	Значение критерия r
	$f_1(t)$	$f_2(t)$		
Экспоненциальное и логистическое распределения	$\lambda = 20,04$	$\mu = 197$ $\lambda = 17,6$	0,23	$r_1 = 2,451$ $r_2 = 2,741$
Экспоненциальное и нормальное распределения	$\lambda = 26,9$	$\sigma = 22,67$ $\mu = 190,04$	0,25	$r_1 = 3,26$ $r_2 = 2,12$
Распределение Вейбула—Гнеденко и распределение Вейбула—Гнеденко	$c_1 = 2,339$ $a_1 = 55,1$ $t_0 = 141$	$c_2 = 0,741$ $a_2 = 25$	0,755	$r_1 = 3,101$ $r_2 = 2,859$

Важное свойство энтропийно-параметрического критерия r состоит в том, что в пространстве приведенных оценок ξ и η нулевой гипотезы критерий позволяет проводить сравнительный анализ различных симметричных распределений. Модель принимается адекватной, если энтропийно-параметрический критерий не превышает своего предельного значения, установленную с помощью уровня значимости и рассчитанного на основе соотношения (20).

Из анализа результатов таблицы следует, что энтропийно-параметрический критерий не превышает предельного значения 3,04, ограниченного уровнем значимости α , равного 0,01, для двух смесей распределений. Первая смесь распределений содержит составляющие экспоненциального и логистического распределения, вторая смесь состоит из двух распределений Вейбулла—Гнеденко. При выборе из двух полученных распределений предпочтение следует отдать смеси экспоненциального и логистического распределения. Таким образом, энтропийно-параметрический критерий является эффективным инструментом проверки адекватности статистической модели ионных токов на эпикардиальной поверхности.

Список литературы

1. Орлов В. Н. Руководство по электрокардиографии. М.: Медицина, 1983. 527 с.
2. Бодин О. Н., Митрохина Н. Ю. Регуляризация решения обратной задачи электрокардиографии в компьютерной диагностической системе "Кардиовид" // Мехатроника, автоматизация, управление. 2008. № 11. С. 37—40.
3. Вёгер В. Информация об измеряемой величине как основа формирования функции плотности вероятности // Измерительная техника. 2003. № 9. С. 3—9.
4. Полосин В. Г., Першенков П. П. Информационный способ проверки гипотез несимметричных распределений // Измерительная техника. 2013. № 12. С. 8—10.
5. Полосин В. Г., Бодин О. Н., Балахонова С. А. Статистические модели трансмембранного потенциала миокарда в компьютерной диагностической системе "Кардиовид" // Математическая биология и биоинформатика: V Международная конф., г. Пушкино, 19—24 октября 2014 г.: Доклады / Под ред. В. Д. Лахно. М.: МАКС Пресс, 2014. 228 с.
6. Лемешко Б. Ю. О задаче идентификации закона распределения случайной составляющей погрешности измерения // Метрология. 2004. № 7. С. 8—17.
7. Брандт З. Анализ данных. Статистические и вычислительные методы для научных работников и инженеров. М.: Мир, Издательство АСТ, 2003. 686 с.
8. Vasicsek O. A test for normality based on sample entropy // JRSS. 1976. Vol. 38, N. 1. P. 54—59.
9. Dudewcz E. J., van der Meulen E. C. Entropy-based tests of uniformity // JASA. 1981. Vol. 76, N. 376. P. 967—974.
10. Kolossvary I., Telek M. Explicit identification of the class of order 3 matrix exponential distributions. 2011. At Eurandom, Eindhoven, Netherlands.
11. Новицкий П. В., Зограф И. А. Оценка погрешностей результатов измерений. Л.: Энергоатомиздат. 1985. 284 с.
12. Боровков А. А. Математическая статистика. Оценка параметров проверки гипотез. М.: Наука., Главная редакция физико-математической литературы, 1984. 472 с.
13. Тычков А. Ю., Чураков П. П., Кривоногов Л. Ю. Автоматизированная система обработки и анализа электрокардосигналов в условиях интенсивных помех различного вида // Известия высших учебных заведений. Поволжский регион. Технические науки. 2011. № 1 (17). С. 117—125.

V. G. Polosin, Ph.D., Associate Professor,
Senior Lecturer of the Department "Physics", polosin-v@sura.ru,
O. N. Bodin, Professor of the Department
"Information Measuring Techniques and Metrology", bodin_o@inbox.ru
Penza State University, Penza

Entropy-Parametric Criterion for the Establishing the Adequacy of the Distribution Model of Ionic Currents of the Myocardium

The article is devoted to reviewing the adequacy of the distribution model of the ion currents the myocardium based on the entropy-parametric criterion. This paper considers the possibility of describing the various distributions in a uniform space of estimates of the entropy factor and of the counter kurtosis. These estimates were centered with respect to the null hypothesis of distribution, for which the entropy factor and counter kurtosis are known a priori. Estimates of the entropy factor and of the ratio kurtosis have been given as the ratio on their standard deviation that made it possible to build them into a single space. We have proposed a method of establishing validity of the hypothesis of distribution. It is based on the entropy-parametric criterion that is calculated by values of the random sample of potential. These samples were calculated for the surface of the epicardium. It is shown that in the space of relative estimates of the entropy factor and of the counter kurtosis it is possible to establishing the justice and the adequacy for the statistical model that is built as the sum of the distributions. In particular, we have showed that in the space of the reduced estimates can be established the integral part of the additive model of the distributions of ionic currents. This is based on control of the weighting factor component of the model. We obtain expressions the dependence of the values of the entropy-parametric criterion when the significance level criterion. This expression allows you to assess the adequacy of the distribution model of myocardial ion currents to the sample values of the ion currents in the selected area of the space that is defined as the ratio of estimates of the entropy factor and counter kurtosis on their known values.

Keywords: electrocardiosignal, the entropy factor, the counter kurtosis, the entropy-parametric criterion, a single space, to establishing the justice, the surface of the epicardium, the significance level criterion, the myocardial ion currents, model of ionic currents of the myocardium

