

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Том 22
2016
№ 11

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

СОДЕРЖАНИЕ

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ

Левин В. И. Непрерывно-логическая модель решения комбинаторных задач . . . 803

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ И СЕТИ

Дворников А. А., Восков Л. С., Саксонов Е. А., Ефремов С. Г. Метод построения оптимального наложенного канала для беспроводной сенсорной сети . 812

СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ

Яковлев А. А., Камаев В. А., Сорокин В. С., Мишустина С. Н. Поисквое конструирование систем охлаждения на основе инженерно-физического подхода 819

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ИЗОБРАЖЕНИЙ

Бимаков Е. В. Метод виртуального зондирования сцены на базе воксельной технологии обработки сцен 827

ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ

Беляков С. Л., Белякова М. Л., Глушков А. А. Метапреобразования образов при поиске достоверных решений в интеллектуальных геоинформационных системах 836

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ

Машкина И. В., Сенцова А. Ю. Обеспечение информационной безопасности системы облачных вычислений 843

Щеглов К. А., Щеглов А. Ю. Вопросы корректности и универсальности подхода к моделированию надежностных параметров и характеристик угроз атак . . 854

БАЗЫ ДАННЫХ

Альгулиев Р. М., Алекперова И. Я. Проблема BIG DATA в нефтегазовой отрасли: состояние и перспективы. 862

Журнал в журнале

НЕЙРОСЕТЕВЫЕ ТЕХНОЛОГИИ

Пекунов В. В. Компактное описание вариативных полей физических величин в повторяющихся задачах моделирования атмосферного переноса загрязнителей 871

Кургалин С. Д., Туровский Я. А., Борзунов С. В., Адаменко А. А. Теоретические аспекты оптимизации эволюционного обучения нейронных сетей с использованием "изолятов" 875

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.
ЛЬВОВИЧ Я. Е., д. т. н., проф.
МАРТЫНОВ В. В., 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.
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

MODELING AND OPTIMIZATION

Levin V. I. Continuous-Logical Model for Solution of Combinatorial Problems . . . 803

COMPUTING SYSTEMS AND NETWORKS

Dvornikov A. A., Voskov L. S., Saksonov E. A., Efremov S. G. A Method for Organization of an Optimal Overlay Channel in Wireless Sensor Networks 812

CAD-SYSTEMS

Yakovlev A. A., Kamaev V. A., Sorokin V. S., Mishustina S. N. Search Engine Design Systems Cooling Based Engineering and Physical Approach 819

DIGITAL PROCESSING OF SIGNALS AND IMAGES

Bimakov E. V. The Method of Virtual Probing of Scene Based on the Voxel Technology of Scene Processing 827

GEOINFORMATION SYSTEMS

Belyakov S. L., Belyakova M. L., Glushkov A. A. Images Meta-Transformation when Searching for Reliable Solutions in Intelligent Geographic Information Systems . . 836

CRYPTOSAFETY INFORMATION

Mashkina I. V., Sentsova A. Yu. Information Security of Cloud Computing System . 843

Shcheglov K. A., Shcheglov A. Yu. Correctness and Versatility Problems of Modeling Attack Threats Reliability Parameters and Characteristics Approach 854

DATABASE

Aliguliyev R. M., Alakperova I. Ya. BIG DATA Problem in Oil and Gas Industry: Current State and Prospects 862

Journal-in-journal

NEUROTECHNOLOGIES

Pekunov V. V. The Compact Description of Variative Fields of Physical Variables in the Repetitive Tasks of Simulation of Atmospheric Pollution's Transfer 871

Kurgalin S. D., Turovsky Ya. A., Borzunov S. V., Adamenko A. A. Theoretical Aspects of Optimization of Evolutionary Learning of Neurochips Using "Isolates" 875

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.
Kurechik V. M., Dr. Sci. (Tech.), Prof.
Ljovitch Ya. E., Dr. Sci. (Tech.), Prof.
Martynov V. V., 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.
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"

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ MODELING AND OPTIMIZATION

УДК 519.715

В. И. Левин, д-р техн. наук, проф., e-mail: vilevin@mail.ru,
Пензенская государственная технологическая академия

Непрерывно-логическая модель решения комбинаторных задач

Сформулирован класс комбинаторных задач, эквивалентных задаче определения взаиморасположения n последовательностей интервалов. Показано, что адекватной математической моделью решения поставленной задачи является конечный динамический автомат без памяти, адекватным математическим аппаратом — непрерывная логика. Построены алгоритмы решения. Приведен пример.

Ключевые слова: непрерывная логика, комбинаторная задача, автоматная модель, интервал

Введение

Многие задачи обработки данных, планирования работ, проектирования систем управления и т. д. сводятся математически к решению комбинаторной задачи — определению взаиморасположения n последовательностей интервалов (временных или пространственных) и нахождению условий, при которых это взаиморасположение имеет тот или иной качественный характер. Приведем несколько примеров таких задач.

1. Пусть имеются последовательность A временных интервалов, в которых некоторое основное техническое устройство работоспособно, и последовательность B временных интервалов, в которых работоспособно резервное устройство. Система "основное и резервное устройства" работоспособна, если работоспособно хотя бы одно из двух ее устройств — основное или резервное. Таким образом, для того чтобы установить последовательность интервалов работоспособности всей системы, надо определить взаиморасположение последовательностей интервалов работоспособности основного устройства (последовательность A) и резервного устройства (последовательность B) и найти те промежутки времени, на которых действуют интервалы хотя бы одной из последовательностей — A или B ; они и будут интервалами работоспособности системы. А для того чтобы, например, установить, когда система работоспособна на произвольном заданном отрезке времени T , надо найти условия, при которых расположение последовательностей интервалов A и B таково, что на отрезке T интервалы последовательности B накрывают промежутки между интервалами последовательности A .

2. Рассмотрим последовательность A временных интервалов, в которых некоторая организация (ма-

газин, банк, ремонтная мастерская и т.д.) выполняет обслуживание клиентов, и последовательность B временных интервалов, в которых некоторый клиент может посетить обслуживающую организацию. Чтобы установить последовательность промежутков времени возможного обслуживания клиента, надо определить взаиморасположение последовательностей интервалов A и B и выявить те промежутки времени, где действуют интервалы обеих последовательностей — A и B ; они и будут промежутками времени возможного обслуживания клиента. А для того чтобы установить, к примеру, когда организация может обслужить клиента при его обращении в любой доступный для него момент времени, надо найти условия, при которых расположение последовательностей интервалов A , B такое, что интервалы последовательности A накрывают все интервалы последовательности B .

3. Пусть заданы последовательность A временных интервалов, в которых председатель совета может провести его заседание, и последовательности B_1, \dots, B_{10} временных интервалов, в которых члены совета 1, ..., 10 могут участвовать в этом заседании. Заседание совета возможно только при участии в нем председателя и не менее 5 любых членов совета. Для того чтобы установить последовательность промежутков времени, в которых возможно проведение заседания совета, надо определить взаиморасположение последовательностей интервалов A , B_1, \dots, B_{10} и найти те промежутки времени, на которых одновременно действуют интервалы последовательности A и интервалы каких-то 5 или более из 10 последовательностей B_1, \dots, B_{10} ; они и будут промежутками времени возможного проведения заседания совета. А для того чтобы установить, например, когда проведение заседания совета оказывается возможным на

произвольном заданном отрезке времени T , надо найти условия, при которых взаиморасположение последовательностей интервалов A, B_1, \dots, B_{10} такое, что отрезок T накрывается каким-либо интервалом последовательности A и какими-либо 5 или более интервалами из 5 или более соответствующих последовательностей B_1, \dots, B_{10} .

Задача определения взаиморасположения n последовательностей интервалов, примеры которой приведены выше, является комбинаторной задачей. При решении комбинаторных задач данного типа можно применять различные переборные методы. Но большие недостатки этих методов (быстрый рост сложности вычислений при увеличении размеров задачи; неаналитический (поисковый) характер алгоритма решения; невозможность обозримого представления алгоритма решения в случае высокой размерности задачи) заставляют искать другие пути решения данной задачи. Возможный подход состоит в том, чтобы найти для задачи готовую адекватную математическую модель, которая глубоко и детально разработана на базе какого-нибудь удобного аналитического аппарата. В данной статье показано, что такой моделью может служить динамический конечный автомат [1–4], а таким аппаратом — непрерывная (бесконечнозначная) логика [1, 2, 5, 6].

1. Постановка задачи

Математическая постановка решаемой задачи может быть описана следующим образом. Задано n конечных последовательностей непересекающихся интервалов

$$\begin{aligned} A_1 &= (a_{11}, b_{11}), (a_{12}, b_{12}), \dots, (a_{1m_1}, b_{1m_1}); \\ A_2 &= (a_{21}, b_{21}), (a_{22}, b_{22}), \dots, (a_{2m_2}, b_{2m_2}); \\ &\dots\dots\dots \\ A_n &= (a_{n1}, b_{n1}), (a_{n2}, b_{n2}), \dots, (a_{nm_n}, b_{nm_n}). \end{aligned} \quad (1)$$

Требуется: 1) определить взаиморасположение имеющейся системы последовательностей интервалов (1); 2) найти условия, при которых это взаиморасположение имеет тот или иной качественный характер. Сформулированные две задачи различаются тем, что первая из них нацелена на то, чтобы по заданному положению всех интервалов всех последовательностей (1) определить взаиморасположение любых комбинаций (по две, по три и т. д.) любых подпоследовательностей последовательностей (1), в то время как вторая имеет целью найти условия, накладываемые на положение всех интервалов всех подпоследовательностей (1), при которых указанное взаиморасположение имеет тот или иной требуемый вид. Таким образом, задача 1 является задачей анализа имеющейся системы последовательностей интервалов (1), задача 2 — задачей синтеза такой системы. Задачи анализа и синтеза системы последовательностей интервалов

вида (1) мы будем решать, используя адекватную системе (1) математическую модель конечного динамического автомата без памяти и математический аппарат непрерывной логики, необходимый для адекватного описания автомата. Решение предлагается получать в аналитической форме суперпозиции непрерывно-логических операций, которая одновременно дает логический алгоритм получения численного решения.

Заметим, что аппарат непрерывной логики перекликается с аппаратом нечеткой логики, введенным в свое время для описания операций над нечеткими множествами [7], когда каждой теоретико-множественной операции устанавливается во взаимно однозначное соответствие некоторая логическая операция (объединению — дизъюнкция, пересечению — конъюнкция, дополнению — отрицанию) над мерами принадлежности элемента множествам-операндам. Поскольку мера принадлежности элемента множеству всегда заключена в отрезке $[0, 1]$, все операции нечеткой логики определяются также на этом отрезке. Для непрерывной логики такое ограничение не требуется, поэтому непрерывно-логические операции определяются на произвольном отрезке $[A, B]$, где A и B — любые вещественные числа, удовлетворяющие условию $A < B$. Это свободное определение носителя алгебры непрерывной логики имеет решающее прикладное значение, в частности, для теории динамических автоматов, где динамические процессы рассматриваются на отрезке времени $[0, \infty)$. Именно такая форма непрерывной логики позволила построить на ее базе аналитическую теорию динамических автоматов [1–4], которая стала адекватной математической моделью для решения поставленной выше задачи.

Еще один кажущийся пригодным подход к решению поставленной задачи — это использование временных логик [8, 9]. Всякая временная логика базируется на основных отношениях между событиями во времени (пересекаться, примыкать, быть позже и т.д.), дополненных в случае метрических шкал информацией о длительности событий и их положении на шкале. Задача временной логики — анализ текстов на естественном языке для установления временных отношений между отдельными событиями, описанными в тексте. При этом часть отношений устанавливается путем непосредственного смыслового анализа текста, а другая часть — путем логического вывода из уже установленных первичных отношений с помощью эвристически формулируемых правил такого вывода. Результатом анализа являются сугубо качественные и вдобавок неточные (из-за использования эвристики) временные отношения между событиями типа "событие A произошло значительно раньше события B " либо временная идентификация событий типа "событие A произошло очень давно". Но сформулированная выше задача анализа имеющейся

системы последовательностей интервалов (1) является строго количественной задачей, предполагающей точное математическое (аналитическое) решение. Поэтому применять для ее решения временные логики невозможно. Что же касается второй сформулированной выше задачи — синтеза системы последовательностей интервалов (1) с нужными временными свойствами, то эта задача, хотя бы в качественной и неточной постановке, в рамках временных логик вообще не рассматривается.

2. Динамические конечные автоматы и непрерывная логика

Динамический конечный автомат (ДА) без памяти [1, 2, 4] представляет собой математическую модель в виде $(n, 1)$ -полосника (рис. 1), реализующего на выходе y булеву логическую функцию своих входов x_1, \dots, x_n :

$$y = f(x_1, \dots, x_n), \quad x_1, \dots, x_n, y \in \{0, 1\}. \quad (2)$$

На входы данного автомата подаются входные двоичные динамические процессы (рис. 1)

$$\begin{aligned} x_1(t) &= 1(a_{11}, b_{11})0(-, -)1(a_{12}, b_{12}) \dots 1(a_{1m_1}, b_{1m_1}); \\ &\text{-----} \\ x_n(t) &= 1(a_{n1}, b_{n1})0(-, -)1(a_{n2}, b_{n2}) \dots 1(a_{nm_n}, b_{nm_n}), \end{aligned} \quad (3)$$

в которых $1(a, b)$ означают интервалы времени единичных значений процесса (импульсы), а $0(-, -)$ — промежуточные интервалы времени нулевых значений процесса (паузы). С выхода ДА (рис. 1) снимается выходной двоичный динамический процесс

$$y(t) = 1(c_1, d_1)0(-, -)1(c_2, d_2) \dots 1(c_m, d_m), \quad (4)$$

соответствующий поданным входным процессам (3) ДА и его реализуемой булевой логической функции (2). Основной задачей для ДА без памяти является задача отыскания выходного динамического процесса $y(t)$ по его известным входным динамическим процессам $x_1(t), \dots, x_n(t)$ и реализуемой логической функции f . В 1971—1972 гг. автором было установлено, что эта задача может быть решена в аналитической форме для любого ДА без памяти, имеющего любое число входов, входные процессы и реализуемую функцию, с помощью математического аппарата непрерывной (бесконечнозначной) логики (НЛ) [1, 2, 4, 5]. Определяется НЛ

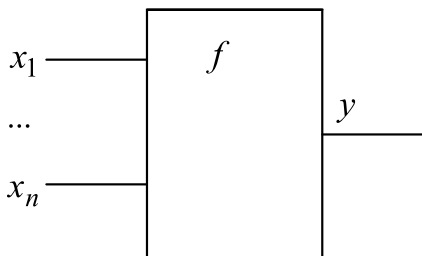


Рис. 1. Математическая модель конечного ДА

следующим образом. Пусть несущее множество $C = [A, B]$ — произвольный отрезок на оси вещественных чисел. Тогда для любых чисел $a, b, e \in C$ можно ввести следующие логические операции:

$$a \vee b = \max(a, b) \text{ — дизъюнкция,} \quad (5)$$

$$a \wedge b = \min(a, b) \text{ — конъюнкция,} \quad (6)$$

$$\bar{e} = A + B - e \text{ — отрицание.} \quad (7)$$

Операции непрерывной логики (5)—(7) подобны соответствующим операциям двужанной (булевой) логики (где несущее множество $C = \{0, 1\}$) и обобщают их на случай непрерывного несущего множества. В НЛ сохраняют силу некоторые законы двужанной логики:

$$a \vee a = a, a \wedge a = a \text{ — тавтологии;} \quad (8)$$

$$a \vee b = b \vee a, a \wedge b = b \wedge a \text{ — переместительный;} \quad (9)$$

$$(a \vee b) \vee c = a \vee (b \vee c), (a \wedge b) \wedge c = a \wedge (b \wedge c) \text{ — сочетательный;} \quad (10)$$

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c), a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c) \text{ — распределительный;} \quad (11)$$

$$\overline{a \vee b} = \bar{a} \wedge \bar{b}, \overline{a \wedge b} = \bar{a} \vee \bar{b} \text{ — де Моргана;} \quad (12)$$

$$a \vee (a \wedge b) = a, a \wedge (a \vee b) = a \text{ — поглощения.} \quad (13)$$

Кроме них, в НЛ действуют некоторые важные специфические законы, например, законы оценки и упрощения логического выражения:

$$a \vee b \geq a, b; a \wedge b \leq a, b, \quad (14)$$

$$\begin{aligned} a_1 \vee \dots \vee a_{i-1} \vee a_i \vee a_{i+1} \vee \dots \vee a_m &= \\ = a_1 \vee \dots \vee a_{i-1} \vee a_{i+1} \vee \dots \vee a_m & \\ \text{при } a_i \leq a_k (k \neq i), & \end{aligned} \quad (15)$$

$$\begin{aligned} a_1 \wedge \dots \wedge a_{i-1} \wedge a_i \wedge a_{i+1} \wedge \dots \wedge a_m &= \\ = a_1 \wedge \dots \wedge a_{i-1} \wedge a_{i+1} \wedge \dots \wedge a_m & \\ \text{при } a_i \geq a_k (k \neq i). & \end{aligned} \quad (16)$$

Идея отыскания выходного процесса ДА без памяти по его заданным входным процессам и реализуемой логической функции проста и изложена ниже. Обозначим:

1 — двоичный динамический процесс, принимающий постоянное значение 1;

0 — двоичный динамический процесс, принимающий постоянное значение 0;

1' — изменение значения двоичного динамического процесса $0 \rightarrow 1$;

0' — изменение значения двоичного динамического процесса $1 \rightarrow 0$;

1'_a — изменение 1' в момент времени a ;

0'_b — изменение 0' в момент времени b ;

1'_a 0'_b — импульс $1(a, b)$ в интервале времени (a, b) ;

0'_a 1'_b — пауза $0(a, b)$ в интервале времени (a, b) .

Любой двоичный динамический процесс можно выписать в виде последовательности импульсов и пауз (как в (3)) либо последовательности изме-

нений значения процесса. Например, процесс на рис. 2 можно записать в виде

$$x(t) = 1(-\infty, a)0(-, -)1(b, e)0(-, \infty)$$

или $x(t) = 0'_a 1'_b 0'_e$.

Число изменений значения двоичного процесса называется его глубиной, например, глубина процесса на рис. 2 равна 3. Для системы (вектора) нескольких двоичных процессов соответствующим понятием является векторная глубина. Например, векторная глубина системы из двух процессов вида $x_1(t) = 1(a, b)$, $x_2(t) = 1(c, d)$ равна (2,2).

Покажем на примерах, как с помощью НЛ определить выходной процесс ДА без памяти по его входным процессам и реализуемой логической функции. Ограничимся простейшими ДА — двух-входовыми дизъюнктами и конъюнктами, реализующими булевы логические функции дизъюнкции \vee и конъюнкции \wedge :

$$y = x_1 \vee x_2 = \begin{cases} 0 & \text{при } x_1 = x_2 = 0; \\ 1, & \text{в иных случаях;} \end{cases} \quad (17)$$

$$y = x_1 \wedge x_2 = \begin{cases} 1 & \text{при } x_1 = x_2 = 1; \\ 0, & \text{в иных случаях,} \end{cases}$$

и простейшими входными процессами с глубиной не выше 1. Пусть надо найти выходной процесс конъюнктора на входные процессы $x_1(t) = 1'_a$, $x_2(t) = 0'_b$. Очевидно, искомым процесс равен одиночному импульсу $1(a, b)$ или тождественному 0 в зависимости от того, что больше: b или a . Поэтому, интерпретируя тождественный нуль как одиночный импульс с совмещенными началом и концом, можем записать искомым процесс в виде

$$y(t) = x_1(t) \wedge x_2(t) = 1'_a \wedge 0'_b = \begin{cases} 1(a, b) & \text{при } b > a; \\ 0 = 1(a, a) & \text{при } b \leq a. \end{cases} \quad (18)$$

Отсюда с помощью операции дизъюнкции НЛ \vee окончательно находим

$$1'_a \wedge 0'_b = 1(a, a \vee b). \quad (19)$$

Выходные процессы в дизъюнкторе и конъюнкторе при всех остальных входных процессах с глубиной не выше 1 получаются аналогично:

$$0 \wedge 0'_a = 0 \wedge 1'_b = 0; \quad 1 \wedge x'_a = x'_a; \quad 0'_a \wedge 0'_b = 0'_{a \wedge b};$$

$$1'_a \wedge 1'_b = 1'_{a \vee b}; \quad 1'_a \wedge 0'_b = 1(a, a \vee b);$$

$$1 \vee 0'_a = 1 \vee 1'_b = 1; \quad 0 \vee x'_a = x'_a; \quad 0'_a \vee 0'_b = 0'_{a \vee b};$$

$$1'_a \vee 1'_b = 1'_{a \wedge b}; \quad 1'_a \vee 0'_b = 0(b, a \vee b). \quad (20)$$

Формулы (20) наглядно демонстрируют удобство и адекватность аппарата НЛ как средства отыскания выходных процессов ДА без памяти по их входным процессам и реализуемой логической функции.

Если входные процессы дизъюнктора или конъюнктора имеют кратность выше 1, то отыскание

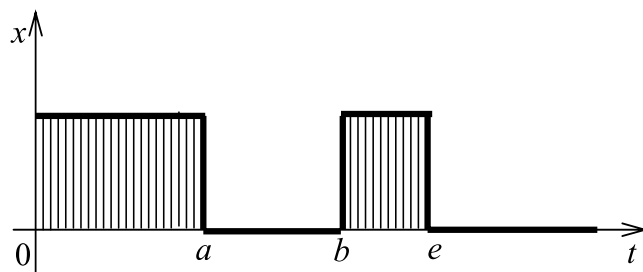


Рис. 2. Выходной процесс ДА без памяти

их выходных процессов требует применения формальных методов. Основные из них — прямой метод, метод декомпозиции и метод инверсии. *Прямой метод* основан на полном переборе всех возможных случаев взаимного расположения входных процессов элемента. Для каждого случая выходной процесс элемента записывается в явном виде отдельно. Общее выражение этого процесса получается из частных с использованием непрерывной логики. *Метод декомпозиции* заключается в том, что один из двух входных процессов элемента $x_1(t)$, $x_2(t)$, например $x_1(t)$, разбивается на два последовательных подпроцесса $x_{11}(t)$ и $x_{12}(t)$. Затем находятся составляющие выходные процессы $y_1(t)$ и $y_2(t)$ — реакции элемента на составляющие входные процессы $\{x_{11}(t), x_2(t)\}$ и $\{x_{12}(t), x_2(t)\}$. Если $y_1(t)$ и $y_2(t)$ не пересекаются во времени участками, содержащими все изменения значения процесса, то процесс $y(t)$ определяем как последовательность процессов $y_1(t)$ и $y_2(t)$. *Метод инверсии* основан на формулах

$$\overline{x_1(t) \vee x_2(t)} = \overline{x_1(t)} \wedge \overline{x_2(t)},$$

$$\overline{x_1(t) \wedge x_2(t)} = \overline{x_1(t)} \vee \overline{x_2(t)}, \quad (21)$$

вытекающих из закона де Моргана двузначной (булевой) логики и позволяющих по уже известной реакции дизъюнктора (конъюнктора) на входные процессы $x_1(t)$, $x_2(t)$ легко определить реакцию конъюнктора (дизъюнктора) на входные процессы $\overline{x_1(t)}$, $\overline{x_2(t)}$. Используя эти методы, легко получить формулы для выходных процессов дизъюнктора и конъюнктора при различных входных процессах с глубиной (1,2):

$$0'_a \vee 1(b, c) = 0(a, a \vee b)1(-, a \vee c);$$

$$1'_a \vee 1(b, c) = 1(a \wedge b, c)0(-, a \vee c);$$

$$0'_a \vee 0(b, c) = 0(a \vee b, a \vee c);$$

$$1'_a \vee 0(b, c) = 0(a \wedge b, a \wedge c);$$

$$0'_a \wedge 1(b, c) = 1(a \wedge b, a \wedge c);$$

$$1'_a \wedge 1(b, c) = 1(a \vee b, a \vee c);$$

$$0'_a \wedge 0(b, c) = 0(a \wedge b, c)1(-, a \vee c);$$

$$1'_a \wedge 0(b, c) = 1(a, a \vee b)0(-, a \vee c); \quad (22)$$

входных процессах с глубиной (2,2):

$$\begin{aligned}
 &1(a, b) \vee 1(c, d) = \\
 &= 1[a \wedge c, (a \wedge d) \vee (b \wedge c)]0(-, -)1(a \vee c, b \vee d); \\
 &1(a, b) \wedge 1(c, d) = 1[a \vee c, a \vee c \vee (b \wedge d)]; \\
 &0(a, b) \vee 0(c, d) = 0[(a \wedge d) \vee (b \wedge c), b \wedge d]; \\
 &0(a, b) \wedge 0(c, d) = \\
 &= 0[a \wedge c, (a \wedge d) \vee (b \wedge c)]1(-, a \vee c)0(-, b \vee d); \quad (23) \\
 &0(a, b) \vee 1(c, d) = \\
 &= 0(a \wedge c, b \wedge c)1(-, a \vee d)0(-, b \vee d); \\
 &0(a, b) \wedge 1(c, d) = \\
 &= 1(a \wedge c, a \wedge d)0(-, b \vee c)1(-, b \vee d)
 \end{aligned}$$

и т.д. Аналогично находятся НЛ-выражения выходных процессов многовыходовых дизъюнктов и конъюнктов, реализующих многоместные булевы дизъюнкцию и конъюнкцию, аналогичные их двуместным прототипам (17), например:

$$\begin{aligned}
 &0'_a \wedge 0'_b \wedge \dots \wedge 0'_d = 0'_{a \wedge b \wedge \dots \wedge d}; \\
 &1'_a \wedge 1'_b \wedge \dots \wedge 1'_d = 1'_{a \vee b \vee \dots \vee d}; \\
 &1'_a \wedge 1'_b \wedge \dots \wedge 1'_d \wedge 0'_e \wedge 0'_g \wedge \dots \wedge 0'_f = \\
 &= 1[a \vee b \vee \dots \vee d; \\
 &a \vee b \vee \dots \vee d \vee (e \wedge g \wedge \dots \wedge f)]; \quad (24) \\
 &0'_a \vee 0'_b \vee \dots \vee 0'_d = 0'_{a \vee b \vee \dots \vee d}; \\
 &1'_a \vee 1'_b \vee \dots \vee 1'_d = 1'_{a \wedge b \wedge \dots \wedge d}; \\
 &1'_a \vee 1'_b \vee \dots \vee 1'_d \vee 0'_e \vee 0'_g \vee \dots \vee 0'_f = \\
 &= 0[e \vee g \vee \dots \vee f; \\
 &e \vee g \vee \dots \vee f \vee (a \wedge b \wedge \dots \wedge d)].
 \end{aligned}$$

3. Идея и метод решения

Пусть интервалы в системе последовательностей интервалов (1) интерпретируются как временные интервалы. Тогда системе (1) можно поставить во взаимно однозначное соответствие совокупность двоичных динамических процессов $x_i(t)$ вида (3). Именно i -й процесс этой совокупности соответствует i -й последовательности системы ($i = \overline{1, n}$), причем k -й импульс процесса соответствует k -му интервалу данной последовательности ($k = \overline{1, m_i}$). Иными словами, двоичная переменная x_i ($i = \overline{1, n}$) является индикатором присутствия какого-то интервала i -й последовательности интервалов (1): $x_i = 1$ означает присутствие, а $x_i = 0$ — отсутствие интервала.