References

1. Orlov V. N. *Rukovodstvo po elektrokardiografii* (Guide electrocardiography.) Moscow, Meditsine, 1983, 527 p.
2. Bodin O. N., Mitrohina N. Yu. Regularizatsiya resheniya obratnoy zadachi elektrokardiografii v kompyuternoy diagnosticheskoy sisteme "Kardiovid" (Regularization of the inverse problem of electrocardiography in the computer diagnostic system "Kardiofit"), *Mehatronika. Avtomatizatsiya. Upravlenie* (Mechatronics. Automation. Control.), 2008, no. 11, pp. 37–41.
3. Vyoger V. Informatsiya ob izmeryaemoy velichine kak osnova formirovaniya funktsii plotnosti veroyatnosti (Information about the measured value as a basis of the probability density function), *Izmeritel'naya tekhnika* (Measurement Technique), 2003, no. 9, pp. 3–9.
4. Polosin V. G., Pershenkov P. P. Informatsionnyy sposob proverki gipotez nesimmetrichnykh raspredeleniy (Information — theoretic method for hypothesis testing with nonsymmetric distributions), *Izmeritel'naya tekhnika* (Measurement Techniques), 2013, no. 12, pp. 8–10. (Vol. 56, no. 12, march, 2014, pp. 1318–1322).
5. Polosin V. G., Bodin O. N., Balahonova S. A. Statisticheskie modeli transmembrannogo potentsiala miokarda v kompyuternoy diagnosticheskoy sisteme "Kardiovid" (Statistical models of the transmembrane potential of the myocardium in the computer diagnostic system "Kardiofit"), *Matematicheskaya biologiya i bioinformatika: V Mezhdunarodnaya konf.* (Mathematical Biology and Bioinformatics: International conf), g. Puschino, 19–24 okt. 2014: Doklady, Ed. V. D. Lahno, Moscow, MAKS Press, 2014, 228 p.
6. Lemeshko B. Yu. O zadache identifikatsii zakona raspredeleniya sluchaynoy sostavlyayushey pogreshnosti izmereniya (On the problem of identifying the law of distribution of the random component of measurement error), *Metrologiya* (Metrology), 2004, no. 7, pp. 8–17.
7. Brandt Z. *Analiz dannykh. Statisticheskie i vychislitelnyye metody dlya nauchnykh rabotnikov i inzhenerov* (Data analysis. Statistical and Computational Methods for Scientists and Engineers), Moscow, Mir, OOO "Izdatelstvo AST", 2003, 686 p.
8. Vasicek O. A test for normality based on sample entropy, *JRSS*, 1976, vol. 38, no. 1, pp. 54–59.
9. Dudewcz E. J., van der Meulen E. C. Entropy-based tests of uniformity, *JASA*, 1981, vol. 76, no. 376, pp. 967–974.
10. Kolossvary I., Telek M. Explicit identification of the class of order 3 matrix exponential distributions, *YEQT V "Stochastic Networks and Optimization"*, October 24–26, 2011. At Eurandom, Eindhoven, Netherlands. <http://www.eurandom.nl/events/workshops/2011/YEQTV>
11. Novitskiy P. V., Zograf I. A. Otsenka pogreshnostey rezultatov izmereniy, Leningrad, Energoatomizdat, 1985, 284 p.
12. Borovkov A. A. Matematicheskaya statistika. Otsenka parametrov proverki gipotez (Math statistics. Parameter estimation of testing hypotheses, Moscow, Nauka, Glavnaya redaktsiya fiziko-matematicheskoy literatury, 1984, 472 p.
13. Tyichkov A. Yu., Churakov P. P., Krivonogov L. Yu. Avtomatizirovannaya sistema obrabotki i analiza elektrokardiosignalov v usloviyakh intensivnykh pomekh razlichnogo vida (Automated system for processing and analyzing electrocardiosignals under intensive interference of various kinds), *Izvestiya vysshih uchebnykh zavedeniy. Povolzhskiy region. Tehnicheskie nauki* (Proceedings of the higher educational institutions. Volga region. Technical science), 2011, no. 1 (17), pp. 117–125.

Н. Т. Абдуллаев¹, канд. техн. наук, доц., e-mail: a.namik46@mail.ru,

М. М. Гасанкулиева², диссертант,

¹ Азербайджанский технический университет

² Азербайджанская государственная нефтяная академия

Информативные параметры и флуктуационная динамика электрокардиограмм при фликкер-шумовой спектроскопии

На основе дискретно фиксируемых временных рядов электрокардиографических сигналов проводится анализ спектральной мощности и переходных разностных моментов второго порядка. Такой анализ позволяет получить информативные диагностические параметры, оценивающие функциональное состояние сердечно-сосудистой системы пациента. Одним из факторов, позволяющих оценить флуктуационную динамику электрокардиограмм и выделить вклад нерегулярностей в анализируемые реальные сигналы, может быть вариация используемых частот дискретизации.

Ключевые слова: фликкер-шумовая спектроскопия, электрокардиография, дифференциальная диагностика, информативные параметры, частота дискретизации сигнала

Введение

Основная масса исследований функционального состояния сердечно-сосудистой системы ведется с применением статистической обработки, временного и спектрального анализа ритмограмм и последующей разработкой различных методик, основанных на этих методах.

При исследованиях открытых живых динамических систем, например организма человека, условия эргодичности и стационарности регистрируемых процессов не выполняются, и использование в целях анализа состояния организма осредненных статистических характеристик медицинских показателей может служить лишь грубой оценочной информацией по этим процессам.

В области исследования сердечно-сосудистой системы выделяют в основном методы анализа, строго математически применимые к стационарным сигналам, однако в медицинской практике преобладают нестационарные сигналы, статистические свойства которых изменяются со временем. Часто они состоят из кратковременных высокочастотных компонентов, сопровождаемых длительными низкочастотными составляющими.

Учитывая такой характер биоэлектрических потенциалов, и в частности электрокардиографических сигналов, наиболее приемлемым для их анализа может быть метод нелинейной динамики с вычислением количественных характеристик хаоса. Таковую возможность представляет метод фликкер-шумовой спектроскопии, учитывающий в сложной динамической системе эффект перемежаемости, когда участки хаотических всплесков и скачков чередуются относительно длительными участками ламинарного характера. Анализ сигналов, имеющих такой динамический характер, основывается, как правило, на

применении метода фликкер-шумовой спектроскопии.

Постановка проблемы

При анализе оцифрованных сигналов возникает вопрос о характере информации, которая может быть извлечена из исследуемого сигнала $V(t)$ в случае, если это сложный хаотический сигнал, фиксируемый с частотой дискретизации f_d . Электрокардиографическому сигналу присущи основные для режима "перемежаемости" черты, когда "ламинарные" фазы с относительно небольшими изменениями динамической переменной $V(t)$ на временных интервалах с достаточно большой длительностью T прерываются кратковременными в течение времени τ ($\tau \ll T$) всплесками значений $V(t)$, характерными для последующей во времени "ламинарной" фазы [1]. Анализ электрокардиограммы (ЭКГ) показывает, что она соответствует описанной динамике, когда "всплески" в виде QRS-комплексов чередуются достаточно малыми скачками в виде P - и T -зубцов и протяженными фазами в виде изолинии ("ламинарная" фаза) [2]. Величина и продолжительность таких "скачков", "всплесков" и "ламинарных" участков специфичны для каждой из сердечно-сосудистых систем, обуславливая определенный вклад в соответствующий спектр мощности и определяя совокупность информативных диагностических параметров, характеризующих корреляционные взаимосвязи в таких последовательностях.