Подадим определенную описанным образом совокупность двоичных динамических процессов $x_i(t)$ (3) на входы x_1, \dots, x_n ДА без памяти (см. рис. 1), который реализует некоторую выбранную нами булеву логическую функцию входов $y = f(x_1, \dots, x_n)$ вида (2). Тогда ДА выдаст на выходе y некоторый двоичный динамический процесс $y(t)$ вида (4). Что характеризует этот процесс? Как известно из теории, любая булева логическая функция определяется заданным множеством единичных наборов

значений аргументов, на которых функция принимает значение 1. Таким образом, выбранная нами определенная булева логическая функция f , предназначенная для реализации в ДА без памяти, заставляет этот ДА вырабатывать на выходе определенный двоичный динамический процесс $y(t)$ вида (4), импульсы которого соответствуют тем интервалам времени, где входные процессы $x_i(t)$ (3) ДА принимают набор значений, совпадающий с одним из единичных наборов функции f . Последнее означает, что при подаче на входы ДА без памяти совокупности двоичных процессов $x_i(t)$ (3), взаимно однозначно соответствующих системе последовательностей интервалов (1), выбор для реализации в этом ДА некоторой булевой логической функции $y = f(x_1, \dots, x_n)$ означает выбор соответствующего частного показателя взаиморасположения системы последовательностей интервалов (1), а реализуемый на выходе автомата двоичный процесс $y(t)$ (4) — числовое значение этого показателя. Например, если в качестве функции f выбрана многоместная булева конъюнкция, это означает выбор частного показателя взаиморасположения системы последовательностей интервалов (1) в виде выделения всех случаев, когда интервалы всех n последовательностей (1) пересекаются (поскольку у этой функции есть только один единичный набор (1, 1, ..., 1)). При этом значение данного показателя имеет вид двоичного процесса $y(t)$ (4), импульсы которого соответствуют отрезкам времени, где интервалы всех последовательностей (1) пересекаются.

Итак, в качестве адекватной математической модели для решения задачи анализа системы последовательностей интервалов вида (1) можно выбрать ДА без памяти (см. рис. 1). Входными двоичными динамическими процессами этого динамического автомата является совокупность процессов вида (3), взаимно однозначно соответствующая системе последовательностей интервалов (1), т.е. совокупность процессов, моделирующая эту систему. ДА реализует некоторую, выбранную нами булеву логическую функцию $y = f(x_1, \dots, x_n)$, являющуюся некоторым частным показателем взаиморасположения системы последовательностей интервалов. Тогда на выходе автомата вырабатывается двоичный динамический процесс $y(t)$ (4), дающий числовое значение выбранного частного показателя взаиморасположения интервалов (более точно, выделяющий отрезки времени, где интервалы системы (1) находятся в данном взаиморасположении). То есть выходной процесс (4) ДА-модели на рис. 1 моделирует числовое значение того или иного частного показателя взаиморасположения системы последовательностей интервалов (1), соответствующего выбранной булевой логической функции f , реализуемой ДА.

Алгоритм решения задачи анализа имеющейся системы последовательностей интервалов (1) в соответствии с изложенной идеей приведен ниже.

Шаг 1. Выбирается некоторый частный показатель Π , характеризующий взаиморасположение интервалов системы (1) (если показатель Π уже задан условиями задачи, шаг 1 опускается).

Шаг 2. Строится булева логическая функция $y = f(x_1, \dots, x_n)$, соответствующая показателю Π .

Шаг 3. Строится математическая модель задачи — схема ДА без памяти, реализующая функцию f (см. рис. 1) и получающая на всех n своих входах двоичные динамические процессы $x_i(t)$ (3). Совокупность этих процессов взаимно однозначно соответствует данной системе последовательностей интервалов (1). Выходной двоичный процесс ДА $y(t)$ (4) моделирует показатель взаиморасположения системы интервалов (1) (выделяет периоды, где интервалы находятся в данном взаиморасположении).

Шаг 4. Методами теории динамических автоматов [1–4] по входным процессам ДА-модели $x_1(t), \dots, x_n(t)$ и его реализуемой функции f находится его выходной процесс $y(t)$ (4). Параметры (моменты изменения значений) этого процесса выражаются через аналогичные параметры входных процессов ДА-модели в аналитической форме с помощью операций дизъюнкции \vee и конъюнкции \wedge НЛ.

Шаг 5. Развернув найденные на шаге 4 аналитические выражения параметров выходного процесса ДА-модели $y(t)$, получаем алгоритмы вычисления этих параметров в терминах операций НЛ \vee и \wedge .

Шаг 6. Вычислив параметры выходного процесса ДА-модели $y(t)$ по алгоритмам, найденным на шаге 5, получаем числовое значение процесса, являющееся числовым значением выбранного частного показателя Π (или f) взаиморасположения системы последовательностей интервалов.

Конец алгоритма.

Заметим, что в общем случае решение задачи анализа имеющейся системы последовательностей интервалов (1) может потребовать использования не одного, а нескольких частных показателей взаиморасположения интервалов системы (1). В этом случае получается общая задача анализа, распадающаяся на несколько частных задач, соответствующих указанным частным показателям. Для решения общей задачи анализа следует решить с помощью описанного алгоритма все частные задачи и объединить полученные решения.

Алгоритм решения задачи синтеза системы последовательностей интервалов (1), соответствующей заданным требованиям к взаиморасположению интервалов, строится с использованием описанного выше алгоритма решения задачи анализа системы (1). Он состоит из следующих шагов.

Шаг 1. Решение частичной задачи анализа (с помощью выполнения шагов 1–4 алгоритма анализа) для подлежащей синтезу системы последовательностей интервалов (1) в предположении, что задан показатель взаиморасположения интервалов, а все

параметры системы (1) (координаты начала и конца всех интервалов) заданы в буквенной форме. В результате получается двоичный процесс $y(t)$, моделирующий заданный показатель взаиморасположения интервалов (1) (точнее, содержащий импульсы-периоды, где интервалы находятся в заданном взаиморасположении).

Шаг 2. Составление системы уравнений и неравенств, выражающих в математической форме заданные требования к взаиморасположению интервалов (1). Эта система получается путем выписывания требуемых соотношений ($>$, $=$) между координатами начала и конца соответствующих импульсов процесса $y(t)$. Поскольку эти координаты выражаются через параметры интервалов (1) с помощью операций НЛ, полученная система есть система уравнений и неравенств НЛ.

Шаг 3. Решение системы уравнений и неравенств НЛ, полученной на шаге 2, с помощью специальных методов [1, 2, 5, 6], основанных на принципе последовательного расчленения отдельного уравнения (неравенства) НЛ на несколько более простых уравнений (неравенств). В результате решения указанной системы уравнений (неравенств) НЛ получаются условия на параметры отдельных интервалов (1), при выполнении которых взаиморасположение этих интервалов отвечает заданным требованиям.

Конец алгоритма.

Подчеркнем, что предложенные решения задач анализа и синтеза системы последовательностей интервалов получаются в аналитической форме, в терминах суперпозиции операций НЛ \vee (дизъюнкция) и \wedge (конъюнкция).

4. Примеры и оценка сложности вычислений

Пример 1. Магазин открыт в течение дня с a_{11} до b_{11} и с a_{12} до b_{12} часов, где $b_{11} < a_{12}$. Двое работающих приятелей собираются вместе посетить магазин. Первый из них свободен и может это сделать в промежутке времени с a_{21} до b_{21} часов, аналогично второй — в промежутке с a_{31} до b_{31} часов. Требуется определить периоды времени, в каждом из которых приятели могут реализовать свой план посещения магазина, и установить, при каких условиях это возможно, т.е. данные периоды существуют (не вырождены).

Приятели могут реализовать свой план посещения магазина в те и только те периоды времени, когда магазин открыт, а оба они свободны. Следовательно, для ответа на первый поставленный вопрос надо решить задачу анализа системы последовательностей интервалов

$$A_1 = (a_{11}, b_{11}), (a_{12}, b_{12}); A_2 = (a_{21}, b_{21}); \\ A_3 = (a_{31}, b_{31}),$$

т.е. определить нужное взаиморасположение интервалов данной системы. Конкретно, нас интересуют те периоды времени, в которых взаимо-

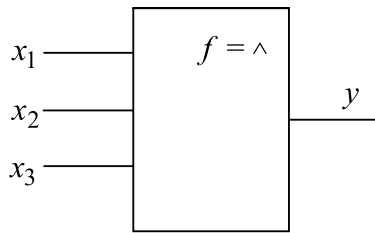


Рис. 3. ДА без памяти, реализующий 3-местную конъюнкцию

расположение таково, что присутствуют интервалы всех трех последовательностей: A_1, A_2, A_3 . Для решения задачи применим алгоритм анализа п. 2.

Шаг 1. Показатель П, характеризующий нужное взаиморасположение интервалов системы A_1, A_2, A_3 , содержательно уже задан условиями задачи.

Шаг 2. Булева логическая функция $y = f(x_1, x_2, x_3)$, соответствующая показателю П, есть трехместная конъюнкция $y = x_1 \wedge x_2 \wedge x_3$.

Шаг 3. Математическая модель задачи — ДА без памяти с тремя входами и одним выходом, реализующий на выходе указанную функцию f своих входов (рис. 3). На входы ДА-модели поступают процессы

$$\begin{aligned} x_1(t) &= 1(a_{11}, b_{11})0(-, -)1(a_{12}, b_{12}), \\ x_2(t) &= 1(a_{21}, b_{21}), \\ x_3(t) &= 1(a_{31}, b_{31}), \end{aligned}$$

которые взаимно однозначно соответствуют системе интервалов (A_1, A_2, A_3) . С выхода ДА снимается двоичный процесс $y(t)$ вида (4), моделирующий показатель П взаиморасположения системы интервалов (A_1, A_2, A_3) .

Шаг 4. По входным процессам $x_1(t), x_2(t), x_3(t)$ ДА-модели и его реализуемой функции $f = \wedge$ вычисляем его выходной процесс $y(t)$ [1, 2, 4], используя готовую формулу $1(a, b) \wedge 1(c, d) = 1[a \vee c, a \vee c \vee (b \wedge d)]$:

$$\begin{aligned} y(t) &= x_1(t) \wedge x_2(t) \wedge x_3(t) = x_1(t) \wedge [x_2(t) \wedge x_3(t)] = \\ &= [1(a_{11}, b_{11})0(-, -)1(a_{12}, b_{12})] \wedge [1(a_{21}, b_{21}) \wedge \\ &\wedge 1(a_{31}, b_{31})] = [1(a_{11}, b_{11})0(-, -)1(a_{12}, b_{12})] \wedge \\ &\wedge 1[a_{21} \vee a_{31}, a_{21} \vee a_{31} \vee (b_{21} \wedge b_{31})] = \\ &= \{1(a_{11}, b_{11}) \wedge 1[\cdot]\}0(-, -)\{1(a_{12}, b_{12}) \wedge 1[\cdot]\} = \\ &= 1\{a_{11} \vee a_{21} \vee a_{31}, a_{11} \vee a_{21} \vee a_{31} \vee \\ &\vee [b_{11} \wedge (a_{21} \vee a_{31} \vee (b_{21} \wedge b_{31}))]\} \times \\ &\times 0(-, -)1\{a_{12} \vee a_{21} \vee a_{31}, a_{12} \vee a_{21} \vee a_{31} \vee \\ &\vee [b_{12} \wedge (a_{21} \vee a_{31} \vee (b_{21} \wedge b_{31}))]\}. \end{aligned}$$

Шаг 5. Найденные на шаге 4 аналитические выражения параметров A, B, C, D выходного процесса $y(t) = 1(A, B)0(-, -)1(C, D)$ ДА-модели дают алгоритмы нахождения указанного процесса в терминах операций НЛ \vee (max) и \wedge (min), а именно, $A = \max(a_{11}, a_{21}, a_{31}), B = \max(A, \min(b_{11}, \max(a_{21}, a_{31}, \min(b_{21}, b_{31}))))$ и т.д.

Шаг 6. По алгоритмам, найденным на шаге 5, находим параметры выходного процесса $y(t)$ ДА-модели, соответствующие конкретным значениям параметров a_{ij}, b_{ij} входных процессов $x_1(t), x_2(t), x_3(t)$. Так, например, при $a_{11} = 9, a_{12} = 14, b_{11} = 13, b_{12} = 20, a_{21} = 12, b_{21} = 16, a_{31} = 11, b_{31} = 15$ находим $A = 12, B = 13, C = 14, D = 15$. Таким образом, есть два периода времени, в которых оба приятеля могут совместно посетить магазин: (12, 13) и (14, 15).

Для того чтобы установить, при каких общих условиях возможно совместное посещение магазина двумя приятелями, надо определить, когда существуют (не вырождены) периоды времени, в которых приятели совместно могут посетить магазин. Так, для ответа на второй поставленный вопрос требуется решить задачу синтеза системы (A_1, A_2, A_3) последовательностей интервалов, т.е. найти условия, при которых взаиморасположение интервалов этой системы имеет нужный качественный характер, благодаря чему и обеспечивается существование указанных периодов времени. Для решения задачи применим алгоритм синтеза п. 2.

Шаг 1. Уже выполнен, так как содержится в шагах 1—4 алгоритма анализа системы (A_1, A_2, A_3) , выполненных выше.