Применение метода фликкер-шумовой спектроскопии [3] позволяет придать информационную значимость последовательностям различных нерегулярностей (всплески, скачки, изломы производных различных порядков) динамических перемен-

ных исследуемых систем и, в частности, сердечно-сосудистой системы организма. Принятие нерегулярностей динамических переменных в качестве информационной основы позволяет классифицировать всю содержащуюся в хаотических сигналах информацию. Как правило, с помощью фликкер-шумовой спектроскопии анализируются спектр мощности сигнала $S(f)$ и переходные разностные моменты второго порядка $\Phi^{(2)}(\tau)$. Информация, извлекаемая из анализа зависимостей $S(f)$ и $\Phi^{(2)}(\tau)$, построенных на основе временных рядов $V(t)$, имеет смысл времени корреляции или параметров, характеризующих потерю корреляционных связей для рассматриваемых нерегулярностей типа всплесков и скачков. В таком случае исследуемый ЭКГ-сигнал удобно представить в виде суммы двух слагаемых: сингулярного члена $V_s(t)$, который формируется лишь всплесками динамической переменной, и регулярного члена $V_r(t)$, который формируется после вычитания всплесков из представленного сигнала и определяется скачками динамической переменной и "ламинарными" фазами. Каждая из составляющих исследуемого сигнала характеризуется совокупностью извлекаемых параметров $T_0, n_0, S_s(0)$ — для сингулярной составляющей и H_1, T_1, σ_1 — для регулярной составляющей сигнала [3, 4].

Для сингулярной составляющей сигнала $V(t)$ параметр T_0 определяет некоторое характерное время, в пределах которого реализуется взаимосвязь измеряемой динамической переменной $V(t)$. Безразмерный параметр n_0 эффективно определяет, каким образом эта взаимосвязь теряется по мере уменьшения частот до значений $1/(2\pi T_0)$. Параметр $S_s(0)$ — это вклад в спектр мощности $S(f)$, определяемый наиболее высокочастотной составляющей, формируемый "нерегулярностями — всплесками".

Для регулярной составляющей сигнала $V(t)$ параметр T_1 определяет характерное время, на котором величины динамических переменных $V(t_i)$ не коррелируют (если их аргументы различаются на величину T_1). Для получения достоверных значений дисперсии σ_1^2 необходимо ее рассчитывать на временных интервалах, превышающих T_1 . При этом параметр H_1 (показатель Херста) показывает по какому закону теряется взаимосвязь измеряемых в разные моменты времени величин $V(t_i)$.

В общем случае при анализе сложного хаотического сигнала, измеряемого при некоторой частоте дискретизации f_d , определяется совокупность указанных параметров, характеризующая корреляционные взаимосвязи в последовательностях нерегулярностей-скачков и нерегулярностей-всплесков, свойственных данному сигналу, определяемому с частотой дискретизации f_d . Таким образом, одним из основных факторов, позволяющим реализовать выделение вклада нерегулярностей в анализируемые

реальные сигналы, является вариация используемых частот дискретизации f_d . Если анализируемый временной ряд получен при достаточно высокой частоте дискретизации f_d , то анализ зависимостей $\Phi^{(2)}(\tau)$ и $S(f)$, рассчитываемых на основе временных рядов, полученных из исходного временного ряда с уменьшающейся частотой дискретизации, позволяет оценить меру "устойчивости" параметров σ_1, T_1 и H_1 (для $\Phi^{(2)}(\tau)$) и меру вариабельности параметров $S_s(0), T_0$ и n_0 (для $S(f)$).

Результаты вычислительного эксперимента и их анализ

Высокая специфичность зависимостей $\Phi^{(2)}(\tau)$ и $S(f)$, получаемых при анализе состояния сложных систем, может быть использована для диагностики болезней, так же как совокупность указанных параметров для их классификации. Анализируются ЭКГ-сигналы четырех типов — норма и сердечные "катастрофические" аритмии, непосредственно угрожающие жизни пациента (желудочковая тахикардия, предсердная фибрилляция, предсердная аритмия). Для выявления особенностей анализируемых сигналов необходимо оценить всю совокупность полученных оцифрованных данных $V(t)$ ЭКГ для указанных состояний сердечно-сосудистой системы. При проведении вычислительного эксперимента использовались экспериментальные данные с общедоступного сайта www.PhysioNet.org [5].

Сигналы снимались со II-стандартного отведения в течение ~60 с с частотой дискретизации $f_d = 500$ Гц и числом точек $N = 29\,859$. Таким образом, был получен временной ряд ЭКГ-сигналов при достаточно высокой частоте дискретизации f_d , поскольку на его основе можно получить набор новых временных рядов при частотах дискретизации, меньших f_d в кратное число раз.

Результаты соответствующего анализа для указанных функциональных состояний сердечно-сосудистой системы при частоте дискретизации ЭКГ-сигналов $f_d = 500$ Гц приведены в табл. 1.

Проведем сравнительный анализ информативных параметров для двух состояний сердечно-сосудистой системы: норма (табл. 2) и желудочковая тахикардия (табл. 3) для частот дискретизации $f_d = 500$ Гц и $f_d = 250$ Гц.

Из полученных таблиц следует, что при увеличении частоты дискретизации f_d возрастает высокочастотный вклад в спектр мощности $S(f)$ за счет включения в анализируемый сигнал "всплесков", соответствующих увеличенной частоте f_d . При этом также происходят изменения зависимости $\Phi^{(2)}(\tau)$ при малых τ , которые обусловлены вкладом локальных изменений в величины "ламинарных" участков сигнала. Поэтому при возрастании f_d изменяются как параметры T_0 и n_0 , характеризующие

Информативные диагностические параметры для различных функциональных состояний сердечно-сосудистой системы

ЭКГ-сигнал	Сингулярная составляющая			Регулярная составляющая		
	$S_s(0)$, отн. ед.	T_0 , отн. ед.	n_0	σ_1 , отн. ед.	H_1	T_1 , отн. ед.
Норма	437,80	0,0042	0,3414	0,55	1,1133	1,5080
Желудочковая тахикардия	197,358	0,0032	0,4123	0,51	1,0913	0,6840
Предсердная фибрилляция	334,364	0,0036	0,3836	0,435	1,1388	0,0640
Предсердная аритмия	43,7105	0,0059	0,4013	0,208	1,1298	1,1560

Таблица 2

Норма

f_d , Гц	N	Сингулярная составляющая			Регулярная составляющая			$\frac{4S(0)}{N}$
		$S_s(0)$, отн. ед.	T_0 , отн. ед.	n_0	σ_1 , отн. ед.	H_1	T_1 , отн. ед.	
500	29 859	437,80	0,0042	0,3414	0,55	1,1133	1,5080	0,05
250	14 930	403,72	0,0028	0,4187	0,5044	1,0845	1,38	0,08

Таблица 3

Желудочковая тахикардия

f_d , Гц	N	Сингулярная составляющая			Регулярная составляющая			$\frac{4S(0)}{N}$
		$S_s(0)$, отн. ед.	T_0 , отн. ед.	n_0	σ_1 , отн. ед.	H_1	T_1 , отн. ед.	
500	29 859	197,358	0,0032	0,4123	0,5180	1,0913	0,6840	0,026
250	14 930	175,80	0,0034	0,3446	0,517	1,5200	0,340	0,10

высокочастотную область зависимости $S(f)$, так и параметры H_1 и T_1 , характеризующие зависимость $\Phi^{(2)}(\tau)$ при малых τ . Величина параметра σ_1 , а также характер спектральной зависимости $S(f)$ изменяются в значительно меньшей степени. Небольшие вариации параметра среднеквадратичного отклонения σ_1 указывают на меньшую зависимость функции $\Phi^{(2)}(\tau)$ от f_d . В то же время анализ сигналов при фликкер-шумовой спектроскопии выявляет динамику изменений параметров H_1 и T_1 при малых τ , а также параметров T_0 и n_0 , характеризующих зависимость $S(f)$ в высокочастотной области. Поскольку зависимость $S(f)$ определяется числом M слагаемых в дискретном выражении для $S(f)$, то при изменении частот дискретизации удобно использовать нормированные выражения, получающиеся умножением $S(f)$ на множитель $1/M = 4/N$. При такой нормировке функциональные различия в зависимостях $S(f)$, обусловленные использованием сигналов, измеренных при разных частотах дискретизации, выявляются более явно.

Таким образом, при анализе сложного хаотического сигнала при фликкер-шумовой спектроскопии

определяется совокупность параметров, характеризующих корреляционные взаимосвязи в последовательностях нерегулярностей-скачков и нерегулярностей-всплесков, свойственные данному сигналу, определяемому с частотой дискретизации f_d . Анализ зависимостей $\Phi^{(2)}(\tau)$ и $S(f)$, рассчитываемых на основе временных рядов с уменьшающей частотой дискретизации, позволяет оценить меру "устойчивости" параметров σ_1 , T_1 и H_1 , определяемых на основе $\Phi^{(2)}(\tau)$, и меру вариабельности параметров $S_s(0)$, T_0 и n_0 , относящихся к зависимости $S(f)$.

Список литературы

1. Малинецкий Г. Г. Введение в нелинейную динамику. М.: Эдиториал УРСС, 2002. 256 с.
2. Стругинский А. В., Мурашко В. В. Электрокардиография. М.: Медпресс-информ, 2004. 320 с.
3. Тимашев С. Ф. Фликкер-шумовая спектроскопия: информация в хаотических сигналах. М.: Физматлит, 2007. 248 с.
4. Тимашев С. Ф., Встовский Г. В. Фликкер-шумовая спектроскопия в анализе хаотических временных рядов динамических переменных и проблема отношения "сигнал — шум" // Электрохимия. 2003. Т. 39. С. 156—169.
5. URL: www.PhysioNet.org

N. T. Abdullayev¹, Associate Professor, e-mail: a.namik46@mail.ru,

M. M. Gasankuliyeva², Graduate Student,

¹ Azerbaijani Technical University

² Azerbaijani State Oil Academy

Informative Parameters and Fluctuation Dynamics of Electrocardiograms at Flicker-Noise Spectroscopy

On the basis of discretely fixed temporary ranks of electrocardiographic signals the analysis of spectral power and the transitional differential moments of the second order is made. Such analysis allows to receive the informative diagnostic parameters estimating a functional condition of cardiovascular system of the patient. The variation of the used sampling frequencies can be one of the factors allowing to estimate fluctuation dynamics of electrocardiograms and to allocate a contribution of irregularities to the analyzed real signals.

Keywords: flicker-noise spectroscopy, electrocardiography, differential diagnostics, informative parameters, frequency of sampling of a signal

References

1. Malineckij G. G. *Vvedenie v nelinejnuju dinamiku*, Moscow: Jeditorial URSS, 2002, 256 p.

2. Strutynskij A. V., Murashko V. V. *Jelektrokardiografija*. Moscow: Medpress-inform, 2004. 320 p.

3. Timashev S. F. *Flicker-shumovaja spektroskopija: informacija v haoticheskih signalah*, Moscow: Fizmatlit, 2007. 248 p.

4. Timashev S. F., Vstovskij G. V. Flicker-shumovaja spektroskopija v analize haoticheskih vrcmennyh rjadov dinamicheskikh perezmenenij i problema odnoshenija "signal — shum", *Jelektrohimija*, 2003, vol. 39, pp. 156–169.

5. URL: www.PhysioNet.org

ИНФОРМАЦИЯ

16–21 мая 2016 г.,

г. Ростов-на-Дону — г. Сочи

1-я Международная научная конференция

"ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ТЕХНИКЕ И НА ПРОИЗВОДСТВЕ" (ИТИ'16)

НАУЧНЫЕ НАПРАВЛЕНИЯ КОНФЕРЕНЦИИ

- Модели представления и методы извлечения знаний
- Интеллектуальный анализ данных
- Нечеткие множества, нечеткие отношения и родственные им формализмы
- Теория возможности и теория свидетельств
- Неклассические логики и правдоподобный вывод
- Моделирование НЕ-факторов в интеллектуальных системах
- Мягкие вычисления и измерения, вычислительный интеллект
- Лингвистические модели и гранулярные вычисления, вычисления со словами и перцептами
- Принятие решений и интеллектуальные системы поддержки принятия решений
- Мультимножества и нечеткие мультимножества
- Гибридные нечеткие и вероятностные модели
- Алгебраические байесовские сети и сети доверия
- Нейронные сети, нейроинформатика и нейроинтеллект
- Когнитивное моделирование и когнитивные агенты
- Классификация, распознавание и обучение
- Эволюционное моделирование, бионические и генетические алгоритмы
- Многоагентные системы и распределенный ИИ
- Искусственная жизнь, интеллект роя, адаптивное поведение
- Прикладные интеллектуальные системы динамического типа
- Автоматизация и интеллектуализация энергетических и транспортных систем
- Защита информации и информационная безопасность
- Интеллектуальные информационные технологии в образовании

КОНТАКТЫ ИТИ'16

Суханов Андрей Валерьевич (+ 7-989-720-65-53; iiti16@rgups.ru)

Самсонов Владимир Леонидович (+ 7-903-470-68-34; samsonov.sem@gmail.com)

УДК 004.934

В. В. Савченко, профессор, зав. каф., e-mail: svv@lunn.ru,
Нижегородский государственный лингвистический университет, г. Нижний Новгород

Распознавание речевых команд методом фонетического декодирования слов с подавлением фонового шума

Предложен новый, помехоустойчивый метод фонетического декодирования слов для коммерческих (недорогих и надежных) систем голосового управления робототехникой. В его основу положена идея "исправления" выборочных оценок автокорреляционных матриц гласных фонем путем вычитания из их элементов на главной диагонали оценки дисперсии шумового фона. Показано, что достигаемый эффект состоит в многократном (11,8 дБ и более) увеличении отношения сигнал/шум на входе решающего устройства.