Шаг 2. Систему уравнений и неравенств НЛ, выражающих требования к взаиморасположению интервалов системы (A_1, A_2, A_3) , обеспечивающие существование нужных периодов времени, получаем, потребовав, чтобы найденный в процессе анализа системы последовательностей (A_1, A_2, A_3) выходной процесс $y(t) = 1(A, B)0(-, -)1(C, D)$ ДА-модели, моделирующий заданный показатель взаиморасположения интервалов, имел невырожденные импульсы, моделирующие нужные нам периоды времени: $B > A$ или $D > C$. Если подставить в систему выражения для A, B, C, D из развернутого выражения процесса $y(t)$, приведенного выше, получим

$$\begin{aligned} a_{11} \vee a_{21} \vee a_{31} \vee [b_{11} \wedge (a_{21} \vee a_{31} \vee (b_{21} \wedge b_{31}))] &> \\ &> a_{11} \vee a_{21} \vee a_{31} \text{ или} \\ a_{12} \vee a_{21} \vee a_{31} \vee [b_{12} \wedge (a_{21} \vee a_{31} \vee (b_{21} \wedge b_{31}))] &> \\ &> a_{12} \vee a_{21} \vee a_{31}. \end{aligned}$$

Шаг 3. Решение полученной на шаге 2 системы неравенств НЛ имеет вид

$$\begin{aligned} b_{11} \wedge [a_{21} \vee a_{31} \vee (b_{21} \wedge b_{31})] &> a_{11} \vee a_{21} \vee a_{31} \text{ или} \\ b_{12} \wedge [a_{21} \vee a_{31} \vee (b_{21} \wedge b_{31})] &> a_{12} \vee a_{21} \vee a_{31}. \end{aligned}$$

Итак, взаиморасположение интервалов системы (A_1, A_2, A_3) , отвечающее хотя бы одному из выписанных неравенств, обеспечивает существование периодов времени, в которых приятели могут совместно посетить магазин. Еще раз обратим внимание, что обе части неравенств (условий) представляют собой выражения, построенные из параметров интервалов системы (A_1, A_2, A_3) с помощью операций НЛ — дизъюнкции \vee и конъюнкции \wedge .

Пример 2. Усложним пример 1, приняв, что число приятелей, собравшихся вместе посетить магазин, равно $n - 1$, при этом первый из них может это осуществить в промежутке времени с a_{21} до b_{21} , второй — в промежутке времени с a_{31} до b_{31} , ..., $(n - 1)$ -й — в промежутке с a_{n1} до b_{n1} . Следовательно, теперь для определения периодов времени, в каждом из которых приятели могут посетить магазин, необходимо решить задачу анализа системы последовательностей интервалов

$$A_1 = (a_{11}, b_{11}), (a_{12}, b_{12}); A_2 = (a_{21}, b_{21}); \\ A_3 = (a_{31}, b_{31}); \dots, A_n = (a_{n1}, b_{n1}),$$

найдя соответствующие периоды времени, в которых присутствуют интервалы из всех n последовательностей A_1, A_2, \dots, A_n . Снова применим алгоритм анализа из п. 3.

Шаг 1. Показатель Π , определяющий нужное взаиморасположение интервалов A_1, A_2, \dots, A_n , задан условиями задачи.

Шаг 2. Соответствующая Π булева логическая функция f — n -местная конъюнкция $y = \bigwedge_{i=1}^n x_i$.

Шаг 3. Математическая модель задачи — ДА без памяти с n входами и одним выходом, где реализуется функция f . На входы ДА поступают процессы, соответствующие системе интервалов A_1, \dots, A_n : $x_1(t) = 1(a_{11}, b_{11})0(-, -)1(a_{12}, b_{12})$, $x_2(t) = 1(a_{21}, b_{21})$, $x_3(t) = 1(a_{31}, b_{31}), \dots, x_n(t) = 1(a_{n1}, b_{n1})$.

Выходной процесс $y(t)$ вида (4) моделирует показатель Π взаиморасположения интервалов A_1, \dots, A_n .

Шаг 4. По входным процессам $x_1(t), \dots, x_n(t)$ ДА-модели и функции $f = \wedge$ находим выходной процесс $y(t)$ [1, 2, 4], используя для этого готовую формулу $\bigwedge_{i=1}^N 1(a_i, b_i) = 1 \left[\bigvee_{i=1}^N a_i, \bigvee_{i=1}^N a_i \vee \left(\bigwedge_{i=1}^N b_i \right) \right]$:

$$y(t) = \bigwedge_{i=1}^n x_i(t) = x_1(t) \wedge \left[\bigwedge_{i=2}^n x_i(t) \right] =$$

$$= [1(a_{11}, b_{11})0(-, -)1(a_{12}, b_{12})] \wedge \left[\bigwedge_{i=2}^n 1(a_{i1}, b_{i1}) \right],$$

и после действий, аналогичных шагу 4 примера 1, находим

$$y(t) = 1 \left\{ a_{11} \vee \bigvee_{i=2}^n a_{i1}, a_{11} \vee \bigvee_{i=2}^n a_{i1} \vee \left[b_{11} \wedge \left(\bigvee_{i=2}^n a_{i1} \vee \left(\bigwedge_{i=2}^n b_{i1} \right) \right) \right] \right\} 0(-, -) 1 \left\{ a_{12} \vee \bigvee_{i=2}^n a_{i1}, a_{12} \vee \bigvee_{i=2}^n a_{i1} \vee \left[b_{12} \wedge \left(\bigvee_{i=2}^n a_{i1} \vee \left(\bigwedge_{i=2}^n b_{i1} \right) \right) \right] \right\}.$$

Шаг 5. Найденные на шаге 4 выражения параметров A, B, C, D выходного процесса $y(t) = 1(A, B)0(-, -)1(C, D)$ ДА-модели дают, как и в

примере 1, алгоритмы вычисления процесса в терминах операций НЛ \vee (max), \wedge (min).

Шаг 6. По алгоритмам, найденным на шаге 5, вычисляем параметры выходного процесса $y(t)$ ДА-модели при конкретных значениях параметров входных процессов $x_1(t), \dots, x_n(t)$, находя периоды (A, B) и (C, D) , в которых все $n - 1$ приятелей могут посетить магазин.

Конец алгоритма.

Для нахождения общих условий, при которых возможно совместное посещение магазина $n - 1$ приятелями, определим, когда существуют (не вырождены) найденные выше периоды времени, в которых приятели могут вместе посетить магазин. Для этого нужно решить задачу синтеза системы последовательностей интервалов (A_1, A_2, \dots, A_n) , найдя условия, при которых взаиморасположение интервалов этой системы имеет нужный характер, обеспечивающий существование данных периодов времени. Как и в примере 1, используем алгоритм синтеза п. 3.

Шаг 1. Этот шаг уже выполнен — см. шаги 1—4 алгоритма анализа системы (A_1, A_2, \dots, A_n) , выполненные выше.

Шаг 2. Выполняется аналогично шагу 2 алгоритма синтеза в примере 1, а именно, требуем, чтобы найденный при анализе системы (A_1, A_2, \dots, A_n) выходной процесс $y(t) = 1(A, B)0(-, -)1(C, D)$ ДА-модели, моделирующий показатель взаиморасположения интервалов системы, имел невырожденные импульсы, моделирующие нужные нам периоды времени: $B > A$ или $D > C$. Подставив сюда выражения A, B, C, D из выражения $y(t)$, полученного выше при анализе системы, найдем условие

$$a_{11} \vee \bigvee_{i=2}^n a_{i1} \vee \left[b_{11} \wedge \left(\bigvee_{i=2}^n a_{i1} \vee \left(\bigwedge_{i=2}^n b_{i1} \right) \right) \right] > \\ > a_{11} \vee \bigvee_{i=2}^n a_{i1} \text{ или} \\ a_{12} \vee \bigvee_{i=2}^n a_{i1} \vee \left[b_{12} \wedge \left(\bigvee_{i=2}^n a_{i1} \vee \left(\bigwedge_{i=2}^n b_{i1} \right) \right) \right] > \\ > a_{12} \vee \bigvee_{i=2}^n a_{i1}.$$

Шаг 3. Решение полученной на шаге 2 системы неравенств НЛ таково:

$$b_{11} \wedge \left[\bigvee_{i=2}^n a_{i1} \vee \left(\bigwedge_{i=2}^n b_{i1} \right) \right] > a_{11} \vee \bigvee_{i=2}^n a_{i1} \\ \text{или } b_{12} \wedge \left[\bigvee_{i=2}^n a_{i1} \vee \left(\bigwedge_{i=2}^n b_{i1} \right) \right] > a_{12} \vee \bigvee_{i=2}^n a_{i1}.$$

Взаиморасположение интервалов системы (A_1, A_2, \dots, A_n) , удовлетворяющее хотя бы одному из двух выписанных неравенств, обеспечивает существование периодов времени, в которых $n - 1$

приятелей могут совместно посетить магазин — (A, B) и (C, D) .

Как видно из примера 2, усложнение задачи в виде увеличения размерности рассматриваемой системы никак не повлияло на возможность ее решения в аналитической форме.

Оценим теперь вычислительную сложность нашего подхода. Базовая задача в этом подходе (анализ системы n последовательностей интервалов) эквивалентна вычислению выходного процесса ДА-модели системы, которая согласно [5, с. 93] имеет сложность

$$N \leq 2mn[\log_2(2m)(l^r - 1)/(l - 1) + \log_2 l(r l^{r+1} - (r + 1)l^r + 1)/(l - 1)^2]$$

элементарных операций. В этом выражении m — число интервалов в каждой из последовательностей, l — число входов в каждом логическом элементе ДА-модели, r — число ступеней в ДА-модели. Реализуя (это всегда возможно) двухступенчатую модель ($r = 2$), получим

$$N \leq 2mn[(\log_2 m + 1)(l + 1) + \log_2 l(2l^2 + 1)].$$

Таким образом, вычислительная сложность предложенного подхода возрастает как степенная функция от размерности задачи $m \times n \times l$, что позволяет применять этот подход к решению задач высокой размерности как в аналитической, так и численной форме.

Заключение

В настоящей работе показано, что изучение класса комбинаторных задач, эквивалентных комбинаторной задаче определения взаиморасположения n последовательностей интервалов, можно осуществлять с помощью математической модели динамического конечного автомата без памяти и математического аппарата непрерывной логики. Такой подход позволяет формально находить алгоритмы решения указанных задач, а также формально анализировать эти решения, например, находить необходимые и достаточные условия их существования. Другим преимуществом предложенного подхода является его применимость к решению задач произвольно высокой размерности.

Список литературы

1. **Левин В. И.** Введение в динамическую теорию конечных автоматов. Рига: Зинатне, 1975. 376 с.
2. **Bochmann D., Roginskij V. N., Levin V. I.** *Dinamische Prozesse in Automaten*. Berlin: Technik, 1977. 285 p.
3. **Левин В. И.** Динамика логических устройств и систем. М.: Энергия, 1980. 228 с.
4. **Левин В. И.** Теория динамических автоматов. Пенза: Изд-во Пенз. гос. ун-та, 1995. 407 с.
5. **Левин В. И.** Бесконечнозначная логика в задачах кибернетики. М.: Радио и связь, 1982. 176 с.
6. **Левин В. И.** Структурно-логические методы исследования сложных систем. М. Наука, 1987. 304 с.
7. **Заде Л.** Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976. 150 с.
8. **Кандрашина Е. Ю., Литвинцева Л. В., Поспелов Д. А.** Пространство и время в системах искусственного интеллекта. М.: Наука, 1988. 328 с.
9. **Кандрашина Е. Ю., Литвинцева Л. В., Поспелов Д. А.** Представление знаний о пространстве и времени в системах искусственного интеллекта. М.: Наука, 1989. 328 с.

V. I. Levin, Dr of Tech Sci, Professor, e-mail: vilevin@mail.ru; Penza State Technological Academy

Continuous-Logical Model for Solution of Combinatorial Problems

The class of combinatorial problems equivalent to the problem of determination of mutual dislocation of interval sequences is formulated. It is shown, that an adequate mathematical model of solution of delivered problem is the finite dynamic automaton without memory, and adequate mathematical means — continuous logic. The algorithms of a solution are constructed. The example is indicated.

Keywords: continuous logic, combinatorial problems, automata model, interval value

References

1. **Levin V. I.** *Vvedenie v dinamicheskuyu teoriyu konechnykh avtomatov* (Introduction to the dynamic theory of finite automata), Riga, Zinatne, 1975, 376 p. (in Russian).
2. **Bochmann D., Roginskij V. N., Levin V. I.** *Dinamische Prozesse in Automaten*, Berlin, Technik, 1977, 285 p.
3. **Levin V. I.** *Dinamika logicheskikh ustroystv i sistem* (Dynamics of logic devices and systems), Moscow, Energiya, 1980, 228 p. (in Russian).
4. **Levin V. I.** *Teoriya dinamicheskikh avtomatov* (The theory of dynamical machines), Penza, Publishing house of Penz. gos. university, 1995, 407 p. (in Russian).
5. **Levin V. I.** *Beskonechnoznachnaya logika v zadachakh kibernetiki* (Infinite-logic problems in cybernetics), Moscow, Radio i svyaz', 1982, 176 p. (in Russian).
6. **Levin V. I.** *Strukturno-logicheskie metody issledovaniya slozhnykh sistem* (Structural and logical methods for complex systems research), Moscow, Nauka, 1987, 304 p. (in Russian).
7. **Zade L.** *Ponyatie lingvisticheskoi peremennoi i ego primeneniye k prinyatiyu priblizhennykh reshenii* (The concept of linguistic variable and its application to the adoption of the approximate solutions), Moscow, Mir, 1976, 150 p. (in Russian).
8. **Kandrashina E.Yu., Litvintseva L. V., Pospelov D. A.** *Prostranstvo i vremya v sistemakh iskusstvennogo intellekta* (Space and time in artificial intelligence systems), Moscow, Nauka, 1988, 328 p. (in Russian).
9. **Kandrashina E.Yu., Litvintseva L. V., Pospelov D. A.** *Predstavleniye znaniy o prostranstve i vremeni v sistemakh iskusstvennogo intellekta* (Knowledge representation of space and time in artificial intelligence systems), Moscow, Nauka, 1989, 328 p. (in Russian).

УДК 004.72; 621.391

- А. А. Дворников**, аспирант, ассистент департамента Компьютерной инженерии
МИЭМ НИУ ВШЭ, e-mail: advornikov@hse.ru,
Л. С. Восков, канд. техн. наук, проф. департамента Компьютерной инженерии
МИЭМ НИУ ВШЭ, e-mail: lvoskov@hse.ru,
Е. А. Саксонов, д-р техн. наук, проф.,
МГУТУ им. К. Г. Разумовского, saksmiem@mail.ru,
С. Г. Ефремов, канд. техн. наук, ст. преподаватель, e-mail: sefremov@hse.ru,
НИУ ВШЭ

Метод построения оптимального наложенного канала для беспроводной сенсорной сети

Предложен метод поиска оптимального маршрута наложенного канала связи поверх беспроводной сенсорной сети, позволяющий повысить эффективность использования свободных телекоммуникационных и энергетических ресурсов беспроводных сенсорных сетей без нарушения их первичной функции.