Ключевые слова: речь, распознавание речи, речевая команда, голосовое управление, шум, помехи, борьба с помехами, защита от помех, помехозащищенность

Введение

Метод фонетического декодирования слов (МФДС) предложен в работах [1–4] для повышения точности и быстродействия автоматического распознавания речи (АРР) в системах голосового управления робототехникой. Со ссылкой на информационную теорию восприятия речи [5] данный метод изначально его авторы позиционировали как альтернативу каноническому подходу к АРР, реализованному в большой группе известных методов [6–9] на основе математического аппарата скрытых марковских моделей речевого сигнала [10]. Их общий принципиальный недостаток — использование многозатратной процедуры выравнивания слов по темпу речи диктора. В основу же МФДС положена идея выделения множества информативных, в частности гласных или вокализованных, фонем $\{x_r\}$ из фонетической базы данных (ФБД) русского языка на роль конечной группы объектов АРР с наиболее ярко выраженными различиями в теоретико-информационном смысле [4]. В данном случае алгоритм АРР сводится к фонетическому послоговому декодированию слов на сильно редуцированном множестве гласных фонем, а процедура динамического выравнивания слов утрачивает всякий смысл. Именно в этом преимуществе авторы МФДС первоначально видели его главное обоснование. Между тем, существует и еще один важный мотив, возможно, не менее актуальный для практики, а именно [3] повышение помехоустойчивости АРР с использованием редуцированного МФДС. Цель статьи — исследование эффективности метода

при действии интенсивного внешнего (фонового) шума, а также поиски путей повышения его помехозащищенности. Указанная цель направлена на преодоление актуальной научно-технической проблемы — защиты АРР от внешних помех [11] на пути к широкому распространению систем и технологий голосового управления робототехникой в быту, на транспорте и на производстве.

1. Постановка задачи

Пусть задана аддитивная модель речевого сигнала

$$x = x_r + \eta, \quad r = \overline{1, R}, \quad (1)$$

на интервале квазистационарности его текущего вокализованного отрезка длиной T_r . При частоте дискретизации сигнала F в результате будем иметь $N_r = FT_r$ отсчетов данных. Рассматривая этот сигнал в режиме неперекрывающегося "скользящего окна" длиной в n отсчетов, получаем $L_r = N_r/n$ независимых в совокупности массивов данных $x_l = x_{r,l} + \eta_l$ размерностью n каждый. Здесь $x_{r,l}$, η_l — n -векторы (столбцы) отсчетов r -й фонемы x_r и фонового шума η соответственно в дискретном времени $l = 1, 2, \dots, L_r$. Так, например, при $F = 8$ кГц, $T_r = 80 \dots 120$ мс и $n = 20$ (рекомендуемые [1–4] значения параметров МФДС) получим объем выборки многомерных наблюдений над речевым сигналом (1) $L_r = 30 \dots 50$, что представляется вполне достаточным для обеспечения высокой помехоустойчивости МФДС.

Предположим далее [3, 12, 13], что и полезный сигнал x_r , и шум η из выражения (1) подчиняются

многомерному гауссовскому закону распределения с нулевыми средними и автокорреляционными $(n \times n)$ -матрицами (АКМ) \mathbf{K}_r и $\mathbf{K}_\eta = \sigma_\eta^2 \mathbf{I}_n$ соответственно, где $n = 10...20 \ll N$, $\sigma_\eta^2 = \text{const}$ — дисперсия или средняя мощность шума, предполагаемого стационарным; \mathbf{I}_n — единичная $(n \times n)$ -матрица. В таком случае гауссовским же будет и закон распределения суммарного речевого сигнала (1) с АКМ, равной $\mathbf{K}_x = \mathbf{K}_r + \sigma_\eta^2 \mathbf{I}_n$. Задача АРР при этом сводится [13] к многоальтернативной проверке статистических гипотез:

$$W_r: \mathbf{K}_x = \mathbf{K}_r, r = \overline{1, R}, \quad (2)$$

в отношении АКМ наблюдаемой фонемы x_r . Это стандартная задача статистической радиотехники [14]. Ее оптимальное (в байесовском смысле) решение приводит [12—14] к R -канальной обработке речевого сигнала (1) с предварительной настройкой каждого канала на заданную ФБД $\{\mathbf{K}_r\}$.

В терминах общесистемного критерия минимума величины информационного рассогласования (ВИР) по Кульбаку—Лейблеру [15] запишем следующее выражение для оптимальной решающей статистики:

$$\rho_r(x) = 0,5[\text{tr}(\widehat{\mathbf{K}}_x \cdot \mathbf{K}_r^{-1}) - \ln|\widehat{\mathbf{K}}_x \cdot \mathbf{K}_r^{-1}| - n], \quad r = \overline{1, R}, \quad (3)$$

где

$$\widehat{\mathbf{K}}_x = L^{-1} \sum_{l=1}^L x_l x_l^T \quad (4)$$

— статистическая оценка АКМ речевого сигнала (1) на интервале действия гласной фонемы x_r ; $\text{tr}(\cdot)$, $|\cdot|$ — соответственно след и определитель квадратной $(n \times n)$ -матрицы, $(\cdot)^T$ — символ транспонирования векторов. При этом учтена естественная центрированность речевого сигнала $\mathbf{M}\{x_r\} = 0$ на выходе речевого тракта диктора [12]. Здесь $\mathbf{M}\{\cdot\}$ — символ математического ожидания. А при учете осуществляемой по МФДС [1—5] нормировки минимальных собственных чисел эталонных АКМ $\{\mathbf{K}_r\}$ к некоторому постоянному уровню $\sigma_0^2 = \text{const}$, перепишем выражение (3) в упрощенном виде:

$$\rho_r(x) = 0,5[\text{tr}(\widehat{\mathbf{K}}_x \cdot \mathbf{K}_r^{-1}) - n]. \quad (5)$$

Или в терминах обеляющего фильтра (ОФ) $(n - 1)$ -го порядка [4]

$$y_{r,l} = a_r^T x_l, l = \overline{1, L}, \quad (6)$$

будем иметь выражение для ВИР [5]

$$\rho_r(x) = 0,5[\sigma_r^2(x)/\sigma_0^2 - 1], r = \overline{1, R}, \quad (7)$$

где

$$\sigma_r^2(x) = L^{-1} \sum_{l=1}^L y_{r,l}^2 \quad (8)$$

— выборочная дисперсия отклика r -го ОФ (6) на речевой сигнал (1);

$$a_r = \mathbf{K}_r^{-1} e / (e^T \mathbf{K}_r^{-1} e) -$$

n -вектор его коэффициентов; $e = (1, 0, \dots, 0)^T$. Здесь $e^T \mathbf{K}_r^{-1} e = \rho_r$ — это средняя мощность речевого сигнала на интервале действия r -й фонемы. Выражения (2)—(8) в совокупности определяют общую формулировку МФДС в задаче АРР. Решение задачи (2) в пользу одной из гипотез W_v , $v \leq R$, принимается здесь по признаку достижения минимизации ВИР из (6) при равенстве $r = v$. Исследуем его помехоустойчивость.

2. Приведенное отношение сигнал/шум

По логике применения МФДС [1—4] точность АРР может быть охарактеризована в общем случае вероятностью перепутывания v -й и r -й фонем:

$$\begin{aligned} \beta_{v \rightarrow r} &\triangleq P\{\rho_r(x) < \rho_v(x) | x \in x_v\} = \\ &= P\{\sigma_r^2(x) < \sigma_v^2(x) | x \in x_v\}, v \neq r \leq R, \end{aligned}$$

которая в первом приближении сводится к следующему выражению [14]:

$$\beta_{v \rightarrow r} = [1 - \Phi_{L,L}(1 + 2\rho_{v/r})]. \quad (9)$$

Здесь обозначено

$$\begin{aligned} \rho_{v/r} &\triangleq \rho_r(x) | x \in x_v = \\ &= 0,5 \left[\frac{\mathbf{M}\{\sigma_r^2(x_v)\}}{\sigma_0^2} - 1 \right] = 0,5 \left[\frac{\sigma_{r/v}^2}{\sigma_0^2} - 1 \right] \quad (10) \end{aligned}$$

— среднее (статистическое) значение ВИР (7) между v -й и r -й фонемами в отсутствие шума; $\Phi_{L,L}(\cdot)$ — интегральная функция Φ -распределения (Фишера) с (L, L) степенями свободы. В качестве примера рассмотрим случай характерных на множестве гласных фонем [3, 4] значений параметров МФДС: $L = 30$ и $2\rho_{v/r} = 0,7...1,0$. Согласно таблицам Φ -распределения [16] в таком случае будем иметь $\beta_{v,r} = 0,07...0,03$, т. е. достаточно мало по меркам фонетического декодирования слов. Ситуация меняется при учете фонового шума из выражения (1).

В этом случае выражение (2) может быть переписано следующим образом:

$$\mathbf{K}_x = \mathbf{K}_r + \sigma_\eta^2 \mathbf{I}_n. \quad (11)$$

Тогда из выражения (10) получим

$$\rho_{v/r} = 0,5 \left[\frac{\sigma_{r/v}^2 + \sigma_{r/\eta}^2}{\sigma_0^2} - 1 \right], \quad v, r \leq R,$$

где $\sigma_{r/\eta}^2 = (a_r^\top \cdot a_r) \sigma_\eta^2 = c_r \sigma_\eta^2$ — приведенная к выходу ОФ (6) дисперсия шума. И как результат, запишем следующее равенство:

$$\beta_{v,r} = \left[1 - \Phi_{L, L} \left(\frac{1 + 2\rho_{v/r} + q_r^{-2}}{1 + q_r^{-2}} \right) \right], \quad (12)$$

где

$$q_r^2 \triangleq (\sigma_0^2 / \sigma_{r/\eta}^2) - \quad (13)$$

приведенное (к выходу ОФ) отношение сигнал/шум (ОСШ). Как видим, величина ОСШ может служить объективной характеристикой точности АРР по МФДС при действии фонового шума определенной дисперсии σ_η^2 .

В связи с введенным понятием отметим, что речевой вокализованный сигнал x_r сильно ослабляется в настроенном на него ОФ (6). Поэтому его приведенная средняя мощность резко уменьшается: $\sigma_0^2 \ll \rho_r$. И наоборот, белый шум на выходе того же ОФ в несколько раз усиливается по своей дисперсии σ_η^2 , так как применительно к гласным фонемам выполняется соотношение $(a_r^\top \cdot a_r) \gg 1$. Нетрудно понять, в таком случае, что при действии интенсивных внешних помех, когда ОСШ сокращается до уровня $q^2 \ll 1$, вероятность перепутывания любых пар фонем согласно (12) резко возрастает и в пределе достигает максимума 0,5. А это означает, что принимаемые согласно МФДС решения в пользу той или иной гипотезы в задаче (2) при действии даже не слишком интенсивного шума по существу принимаются наугад. Понятно, что для всех других методов АРР полученные оценки помехоустойчивости, по крайней мере, не улучшаются. Поэтому далее предпринята попытка всемерно ослабить указанную проблему, хотя бы, применительно к МФДС.

3. Коэффициент подавления шума

Отталкиваясь от выражения (11), преобразуем оценку АКМ из (5) к модифицированному виду

$$\hat{\mathbf{K}}_{x/\eta} = \hat{\mathbf{K}}_x - \sigma_\eta^2 \mathbf{I}_n. \quad (14)$$

По существу, этим взят на вооружение известный [17] метод подавления шума — минимального

собственного числа АКМ. Однако он реализуется в (14) с одной принципиальной оговоркой: дисперсия фонового шума σ_η^2 не является минимальным собственным числом для суммарной матрицы \mathbf{K}_x из выражения (11), как это имеет место в упомянутом здесь методе. В нашем случае ее оценка $\hat{\sigma}_\eta^2$ может быть получена, причем с достаточно высокой степенью точности, по стандартной формуле средней квадратичной величины (8) — в паузах между слогами текущей речевой команды. Напомним при этом [3, 4], что послоговое произнесение команд диктором является принципиальной особенностью МФДС в задачах голосового управления робототехникой.

При учете оценки (14) перепишем выражение (5) следующим образом:

$$\rho_{r/\eta}(x) = 0,5 [\text{tr}(\hat{\mathbf{K}}_x - \hat{\sigma}_\eta^2 \mathbf{I}_n) \mathbf{K}_r^{-1}] - n]. \quad (15)$$

Из этого выражения по аналогии с (7) получим модифицированное выражение для решающей статистики МФДС вида

$$\begin{aligned} \rho_{r/\eta}(x) &= 0,5 [(\sigma_r^2(x) + (\hat{\sigma}_{r/\eta}^2 - \sigma_{r/\eta}^2)) / \sigma_0^2 - 1] = \\ &= 0,5 [(\sigma_r^2(x) + c_r(\hat{\sigma}_\eta^2 - \sigma_\eta^2)) / \sigma_0^2 - 1] \end{aligned} \quad (16)$$

при очевидном равенстве $\sigma_\eta^2 = \mathbf{M}(\hat{\sigma}_\eta^2)$ для несмещенной оценки дисперсии. Решение здесь, как и прежде в МФДС, принимается по принципу минимизации решающей статистики. Нетрудно понять на качественном уровне, что точность такого решения существенно возросла по сравнению с первоначальным вариантом алгоритма (7)—(9) — за счет предпринятого в (14) и (15) подавления шума. Дадим количественную характеристику достигаемого выигрыша по ОСШ на входе решающего устройства. Определим этот выигрыш значением коэффициента подавления фонового шума

$$\gamma \triangleq (\hat{\sigma}_\eta^2 - \sigma_\eta^2) / \sigma_\eta^2 \quad (17)$$

в модифицированном алгоритме (16).