Ключевые слова: телекоммуникационные сети, беспроводные сенсорные сети, поиск маршрута, свободный телекоммуникационный ресурс, свободный энергетический ресурс, недогруженные сети, свободные ресурсы, поиск маршрута минимальной длины с максимальным весом, балансировка нагрузки, наложенные сети, наложенный канал связи

Введение

В статье представлен разработанный метод поиска оптимального маршрута наложенного канала беспроводной сенсорной сети, исключающего выбор технически нереализуемых случаев.

Беспроводные сенсорные сети — это сети датчиков и актуаторов, отличительной особенностью которых является большое число автономных узлов с малым энергопотреблением, малой дальностью связи и небольшой пропускной способностью. Одно из главных преимуществ подобных сетей — захват больших территорий за счет применения механизмов ретрансляции сообщений от узла к узлу. В отличие от классических сетей в беспроводных сенсорных сетях энергия на узлах — важная характеристика при планировании дальнейшей работы.

Беспроводные сети по своей природе широковещательны и могут устанавливать связь со всеми узлами в радиусе действия, что затрудняет анализ подобных сетей. В беспроводных сенсорных сетях радиус действия приемопередатчика для большинства узлов невелик, поэтому узел может взаимодействовать с ограниченным числом узлов в сети. Для изображения таких сетей можно использовать граф, где ребро — потенциально возможная связь с другим узлом.

Не все телекоммуникационные ресурсы сетей телесвязи, в том числе и беспроводные сенсорные, используются полностью [1—4]. Событийная активность сетей, направленность на энергосбережение, неэффективная инфраструктура, инженерная избыточность, мобильность, неравномерное распределение ресурсов в узлах приводят к образованию свободных телекоммуникационных и энергетических остатков в современных беспроводных сенсорных сетях, переводя их в класс недогруженных сетей телесвязи.

Имеется ряд задач, связанных с передачей данных, организацией временных или постоянных каналов телесвязи и телекоммуникационных сетей в ситуациях, когда применение существующих на сегодняшний день технических средств невозможно либо затруднено. Предлагается решить данные задачи за счет использования свободных телекоммуникационных ресурсов беспроводных сенсорных сетей. Данный подход отличается тем, что решение задач не требует организации новых сетей и масштабирования существующих, позволяет выполнять балансировку существующей беспроводной сенсорной сети.

В данной работе рассматривается случай организации обособленного наложенного канала. Взаимовлияние нескольких каналов — задача от-

дельного исследования. Расчет маршрутов канала проводится на приграничных шлюзах в режиме реального времени.

Постановка задачи

Узлы беспроводных сенсорных сетей обладают свободными энергетическим и телекоммуникационным ресурсами. Одним из способов использования свободных телекоммуникационных ресурсов являются наложенные каналы (*overlay channels*) и наложенные сети (*overlay networks*) [5, 6] — телекоммуникационной сети, состоящей из нескольких (возможно, пересекающихся) наложенных каналов.

При поиске оптимального маршрута для канала наложенной сети используются те же алгоритмы, что и в классических сетях, при условии применения графовой модели.

В статье [7], посвященной сбору данных в беспроводной сенсорной сети со множеством стоков, графовая модель дополнена емкостью и интерференцией в канале связи на ребре графа. В работе [8] рассмотрена задача формирования структуры базовой сети телекоммуникационной системы при наличии нескольких провайдеров. Для решения поставленной задачи предложена модификация графовой модели, где веса присваиваются не только ребрам графа, но и узлам. В статье [9] представлена графовая модель, где на узлы и ребра мультиграфа накладываются несколько характеристик для решения задачи адаптивной маршрутизации в корпоративных сетях нескольких провайдеров связи.

В результате анализа литературы обнаружена практика модификации графовой модели для решения более частных задач, в том числе за счет добавления одного или нескольких весов на узел, добавления нескольких весов на ребро.

Имеется связный граф сети, предлагается модель беспроводной сенсорной сети и метод поиска оптимального маршрута для наложенного канала, характеризующиеся тем, что в модели на узлы и ребра накладываются следующие параметры: свободный телекоммуникационный ресурс на узле, свободная энергия на узле, качество связи соединения на ребре и профиль соединения на ребре, — позволяющие более точно отобразить предметную область и исключить выбор нереализуемых маршрутов при применении предложенного метода за счет соблюдения минимумов свободной пропускной способности и энергии на узле, а также минимума качества связи и максимума потерь сообщений при наложении профилей на ребре графа.

Математическая модель наложенного канала

Пусть беспроводная сенсорная сеть представлена в виде графа сети $G_w = (V_w, E_w)$, где $V_w = \{v_1, v_2, \dots, v_m\}$: $m \in \mathbb{N}$, $m > 0$ — множество вер-

шин, $E_w = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$: $n \in \mathbb{N}$, $n > 0$ — множество ребер графа сети, $s \in V_w$ — исток маршрута, $t \in V_w$ — сток маршрута. Вершинам $v = (b, e)$ присвоены два параметра: первый параметр $b \in \mathbb{Z}^+$ — телекоммуникационный остаток (неиспользуемая в настоящее время часть пропускной способности приемопередатчика узла), второй параметр $e \in \mathbb{Z}^+$ — энергетический остаток узла (энергия, которую узел может выделить для наложенного канала без риска выйти из строя в текущий момент). Для каждого ребра графа сети $\varepsilon = (q, p_w)$ определены два параметра: первый — $q \in \mathbb{R}$, $q \in \{0, \dots, 1\}$ — определяет качество связи между двумя узлами так, что значение $q = 1$ соответствует максимальному качеству связи, а $q = 0$ — полному ее отсутствию; второй — $p_w \subseteq P_w$ задает профиль канала беспроводной сенсорной сети между двумя узлами (характеристика трафика в канале беспроводной сенсорной сети). Профиль накладываемого на беспроводную сенсорную сеть канала является важной его характеристикой [10]: он определяет процент потерь, возникающих при наложении потоков сообщений наложенного канала и канала беспроводной сенсорной сети. Состав множества профилей P_w зависит от решаемой задачи.

Если узел получает энергоснабжение извне и энергия потенциально бесконечна, то $e = e_{\max}$, $e_{\max} \in \mathbb{N}$, $e_{\max} > 0$, где e_{\max} — максимально возможное количество энергии на узле.

Матрица совместимости профилей беспроводной сенсорной сети и наложенной сети $\Sigma_c = P_w \times P_n$, где элементы матрицы $\sigma_{c_{ij}} \in \mathbb{R}$, $\sigma_{c_{ij}} \in \{0, \dots, 1\}$, $i = 1, \dots, m$, $j = 1, \dots, n$, обозначают процент потерянных пакетов из-за наложения потоков сообщений канала наложенной сети и беспроводной сенсорной сети, а P_n — множество профилей наложенного канала. Пусть $r = \zeta(a, b)$: $r \in \Sigma_c$, $a \in P_w$, $b \in P_n$ возвращает из матрицы совместимости Σ_c процент потерь сообщений, возникших из-за наложения потока сообщений беспроводной сенсорной сети, соответствующему профилю из множества P_w , и потока сообщений канала наложенной сети, соответствующему профилю из множества P_n .

Пусть $y = w(u, v)$: $u \in V_s$, $v \in V_s$, $y \in \mathbb{R}^+$, — весовая функция ребра графа. Определим весовую функцию $w(u, v)$ следующим образом:

$$w(u, v) = k_b \min\left(\frac{b_u}{k_s}, \frac{b_v}{k_t}\right) + k_e \min(e_u, e_v) + k_a q_{uv} - k_a \zeta(p_{w_{uv}}, p_n) + C_n. \quad (1)$$

Константа $C_n \in \mathbb{R}^+$: $C_n = \max(k_a \zeta(p_{w_{uv}}, p_n))$ компенсирует отрицательную составляющую совместимости каналов связи и смещает функцию на

множество \mathbb{R}^+ . Коэффициенты $k_s = \begin{cases} 1, u = s \\ 2, u \neq s \end{cases}$ и

$k_t = \begin{cases} 1, v = t \\ 2, v \neq t \end{cases}$ позволяют выбирать большую про-

пускную способность в том случае, если узел является истоком либо стоком маршрута. $k_b \in \mathbb{R}^+$, $k_e \in \mathbb{R}^+$, $k_q \in \mathbb{R}^+$, $k_a \in \mathbb{R}^+$ — весовые коэффициенты для предложенного метода, которые задаются исходя из необходимых приоритетов для решения конкретной задачи для свободной пропускной способности на узле, свободного энергетического остатка, качества канала связи, совместимости профилей беспроводной сенсорной сети и наложенной сети, соответственно.

Задача поиска оптимального маршрута для канала наложенной сети сформирована следующим образом. Пусть $R_{st} \subseteq E_w$ — выбранный маршрут от s до t , $p_n \in P_n$ — профиль наложенного канала):

$$\begin{cases} \sum_{\substack{v, u \in V_w; \\ vu \in R_{st}}} w(u, v) \rightarrow \max; \\ b \geq b_{\min} \quad \forall v \in V_w: vu \in R_{st}; \\ e \geq e_{\min} \quad \forall v \in V_w: vu \in R_{st}; \\ q \geq q_{\min} \quad \forall uv \in R_{st}; \\ \zeta(p_{w_{uv}}, p_n) \geq \sigma_{\max} \quad \forall uv \in R_{st}. \end{cases} \quad (2)$$

Метод поиска оптимального маршрута для канала наложенной сети

Предложенный метод состоит из следующих этапов:

Этап 1. Преобразование предложенного выше графа беспроводной сенсорной сети G_w к графу сети G_s сети с помощью Алгоритма 1.

Этап 2. Выбор алгоритма поиска кратчайшего пути максимальной стоимости исходя из минимальной асимптотической вычислительной сложности.

Этап 3. Поиск кратчайшего пути максимальной стоимости с помощью выбранного алгоритма.

Оптимальный набор алгоритмов поиска кратчайшего пути максимальной стоимости зависит от решаемой задачи.

Алгоритм 1: Преобразование графа беспроводной сенсорной сети G_w в граф сети G_s

Алгоритм 1 служит для преобразования беспроводной сенсорной сети в граф сети, опираясь на информацию о стоке и истоке сети, а также на профиль запланированного для организации наложенного канала. Полученный на выходе сетевой граф может быть обработан любым алгоритмом, работающим на взвешенных неориентированных сетевых графах.

В Алгоритме 1 используются следующие входные и выходные переменные:

- $G_w = (V_w, E_w)$ — беспроводная сенсорная сеть, которую необходимо преобразовать в граф сети;
- $s \in V_w$ — узел-исток наложенного канала (инициатор соединения);
- $t \in V_w$ — узел-сток наложенного канала (удаленная сторона);
- $p_n \in P_w$ — профиль канала наложенной сети (характеристика трафика в создаваемом наложенном канале);
- $N: \forall n_c \in n: n \in V_w$ — вектор характеристик узлов (b и e , применительно к исходной задаче, где b — телекоммуникационный остаток, а e — энергетический остаток на узле);
- $L: \forall l_c \in e: e \in E_w$ — вектор характеристик соединений между узлами (q и p_w , применительно к текущей задаче, где q — качество связи, а p_w — профиль трафика между двумя узлами);
- $N_{\min} = \{r_1, \dots, r_{|M|}\}$ — вектор ограничений снизу на характеристики узла. Значение входит в ограничение включительно (b_{\min} и e_{\min} , применительно для текущей задачи, где b_{\min} — ограничение снизу на свободный телекоммуникационный ресурс, а e_{\min} — ограничение снизу на свободный энергетический остаток на узле);
- $L_{\min} = \{w_1, \dots, w_{|L|}\}$ — вектор ограничений снизу на характеристики соединения между узлами (q_{\min} и 0, применительно к текущей задаче, где q_{\min} — ограничение снизу на качество связи между узлами). Значение входит в ограничение включительно;
- $N_{\max} = \{z_1, \dots, z_{|M|}\}$ — вектор ограничений сверху на характеристики узла (0 и 0, применительно к текущей задаче);
- $L_{\max} = \{x_1, \dots, x_{|L|}\}$ — вектор ограничений сверху на характеристики соединения между узлами (0 и σ_{\max} , применительно к текущей задаче, где σ_{\max} — максимальные потери пакетов при наложении потоков сообщений беспроводной сенсорной сети и наложенного канала).

Алгоритм 1 состоит из следующих шагов:

Шаг 1. Создать копию исходного графа W_s с новым именем G_s .

Шаг 2. Проверить все узлы v графа G_s .

Шаг 3. Проверить все характеристики p узла v . Проверить, удовлетворяет ли значение характеристики p узла v нижнему и верхнему пределу:

1) если нет — удалить узел v из графа G_s и все связанные с ним каналы.

Шаг 4. Проверить все дуги l от узла v .

Шаг 5. Проверить все характеристики p дуги l , которая еще не была проверена.

Проверить, удовлетворяет ли значение характеристики p дуги l нижнему и верхнему пределу:

1) если нет — удалить дугу l из графа G_s .

Шаг 6. Рассчитать вес дуги l , используя формулу (1).

Шаг 7. Проверить достижимость узла s и t в графе G_s любым алгоритмом, работающим на неориентированных графах.

На основе описания алгоритма проведена аналитическая оценка сложности Алгоритма 1:

$$O(|V|*|E|*|N|*|L|), \quad (3)$$

где V — множество узлов графа беспроводной сенсорной сети; E — множество ребер графа беспроводной сенсорной сети; N — множество ограничений на узел графа беспроводной сенсорной сети; L — множество ограничений на ребро беспроводной сенсорной сети.