Следуя известной схеме вычислений [14], раскроем выражение (17) в явном виде $\gamma = L^{-1} \chi_L^2 - 1$, где χ_L^2 — случайная величина, распределенная по закону χ^2 -Пирсона с L степенями свободы [16]. Ее математическое ожидание равно объему многомерной выборки наблюдений L , а среднее квадратичное отклонение равно $\sqrt{2L}$. Квадрат их отношения $h^2 = L/2$ и определяет в конечном итоге значение достигаемого выигрыша в ОСШ при применении модифицированного МФДС (15) с подавлением шума η . Например, при $L = 30 \dots 50$ выигрыш составит 11,8...14,0 дБ. Чем больше объем многомерной выборки речевого сигнала на интервалах его вокализации гласными фонемами, тем выше эф-

фективность предложенного метода защиты АРР от шумовых помех.

Заключение

Системы голосового управления робототехникой привлекают к себе внимание исследователей на протяжении нескольких десятков лет в связи, главным образом, с почти неограниченной сферой их возможного применения: от бытовых приборов и проектов типа "умного дома" до систем специального назначения. К сожалению, их распространение на практике до настоящего времени сдерживается недостаточно высокой степенью надежности АРР в условиях действия интенсивных внешних акустических помех, таких как шум улицы, звуки от проезжающего транспорта, речь посторонних лиц и т.п.

МФДС — это один из новейших методов АРР, основанный на принципах слоговой фонетики русского языка и нацеленный на преодоление проблемы защиты АРР от помех за счет использования ограниченного числа $R < \infty$ минимальных речевых единиц — с максимальными отличиями друг от друга в информационной метрике Кульбака—Лейблера. Помехоустойчивость АРР еще более возрастает в предложенной выше модификации МФДС с автоматическим подавлением фонового шума. Показано, что достигаемый при подавлении шума эффект состоит в многократном (11,8 дБ и более) увеличении ОСШ на входе решающего устройства. На основании данного результата можно сделать вывод о значительных перспективах расширения сферы распространения голосового управления робототехникой за счет использования предложенного метода АРР.

Таким образом, благодаря проведенному исследованию предложен на основе МФДС новый, эффективный метод защиты АРР от помех, предназначенный для использования в условиях действия интенсивного внешнего шума, когда известные технологии голосового управления не отвечают требованиям к их точности и надежности. Если говорить об основных ограничениях на его оптимальность и область применения, то они связаны, прежде всего, с предположением о гауссовском распределении речевого сигнала в постановочной части статьи. Однако указанное ограничение только на первый взгляд кажется чрезмерно жестким. В работе [14] строго показано, что гауссовский закон в его адаптивном варианте теоретически строго обусловлен общесистемным принципом максимальной энтропии наблюдений для широкого круга алгоритмов, использующих моменты распределения не выше второго порядка. Отметим, что подавляющее большинство известных методов АРР относится, как известно [10], именно к такому кругу алгоритмов,

поскольку основывается на методах корреляционной и спектральной обработки речевого сигнала.

Список литературы

1. Патент № 111944 (полезная модель). Устройство для фонетического анализа и распознавания речи / В. В. Савченко, А. В. Савченко, Д. Ю. Акатьев / Роспатент: по заявке № 2011125526/08 от 21.06.2011. — Оpubл. 27.12.2011, бюл. № 6.
2. Савченко А. В., Савченко В. В. Разработка быстродействующих алгоритмов автоматического распознавания голосовых команд с регулируемой точностью и надежностью на основе принципов слоговой фонетики русского языка и метода фонетического декодирования слов в информационной метрике Кульбака—Лейблера: Тезисы докладов Всероссийской научно-технической конференции и выставки, посвященной итогам реализации федеральной целевой программы "Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007—2013 годы": М.: МИСиС, 2013. С. 184—185.
3. Савченко А. В. Адаптивный алгоритм распознавания речи на основе метода фонетического декодирования слов в задаче голосового управления // Информационные технологии. 2013. № 4. С. 34—39.
4. Савченко В. В., Савченко А. В. Метод фонетического декодирования слов в информационной метрике Кульбака—Лейблера для систем автоматического анализа и распознавания речи с повышенным быстродействием // Информационно-управляющие системы. 2013. № 2. С. 7—12.
5. Савченко В. В., Пономарев Д. А. Оптимизация фонетической базы данных по группе дикторов на основе информационной теории восприятия речи // Информационные технологии. 2009. № 12. С. 7—12.
6. Pat. US 6301560 B1, Int. CI G10L15/22. Discrete speech recognition system with ballooning active grammar / Inventor: Masters S. P. / Assignee: Microsoft Corporation. — Pub. Date 09.10.2001.
7. Schuster M. Speech Recognition for Mobile Devices at Google // Proc. of the 11th Pacific Rim international conference on Trends in Artificial Intelligence, 2010. LNCS Vol. 6230. С. 8—10.
8. Pat. US 8175883 B2, Int. CI G10L21/00 (2006.01). Speech recognition system and method / Inventor: Grant R., Gregor P. / Assignee: Nuance Communications Inc. — Pub. Date 08.05.2012.
9. Pat. US 0016678 A1, Int. CI G10L21/00 (2006.01). Intelligent automated assistant / Inventor: Gruber T., Cheyer A., Kittlaus D., Guzzoni D., Brigham C., Giuliano R., Bastea-Forte M., Sessler H. / Assignee: Apple Inc. — Pub. Date 19.01.2012.
10. Rabiner L. A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition // Proc. of the IEEE. 1989. Vol. 77 (2). P. 257—285.
11. Крашенинников В. Р., Армер А. И., Крашенинникова Н. А., Кузнецов В. В., Хвостов А. В. Некоторые задачи, связанные с распознаванием речевых команд на фоне интенсивных шумов // Инфокоммуникационные технологии. 2008. Т. 6, № 1. С. 72—75.
12. Levinson S. C. Mathematical models for speech technology. Chichester: England: John Wiley & Sons Ltd, 2005. 261 p.
13. Савченко В. В. Автоматическая обработка речи по критерию минимума информационного рассогласования на основе метода обеляющего фильтра // Радиотехника и электроника. 2005. Т. 50, № 3. С. 309—314.
14. Савченко В. В. Различение случайных сигналов в частотной области // Радиотехника и электроника. 1997. Т. 42, № 4. С. 426—431.
15. Kullback S. Information Theory and Statistics. New York: Dover Publications, 1997. 399 p.
16. Мюллер П., Нойман П., Шторм Р. Таблицы по математической статистике: Пер. с нем. М.: Финансы и статистика, 1982. 278 с.
17. Кошелев В. И., Андреев В. Г. Оценка собственных значений в задаче обработки эхо-сигналов // 100-летие начала использования электромагнитных волн для передачи сообщений и зарождения радиотехники: тезисы докладов Международной конференции, г. Москва, 4—6 мая 1995 г.: М.: Изд-во журнала "Радиотехника", 1995. С. 180—181.