Ниже представлен пример преобразования исходного графа беспроводной сенсорной сети (рис. 1) в граф сети (рис. 2) с помощью Алгоритма 1.

Преобразование выполнено для профиля канала наложенной сети 1 (см. рис. 1). В процессе выполнения Алгоритма 1 из исходного графа (см. рис. 1) были удалены следующие элементы:

1. Узел A из-за малого свободного телекоммуникационного ресурса.
2. Узел F из-за малого свободного энергетического ресурса.
3. Канал связи c из-за низкого качества связи.
4. Канал связи i из-за низкой совместимости профиля наложенного канала и профиля беспроводной сенсорной сети.
5. Канал связи n из-за низкой совместимости профиля наложенного канала и профиля беспроводной сенсорной сети.

При выполнении алгоритма использовались следующие ограничения при весовых коэффициентах по умолчанию ($k_b = 1, k_e = 1, k_q = 100, k_\sigma = 100$) (формула 1):

1. Свободный телекоммуникационный ресурс не может быть меньше 200.
2. Свободный энергетический ресурс не может быть меньше 100.
3. Качество связи между узлами должно быть больше 0,3 (30 %).
4. Потери пакетов при совмещении каналов беспроводной сенсорной сети и каналов наложенной сети не должны превышать 0,45 (45 %).

Полученный граф G_s используется в алгоритме поиска кратчайшего пути максимальной стоимости, который завершает процедуру поиска маршрута для канала наложенной сети поверх беспроводной сенсорной сети.

Ниже показана ситуация, когда полученный граф G_s оказывается несвязным (рис. 3).

Граф G_s на рис. 3 получен при тех же условиях, но для профиля канала наложенной сети 2 (см. рис. 1).

В процессе выполнения Алгоритма 1 из исходного графа (см. рис. 1) были удалены следующие элементы:

1. Узел A из-за малого свободного телекоммуникационного ресурса.
2. Узел F из-за малого свободного энергетического ресурса.
3. Канал связи c из-за низкого качества связи.

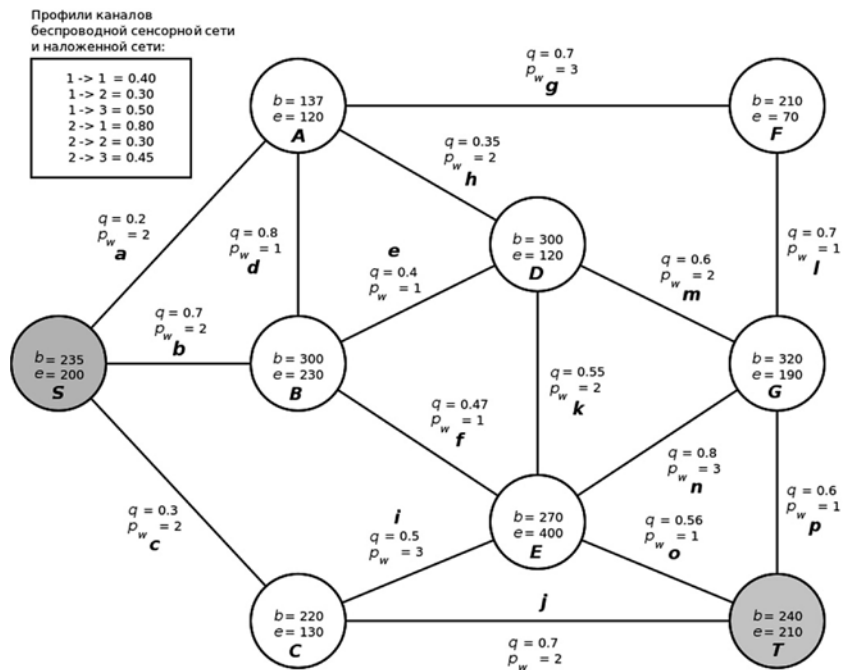


Рис. 1. Граф беспроводной сенсорной сети, соответствующий предложенной модели

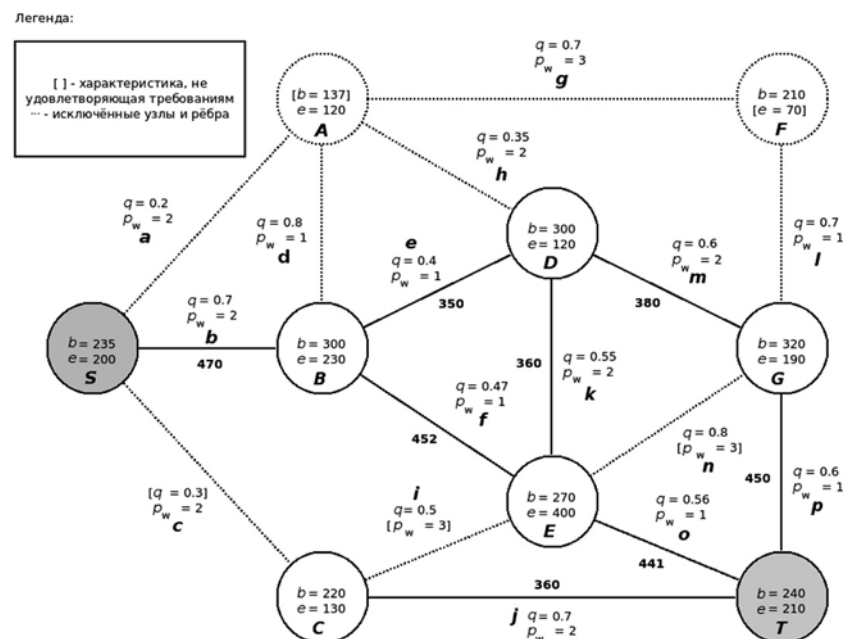


Рис. 2. Граф сети поверх графа беспроводной сенсорной сети, соответствующего предложенной модели

4. Канал связи e из-за низкой совместимости профиля наложенного канала и профиля беспроводной сенсорной сети.

5. Канал связи f из-за низкой совместимости профиля наложенного канала и профиля беспроводной сенсорной сети.

6. Канал связи o из-за низкой совместимости профиля наложенного канала и профиля беспроводной сенсорной сети.

7. Канал связи p из-за низкой совместимости профиля наложенного канала и профиля беспроводной сенсорной сети.

В этом случае наложенный канал между двумя узлами проложить невозможно.

Экспериментальная часть

Сложность метода (формула (4)) зависит от сложности алгоритма преобразования графа беспроводной сенсорной сети в граф сети (Алгоритм 1) и от выбранного алгоритма поиска кратчайшего пути максимального веса:

$$\zeta(V, E, N, L) = \zeta_{alg_1}(V, E, N, L) + \zeta_{sp}(V, E, N, L), (4)$$

где ζ — общая сложность метода; ζ_{alg_1} — сложность Алгоритма 1; ζ_{sp} — сложность алгоритма выбора кратчайшего пути с максимальной сложностью; V — число узлов в графе беспроводной сенсорной сети; E — число ребер в графе беспроводной сенсорной сети; N — число ограничений на узел графа беспроводной сенсорной сети; L — число ограничений на ребро графа беспроводной сенсорной сети.

К формуле (4) есть ряд ограничений:

1. В рассматриваемом варианте метода используются два ограничения на узел беспроводной сен-

сорной сети, поэтому аргументом N можно пренебречь, а его влияние перейдет в мультипликативную константу каждого из слагаемых.

2. В рассматриваемом варианте метода используются два ограничения на ребро беспроводной сенсорной сети, поэтому аргументом L можно пренебречь, а его влияние перейдет в мультипликативную константу каждого из слагаемых.

3. Для определения экспериментальной оценки сложности метода необходимо построить беспроводную сенсорную сеть, выполнив процедуру расположения точек на плоскости и образования ребер между ними. Большинство существующих методов не позволяет указать число ребер графа и узлов одновременно. На чаще других на практике применяется метод построения сети с помощью метода триангуляции, где число ребер зависит линейно от числа узлов. Использование данного метода позволяет избавиться от зависимости сложности метода от числа ребер в графе беспроводной сенсорной сети, переведя данный фактор в мультипликативную константу.

При учете данных факторов формула (4) принимает вид:

$$\zeta(V, E) = \zeta_{alg_1}(V) + \zeta_{sp}(V, E). (5)$$

В связи с адаптивным выбором алгоритма поиска кратчайшего маршрута максимальной стоимости его выбор неизвестен заранее. С точки зрения оценки сложности метода достаточно установить сложность Алгоритма 1.

Ранее аналитически установлено (формула (3)), что Алгоритм 1 линейно зависит от числа узлов. В таком случае, формула сложности Алгоритма 1 (формула (5)) выглядит следующим образом:

$$\zeta_{alg_1}(V) = A_{alg_1} * V + b_{alg_1}, (6)$$

где A_{alg_1} — мультипликативная константа Алгоритма 1, b_{alg_1} — константа смещения Алгоритма 1.

Для восстановления формулы сложности алгоритма достаточно определить A_{alg_1} и b_{alg_1} . Математические расчеты проведены в математической среде Scilab с использованием набора инструментов для проведения расчетов на графах metanet (см. рис. 3). В качестве оценки сложности алгоритма принимается время его выполнения.

На рис. 4 зависимость времени исполнения алгоритма имеет линейный характер. Шум на графиках обусловлен случайным фактором при создании исследуемых беспроводных сенсорных сетей.

Легенда:

[] - характеристика, не удовлетворяющая требованиям
 --- - исключённые узлы и рёбра

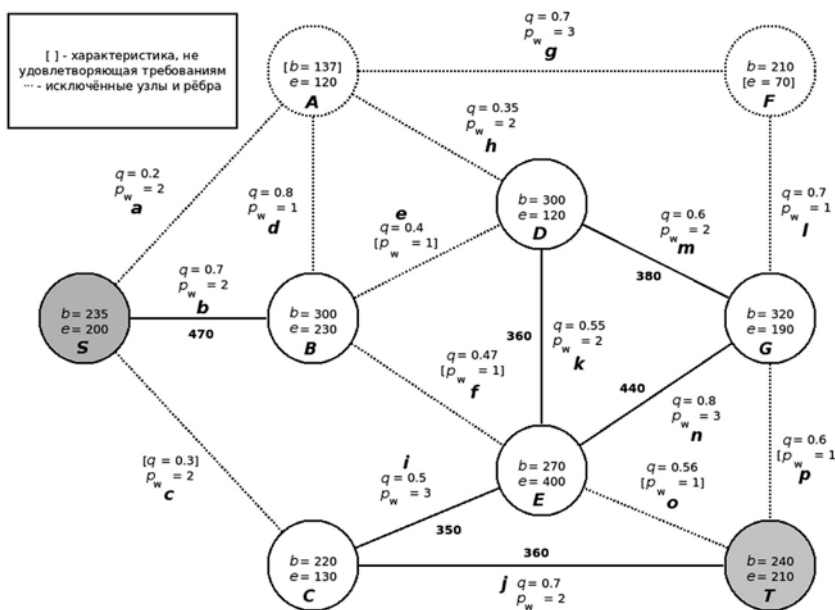


Рис. 3. Граф сети поверх графа беспроводной сенсорной сети, соответствующего предложенной модели (несвязный случай)

По результатам эксперимента получены следующие константы:

- $A_{alg_1} \approx 498,4 \cdot 10^{-5}$ с.
- $b_{alg_1} \approx -836,8 \cdot 10^{-5}$ с.
- $x_{0_{alg_1}} = 2$ — точка, до которой сложность алгоритма принято считать бесконечно малой (значение указывается включительно).
- $\Delta_{alg_1} \approx 8 \cdot 10^{-6}$ с — средняя абсолютная погрешность между восстановленной зависимостью f_r и экспериментальной f_s .
- $\delta_{alg_1} \approx 1,8\%$ — средняя относительная погрешность между восстановленной зависимостью f_r и экспериментальной f_s .

Линейная зависимость Алгоритма 1 от числа узлов сети подтвердилась экспериментально (см. рис. 3). По результатам эксперимента видно, что в беспроводных сенсорных сетях до 20 узлов включительно временные затраты на Алгоритм 1 незначительны (до 100 мс), в сетях от 20 до 100 включительно узлов работа Алгоритма 1 вызывает приемлемые для поставленной задачи задержки (до 500 мс), в сетях с числом узлов более 200 узлов использовать Алгоритм 1 не рекомендуется ввиду большого времени выполнения (более 1 с). Все оценки приведены с учетом ожидаемого увеличения времени выполнения за счет выбранного алгоритма поиска кратчайшего маршрута.

Практическое применение

Полученные ресурсы могут быть направлены на оказание телекоммуникационных услуг внешним клиентам беспроводной сенсорной сети.

Существует ряд ситуаций, когда предложенный метод может быть эффективно применен на практике при организации:

- связи при чрезвычайных ситуациях;
- связи в условиях телекоммуникационной блокады;
- связи в труднодоступных местах;
- высоконадежных каналов связи.

Предложенный метод позволяет найти маршрут и организовать связь при чрезвычайных ситуациях [12], в том числе при катастрофах, для служб спасения тогда, когда иные средства телесвязи вышли из строя, за счет способности беспроводных сенсорных сетей к самовосстановлению.

Большое число узлов в беспроводных сенсорных сетях, в том числе и телекоммуникационных стоков, позволяет организовывать связь в ситуациях информационной войны, когда связь с клиен-

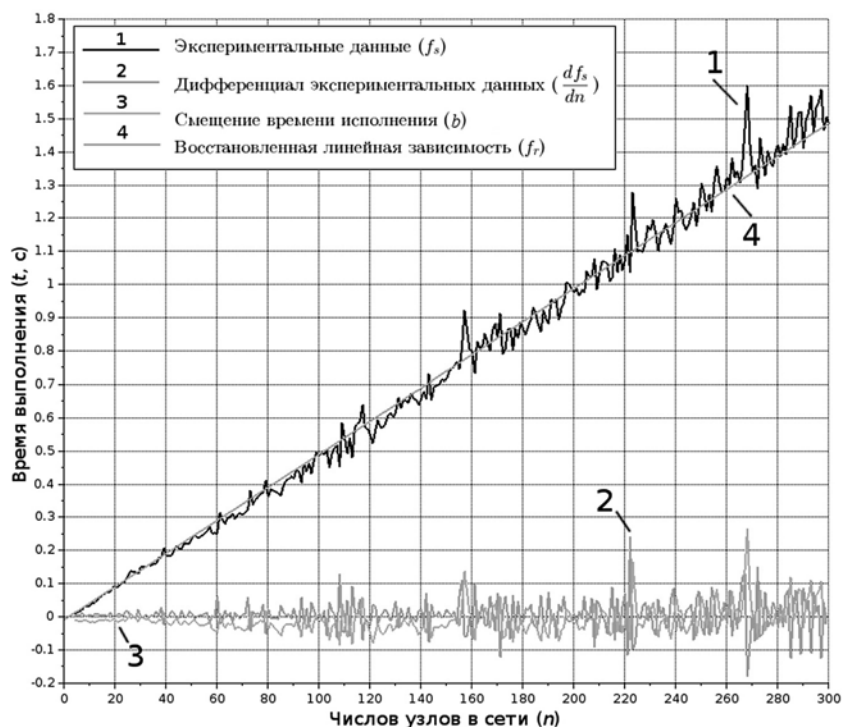


Рис. 4. Результат экспериментального определения сложности Алгоритма 1

том на другой стороне заблокирована по основным телекоммуникационным каналам связи [13]. Возможна организация высоконадежных каналов связи, которые могут быть устойчивы как к естественным, так и искусственным помехам.

Использование предложенного метода в беспроводных сенсорных сетях мониторинга окружающей среды [14] позволяет организовывать связь в тех местах, где иные средства телекоммуникации слишком дороги либо недоступны. Допустимо использование приложения как средства экстренной связи в лесах и иных территориях, где может оказаться человек либо группа людей, вдали от населенных пунктов.

Заключение

Сформирована постановка задачи, разработан метод выбора оптимального маршрута для наложенного канала в беспроводной сенсорной сети. Проведена экспериментальная оценка сложности ключевого алгоритма предложенного метода, которая подтвердила его линейность, восстановлена функция сложности. Установлена область применимости предложенного метода.

По результатам работы остаются нерешенными следующие задачи:

- взаимовлияние и взаимokoординация нескольких наложенных каналов.
- определение достижимости узлов в наложенной сети на раннем этапе (до выполнения преобразования графа и поиска маршрута).

Список литературы

1. **Титов И. М.** Исследование модели трафика сервера данных по результатам измерений трафика мультимедийного ресурса // *T-Comm — Телекоммуникации и Транспорт*. 2011. № 6. С. 46—49.
2. **Кириченко Л. О., Радвиллова Т. А., Стороженко А. В.** Алгоритм предупреждения перегрузки компьютерной сети путем прогнозирования средней длины очереди // *Збірник Наукових Праць Харківського Університету Повітряних Сил*. 2007. № 3. С. 94—97.
3. **Маликова Е. Е.** Метод повышения пропускной способности систем телеметрии и мониторинга на базе беспроводных сетей // *T-Comm — Телекоммуникации и Транспорт*. 2010. № 7. С. 36—38.
4. **Прокопьев А. В.** Разработка и исследование моделей нагрузки в беспроводных сенсорных сетях: Дис. на соиск. учен. степ. канд. тех. наук (05.12.13). СПб.: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича. 2012.
5. **Дорт-Гольц А. А.** Разработка и исследование метода балансировки трафика в пакетных сетях связи: Дис. на соиск. учен. степ. канд. тех. наук (05.12.13). СПб.: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича. 2014. 168 с.
6. **Агеев Д. В.** Синтез наложенных телекоммуникационных сетей с учетом единиц данных и применении многослойного графа // *Восточно-Европейский журнал передовых технологий*. 2013. Т. 4. № 9 (64). С. 49—51.
7. **Chen S., Colbeth M., Dinh H., Kim Y., Wang B.** Data collection with multiple sinks in wireless sensor networks // *Wireless Algorithms, Systems, and Applications*. Springer, 2009. P. 284—294.
8. **Горшков Г. С., Никитин Е. В., Саксонов Е. А.** Задача формирования структуры базовой сети // *Вестник Тихоокеанского государственного университета*. 2010. № 2 (17). С. 59—66.
9. **Корячко В. П., Перепелкин Д. А., Иванчикова М. А.** Алгоритм адаптивной маршрутизации в корпоративных сетях нескольких провайдеров связи // *Вестник РГРТУ*. 2013. № 44. С. 52—56.
10. **Дворников А. А.** Агентно-ориентированное моделирование совместимости каналов беспроводной сенсорной и наложенной сетей // *Качество, инновации, образование*. 2015. № 9 (124). С. 34—40.
11. **Корячко В. П., Перепелкин Д. А.** Разработка и исследование математической модели многопутевой адаптивной маршрутизации в сетях связи с балансировки нагрузки // *Электро-связь*. 2014. № 12. С. 27—31.
12. **Fujiwara T., Nakayama S., Iida N., Watanabe T.** A wireless network scheme enhanced with ad-hoc networking for emergency communications // *Power*. 2003. Vol. 10.
13. **Leberknight C. S., Chiang M., Poor H. V., Wong F.** A taxonomy of Internet censorship and anti-censorship. *Princeton Tech Report*, 2012. 28 p.
14. **Oliveira L. M., Rodrigues J. J.** Wireless Sensor Networks: a Survey on Environmental Monitoring // *J. Commun.* 2011. Vol. 6, N 2. P. 143—151.

A. A. Dvornikov, Teaching Assistant MIEM NRU HSE, Post-Graduate Student, e-mail: advornikov@hse.ru
L. S. Voskov, Professor MIEM NRU HSE, PhD, e-mail: lvoskov@hse.ru
E. A. Saksonov, Professor MSUTM named after K. G. Razumovskiy, PhD, e-mail: saksmiem@mail.ru
S. G. Efremov, Senior Lecturer NRU HSE, PhD, e-mail: sefremov@hse.ru

A Method for Organization of an Optimal Overlay Channel in Wireless Sensor Networks

A new method of optimal route search for an overlay channel built on top of a wireless sensor network is proposed. It allows to increase utilization of unused telecommunication and energy resources in wireless sensor networks without loss of their primary functions.

Keywords: telecommunication networks, wireless sensor networks, route search, unused telecommunication resources, unused energy resources, underloaded network, shortest path with a maximum weight problem, load balancing, overlay networks, overlay channels.

References

1. **Titov I. M.** Issledovanie modeli trafika servera dannyh po rezul'tatam izmerenij trafika mul'timedijnogo resursa, *T-Comm — Telekommunikacii i Transport*, 2011, no. 6, pp. 46—49 (in Russian).
2. **Kirichenko L. O., Radivilova T. A., Storozhenko A. V.** Algoritm preduprezhdenija peregruzki komp'yuternoj seti putem prognozirovaniya srednej dliny ocheredi, *Zbirnyk Naukovykh Prac' Harkivs'kogo Universytetu Povitrijnyh Syl*, 2007, no. 3, pp. 94—97.
3. **Malikova E. E.** Metod povyshenija propusknoj sposobnosti sistem telemetrii i monitoringa na baze besprovodnyh setej, *T-Comm — Telekommunikacii i Transport*, 2010, no. 7, pp. 36—38 (in Russian).
4. **Prokop'ev A. V.** *Razrabotka i issledovanie modelej nagruzki v besprovodnyh sensornyh setjah* (Published PhD thesis). The Bonch-Bruевич Saint — Petersburg State University of Telecommunications, Saint-Petersburg (in Russian).
5. **Dort-Gol'c A. A.** (2014). *Razrabotka i issledovanie metoda balansirovki trafika v paketnyh setjah svyazi* (Published PhD thesis). The Bonch-Bruевич Saint-Petersburg State University of Telecommunications, Saint-Petersburg, 2012, 168 p. (in Russian).
6. **Ageev D. V.** Cintez nalozhennyh tele-kommunikacionnyh setej s uchetom edinic dannyh i primeneniem mnogoslajnogogo grafa, *Vostochno-Evropejskij zhurnal peredovyh tehnologij*, 2013, vol. 4, no. 9 (64), pp. 49—51 (in Russian).
7. **Chen S., Colbeth M., Dinh H., Kim Y., Wang B.** Data collection with multiple sinks in wireless sensor networks, *Wireless Algorithms, Systems, and Applications*. Springer, 2009, pp. 284—294.
8. **Gorshkov G. S., Nikitin E. V., Saksonov E. A.** Zadacha formirovaniya struktury bazovoj seti, *Vestnik Tihookeanskogo gosudarstvennogo universiteta*, no. 2 (17), 2010, pp. 59—66 (in Russian).
9. **Korjachko V. P., Perepelkin D. A., Ivanchikova M. A.** Algoritm adaptivnoj marshrutizacii v korporativnyh setjah neskol'kih provajдеров svyazi, *Vestnik RGRU*, 2013, no. 44, pp. 52—56 (in Russian).
10. **Dvornikov A. A.** Agentno-orientirovanное modelirovanie sovmestimosti kanalov besprovodnoj sensornoj i nalozhennoj setej (An agent-oriented wireless sensor network and a superimposed network channels compatibility modeling), *Kachestvo, innovacii, obrazovanie*, 2015, no. 9 (124), pp. 34—40 (in Russian).
11. **Korjachko V. P., Perepelkin D. A.** Razrabotka i issledovanie matematicheskoj modeli mnogoputevoj adaptivnoj marshrutizacii v setjah svyazi s balansirovki nagruzki, *Jelektrosvjaz'*, 2014, no. 12, pp. 27—31 (in Russian).
12. **Fujiwara T., Nakayama S., Iida N., Watanabe T. A.** A wireless network scheme enhanced with ad-hoc networking for emergency communications. *Power*, 2003, vol. 10.
13. **Leberknight C. S., Chiang M., Poor H. V., Wong F.** A taxonomy of Internet censorship and anti-censorship, *Princeton Tech Report*, 2012, 28 p.
14. **Oliveira L. M., Rodrigues J. J.** Wireless Sensor Networks: a Survey on Environmental Monitoring, *J. Commun.*, 2011, vol. 6, no. 2. pp. 143—151.

СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ CAD-SYSTEMS

УДК 658.512.2

А. А. Яковлев, д-р техн. наук, профессор, yaa_777@mail.ru,

В. А. Камаев, д-р техн. наук, профессор, vkamaev40@mail.ru,

В. С. Сорокин, аспирант, s.o.r.o.k.i.n@mail.ru,

С. Н. Мишустина, ст. преподаватель, svt4656@mail.ru

Волгоградский государственный технический университет, г. Волгоград

Поисковое конструирование систем охлаждения на основе инженерно-физического подхода

Описана методика поискового конструирования, позволяющая путем построения модели физического принципа действия получить множество возможных технических решений систем охлаждения и определить наиболее перспективные варианты для конструктивной реализации. Выделена предметная область методики и уточнена модель физического принципа действия для систем охлаждения. Разработаны алгоритмы и определены основные структуры данных для информационной поддержки поискового конструирования.

Ключевые слова: поисковое конструирование, физический принцип действия, техническое решение, система охлаждения, рабочее тело

Введение

В данной работе рассматриваются вопросы, которые относятся к начальным стадиям проектирования систем охлаждения, включающим определение структуры будущего устройства и состава функциональных элементов для его последующей конструктивной реализации в виде чертежей. Этот наиболее творческий и ответственный этап проектирования, во многом определяющий качество будущего изделия, называют поисковым конструированием.

Главными задачами этапа являются получение как можно большего числа альтернативных вариантов системы и выбор из них наиболее эффективного для последующей реализации. Традиционный подход заключается в выборе в качестве прототипа схемы системы охлаждения и подборе для нее конструктивных элементов. Такой подбор осуществляется методом "проб и ошибок". Это связано с тем, что многие элементы оказываются функционально несовместимыми и определить такую несовместимость может только сам проектировщик в процессе мысленного моделирования. На практике это ограничивает число вариантов для сравнения и заставляет усомниться в оптимальности выбора лучшего из них, так как огромное множество возможных вариантов выпадает из рассмотрения. Поэтому теоретическая проработка методологии начальных этапов проектирования систем охлаж-

дения, позволяющая автоматизировать этот процесс, является актуальной задачей. Ее решение позволит повысить производительность и качество труда проектировщика на этапе поискового конструирования.

Постановка задачи

В работах [1–3] рассмотрены основные положения и область применения инженерно-физического метода поискового конструирования преобразователей энергии. Его главное преимущество заключается в использовании новой модели физического принципа действия (ФПД), позволяющей на основе принципов феноменологической термодинамики выявить функции конструктивных элементов. Затем путем проведения информационного поиска в научно-технической литературе, патентном фонде, специализированных базах данных и других источниках появляется возможность сформировать матрицу возможных технических решений для дальнейшей конструктивной проработки наиболее перспективных вариантов.

Как показал анализ, теоретические положения инженерно-физического метода могут быть успешно использованы для систем охлаждения, основным функциональным элементом которых также является рабочее тело. Были рассмотрены частные и общие классификации систем охлаждения,

что позволило выделить предметную область данной методики. Она распространяется на системы охлаждения с рабочим телом, находящимся в газообразном и/или жидком состоянии. Наиболее характерными представителями таких технических систем являются пароконденсационные, парожеторные, абсорбционные, газовые, воздушные холодильные установки, системы тонкого распыления, испарительные, вакуумно-испарительные, вакуумно-барботажные и многие другие системы охлаждения.

Метод решения задачи

Несмотря на широкую область применения в системах охлаждения используется небольшое количество физических явлений. К ним относятся фазовые переходы (кипение, испарение, конденсация, сублимация, плавление), дросселирование газов и жидкостей, адиабатическое расширение и расширение с совершением внешней работы, вихревой эффект, абсорбция газов и десорбция газов из растворов, эжекция.