The Speech Recognition Method of Phonetic Decoding of Words with Background Noise Cancellation

This paper studies the usage of the known words phonetic decoding method in the task of automatic speech recognition in robotics voice control systems. Our goal is to improve the efficiency of this method for acute problem of modern speech recognition techniques, namely, insufficient accuracy in case of the presence of intensive background noise. The achievement of this goal can cause an extensive use of voice control systems in many practical applications. Our study is based on modern theoretical computer science. In this paper we proposed the novel noise-proof version of phonetic decoding method for its application in commercial (inexpensive and reliable) intelligent systems with voice control of robotics in life, on transport and on production in the presence of intensive external noise. Our modification explores the idea of correcting the estimated covariance matrices of vowels by subtraction from their diagonal elements of the noise variance. Analytical expression of noise suppression coefficient is synthesized. It is shown that the proposed approach allows to multiple increase of the signal-to-noise-ratio (more than 11,8 dB). In conclusion we discuss the possible future expansion of speech recognition and voice control technologies on the basis of proposed method and give practical recommendations of its usage.

Keywords: speech, speech recognition, speech recognition, speech command, voice-activated control, noise, interference, fight against interference, protection against interference, noise immunity

References

1. **Pat. № 111944** (poleznaja model'). *Ustrojstvo dlja foneticheskogo analiza i raspoznavanija rechi*, V. V. Savchenko, A. V. Savchenko, D. Ju. Akat'ev, Rospatent: po zajavke № 2011125526/08 ot 21.06.2011, opubl. 27.12.2011, bjul. no. 6 (in Russian).
2. **Savchenko A. V., Savchenko V. V.** Razrabotka bystrodejstvujushih algoritmov avtomaticheskogo raspoznavanija golosovyh komand s reguliruemoj tochnost'ju i nadezhnost'ju na osnove principov slogovoj fonetiki russkogo jazyka i metoda foneticheskogo dekodirovanija slov v informacionnoj metrike Kul'baka—Lejblera: Tezisy dokladov Vserossijskoj nauchno-tehnicheskoy konferencii i vystavki, posvjashhennoj itogam realizacii federal'noj celevoj programmy "Issledovanija i razrabotki po prioritetyim napravlenijam razvitiya nauchno-tehnologicheskogo kompleksa Rossii na 2007—2013 gody", Moscow, MISiS, 2013, pp. 184—185.
3. **Savchenko A. V.** Adaptivnyj algoritm raspoznavanija rechi na osnove metoda foneticheskogo dekodirovanija slov v zadache golosovogo upravlenija, *Informacionnye tehnologii*, 2013, no. 4, pp. 34—39.
4. **Savchenko V. V., Savchenko A. V.** Metod foneticheskogo dekodirovanija slov v informacionnoj metrike Kul'baka—Lejblera dlja sistem avtomaticheskogo analiza i raspoznavanija rechi s povyshennym bystrodejstviem, *Informacionno-upravljajushhie sistemy*, 2013, no. 2, pp. 7—12.
5. **Savchenko V. V., Ponomarjov D. A.** Optimizacija foneticheskoy bazy dannyh po gruppe diktov na osnove informacionnoj teorii vosprijatija rechi, *Informacionnye tehnologii*, 2009, no. 12, pp. 7—12.
6. **Pat. US 6301560 B1**, Int. CI G10L15/22. *Discrete speech recognition system with ballooning active grammar* / Inventor: Masters S. P. / Assignee: Microsoft Corporation. — Pub. Date 09.10.2001.
7. **Schuster M.** *Speech Recognition for Mobile Devices at Google*, *Proc. of the 11th Pacific Rim international conference on Trends in Artificial Intelligence*, 2010, LNCS, vol. 6230, pp. 8—10.
8. **Pat. US 8175883 B2**, Int. CI G10L21/00 (2006.01), *Speech recognition system and method* / Inventor: Grant R., Gregor P., Assignee: Nuance Communications Inc., pub. Date 08.05.2012.
9. **Pat. US 0016678 A1**, Int. CI G10L21/00 (2006.01), *Intelligent automated assistant* / Inventor: Gruber T., Cheyer A., Kittlaus D., Guzzoni D., Brigham C., Giuli R., Bastea-Forte M., Saddler H. Assignee: Apple Inc., pub. Date 19.01.2012.
10. **Rabiner L.** A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, *Proc. of the IEEE*, 1989, vol. 77 (2), pp. 257—285.
11. **Krashennikov V. R., Armer A. I., Krashennikova N. A., Kuznecov V. V., Hvostov A. V.** Nekotorye zadachi, svjazannye s raspoznavaniem rechevnyh komand na fone intensivnyh shumov, *Infokommunikacionnye tehnologii*, 2008, vol. 6, no. 1, pp. 72—75.
12. **Levinson S. C.** *Mathematical models for speech technology*, Chichester, England: John Wiley & Sons Ltd., 2005, 261 p.
13. **Savchenko V. V.** Avtomaticheskaja obrabotka rechi po kriteriju minimuma informacionnogo rassoglasovanija na osnove nietoda obelajushhego fil'tra, *Radiotekhnika i jelektronika*, 2005, vol. 50, no. 3, pp. 309—314.
14. **Savchenko V. V.** Razliczenie sluchajnyh signalov v chastotnoj oblasti, *Radiotekhnika i jelektronika*, 1997, vol. 42, no. 4, pp. 426—431.
15. **Kullback S.** *Information Theory and statistics*, Dover Publications, New York, 1997, 399 p.
16. **Mjuller P., Nojman P., Shtorm R.** *Tablicy po matematicheskoj statistike*, Per. s nem., Moscow, Finansy i statistika, 1982, 278 p.
17. **Koshelev V. I., Andreev V. G.** Ocenka sobstvennyh znachenij v zadache obrabotki jeh-signalov, *100-letie nachala ispol'zovanija jelektronnym voln dlja peredachi soobshhenij i zarozhdenija radiotekhniki: tezisy dokladov Mezhdunarodnoj konferencii*, g. Moscow, 4—6 maja, 1995, Moscow, Izd-vo zhurnala "Radiotekhnika", 1995, pp. 180—181.

Адрес редакции:

107076, Москва, Стромьинский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Технический редактор *Е. В. Конова*.

Корректор *Е. В. Комиссарова*.

Сдано в набор 09.11.2015. Подписано в печать 22.12.2015. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ IT116. Цена договорная.

Журнал зарегистрирован в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций.

Свидетельство о регистрации ПИ № 77-15565 от 02 июня 2003 г.

Оригинал-макет ООО "Авансед солюшнз". Отпечатано в ООО "Авансед солюшнз".

119071, г. Москва, Ленинский пр-т, д. 19, стр. 1.