В работах [1–3] описаны основные принципы разработки модели ФПД, основу которой составляет граф ФПД, состоящий из двух типов вершин (характерные точки и объекты окружения) и двух типов дуг (дуги-потоки и дуги-взаимодействия).

Каждому физическому явлению на графе соответствует определенная топологическая структура из характерных точек и дуг обоих типов.

Для физических явлений, лежащих в основе работы систем охлаждения, были разработаны их представления на графе ФПД. Например, охлаждение за счет кипения хладагента на графе представляется следующими узлами (рис. 1). Процесс кипения, осуществляющийся при подводе теплоты от внешнего источника, показан на рис. 1, а. На рис. 1, б показан тот же процесс без подвода теплоты. В первом случае кипение происходит при постоянной температуре, а во втором случае температура хладагента понижается, так как испарение жидкости осуществляется за счет внутренней энергии. Аналогичным образом на графе модели ФПД представляются процессы испарения, конденсации, сублимации и плавления.

На рис. 2, а показано представление процесса дросселирования газа, основанное на резком снижении давления газа при прохождении через суженное отверстие (вентиль, дроссель), а на рис. 2, б показан процесс дросселирования жидкостей. Так как температура кипения жидкости зависит от давления, то жидкость, имея определенную температуру и поступая в область низкого давления, оказывается перегретой по отношению к низкому давлению. Происходит ее бурное кипение с образованием сухого насыщенного пара, что отражается на графе показом дуги, обозначающей фазовое превращение жидкости.

Графическое представление процессов расширения газа с совершением внешней работы показано на рис. 3, а. Адиабатическое расширение газа — процесс, протекающий без теплообмена между рабочим телом (газом) и окружающей средой, представлено на рис. 3, б.

На рис. 4 показан вихревой эффект Ранка—Хильша, который создается с помощью специального устройства — вихревой трубы. В результате происходит разделение теплого и холодного воздуха в закрученном потоке.

Принципиальная схема эжектора и графическое представление процессов, происходящих в нем, показано на рис. 5. Рабочая среда (газ или пар высокого давления) выходит из сопла 1 (вершина v_1^1) и поступает в камеру смешения 2 (вершина v_2^1). Сюда же поступает из всасывающей камеры 3 подлежащая сжатию подсосываемая среда (газ или пар низкого давления). Из камеры смешения смесь

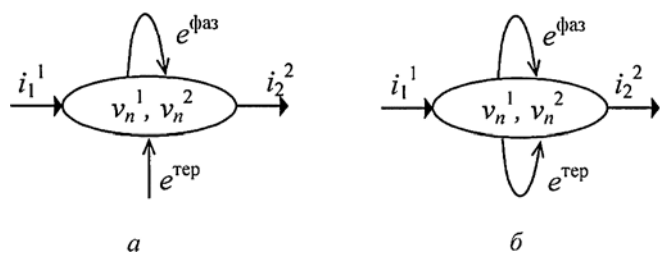


Рис. 1. Графическое представление процесса кипения:
 v_n^1, v_n^2 — хладагент в жидком и газообразном состоянии соответственно; i_1^1 — поток жидкого хладагента; i_2^2 — поток газообразного хладагента; $e^{\text{фаз}}$ — фактор экстенсивности, сопряженный с фазовой формой движения; $e^{\text{тер}}$ — фактор экстенсивности, сопряженный с термической формой движения

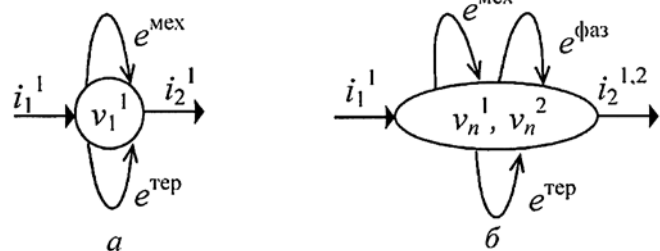


Рис. 2. Графическое представление процессов дросселирования газа и жидкостей

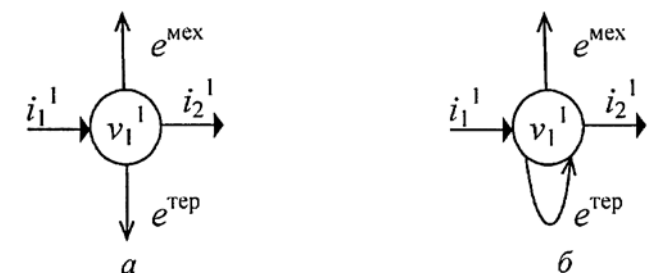


Рис. 3. Графическое представление процессов расширения газа

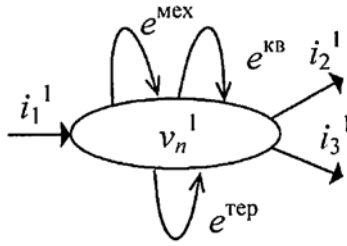


Рис. 4. Графическое представление эффекта Ранка—Хильша:

v_n^1 — газ в вихревой трубе; i_1^1 — поток сжатого газа в сопловом вводе вихревой трубы; i_2^1 — горячий поток газа; i_3^1 — холодный поток газа; $e^{\text{мех}}$ — фактор экстенсивности, сопряженный с механической формой движения; $e^{\text{тер}}$ — фактор экстенсивности, сопряженный с термической формой движения; $e^{\text{кв}}$ — фактор экстенсивности, сопряженный с кинетической (вращения) формой движения

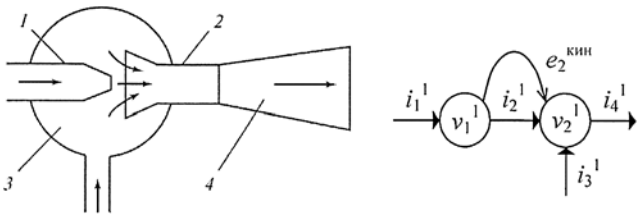


Рис. 5. Графическое представление физических процессов в эжекторе:

v_1^1 — рабочая среда в сопле эжектора; v_2^1 — рабочая и подсосываемая среда в камере смешения; i_1^1 — входной поток рабочей среды; i_2^1 — поток рабочей среды, выходящей из сопла; i_3^1 — поток подсосываемой среды; i_4^1 — выходной поток рабочей среды; $e^{\text{кин}}$ — фактор экстенсивности, сопряженный с кинетической формой движения

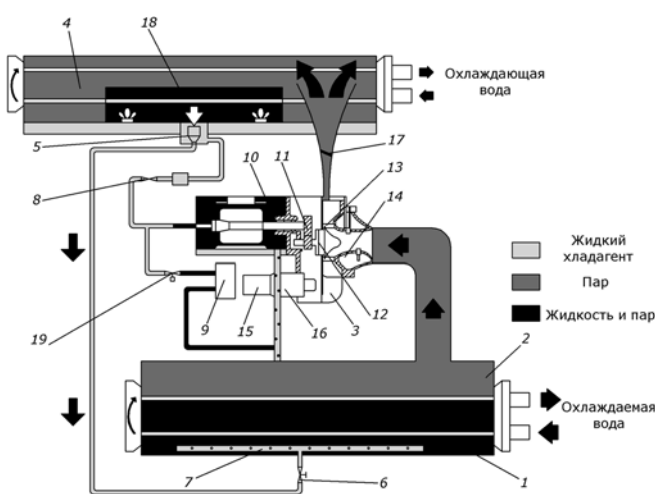


Рис. 6. Схема парокompрессионной холодильной установки:

1 — испаритель; 2 — сухопарник; 3 — радиальный компрессор; 4 — конденсатор; 5 — поплавковый клапан; 6 — регулирующий вентиль I; 7 — распределительная трубка; 8 — регулирующий вентиль II; 9 — масляный теплообменник; 10 — электродвигатель; 11 — трансмиссия; 12 — рабочее колесо; 13 — диффузор; 14 — лопатки направляющего аппарата; 15 — масляный фильтр; 16 — масляный насос; 17 — клапан конденсатора; 18 — камера; 19 — регулирующий вентиль III

этих двух сред поступает в диффузор 4, в котором кинетическая энергия струи переходит в потенциальную энергию с повышением давления.

Все взаимодействия рабочего тела с объектами окружения приводят к изменению его энергии. Однако их функции в технической системе могут существенно различаться. В системах охлаждения главными объектами окружения являются теплодатчики и теплоприемники, термическое взаимодействие с которыми приводит к повышению и понижению температуры рабочего тела.

Для осуществления этих главных процессов необходимы и другие взаимодействия. Они нужны для изменения параметров рабочего тела и его перемещения между характерными точками. По функциональному назначению объекты окружения для систем охлаждения можно разделить на шесть групп:

- теплодатчики (ТО);
- теплоприемники (ТП);
- с функцией изменения параметров рабочего тела (ПРТ);
- с функцией транспортирования рабочего тела (ТРТ);
- истоки рабочего тела (ИРТ);
- стоки рабочего тела (СРТ).

Все объекты окружения могут находиться на разных уровнях иерархии.

Такие представления физических явлений позволяют разрабатывать модели ФПД для самых разных систем охлаждения согласно положениям, изложенным в работах [1, 4, 5]. На рис. 6 в качестве примера приведена принципиальная схема парокompрессионной холодильной установки, а на рис. 7 — разработанная для нее модель ФПД. Опи-

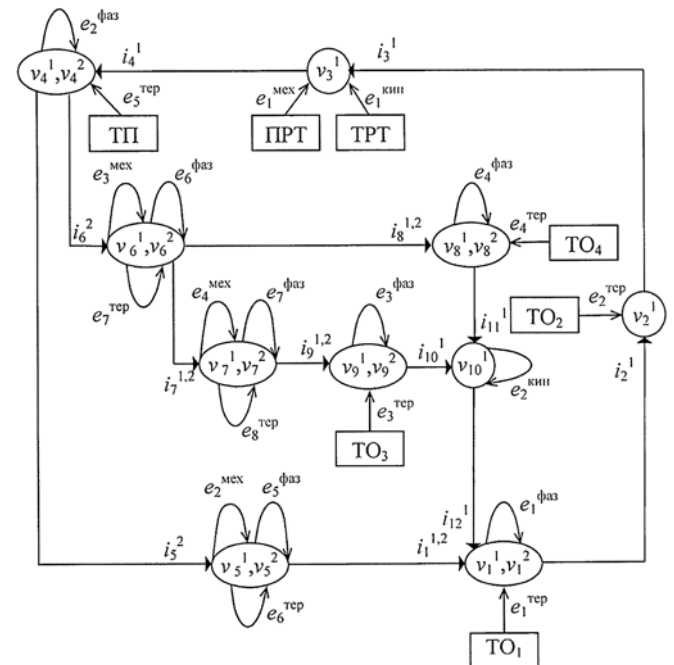


Рис. 7. Граф модели ФПД парокompрессионной холодильной установки

**Описание элементов модели ФПД
парокомпрессионной холодильной машины**

Обозначение элемента	Тип вершины или дуги	Семантическое описание
ТП	Объект окружения	Охлаждающая жидкость (теплоприемник)
ТО ₁	То же	Охлаждаемая жидкость (теплоотдатчик)
ТО ₂	—//—	Регенератор (теплоотдатчик)
ТО ₃	—//—	Масло в масляном контуре теплообменника (теплоотдатчик)
ТО ₄	—//—	Электродвигатель (теплоотдатчик)
ТРГ	—//—	Механический привод радиального насоса (транспортирование хладагента)
ПРТ	—//—	Механический привод радиального насоса (изменение параметров хладагента)
v_1^1, v_1^2	Характерная точка	Хладагент в виде пара и жидкости в испарителе
v_2^1	То же	Парообразный хладагент в сухопарнике
v_3^1	—//—	Хладагент в рабочей камере радиального компрессора
v_4^1, v_4^2	—//—	Хладагент в виде пара и жидкости в конденсаторе
v_5^1, v_5^2	—//—	Хладагент в виде пара и жидкости в регулирующем вентиле I
v_6^1, v_6^2	—//—	Жидкий хладагент в регулирующем вентиле II
v_7^1, v_7^2	—//—	Хладагент в виде пара и жидкости в регулирующем вентиле III
v_8^1, v_8^2	—//—	Хладагент в виде пара и жидкости в кожухе электродвигателя
v_9^1, v_9^2	—//—	Хладагент в виде пара и жидкости в масляном теплообменнике
v_{10}^1	—//—	Парообразный хладагент в соединении трубопроводов
$i_1^{1,2}, i_7^{1,2}, i_8^{1,2}, i_9^{1,2}$	Маршрутная дуга	Потоки хладагента в виде пара и жидкости
$i_2^1, i_3^1, i_4^1, i_{10}^1, i_{11}^1$	То же	Потоки газообразного хладагента
i_{12}^1	—//—	Потоки жидкого хладагента
i_5^2, i_6^2	—//—	Потоки жидкого хладагента
$e_1^{фаз}, e_2^{фаз}, e_3^{фаз}, e_4^{фаз}, e_5^{фаз}, e_6^{фаз}, e_7^{фаз}$	Взаимодействие	Процессы кипения и конденсации хладагента
$e_1^{тер}, e_2^{тер}, e_3^{тер}, e_4^{тер}, e_5^{тер}$	То же	Процессы теплообмена с объектами окружения
$e_6^{тер}, e_7^{тер}, e_8^{тер}$	—//—	Процессы понижения температуры хладагента при дросселировании в регулирующем вентиле
$e_1^{мех}, e_2^{мех}, e_3^{мех}, e_4^{мех}$	—//—	Процессы сжатия и расширения хладагента
$e_1^{кин}$	—//—	Процесс нагнетания парообразного хладагента в радиальном компрессоре
$e_2^{кин}$	—//—	Процесс передачи кинетической энергии между двумя потоками хладагента

сание вершин и дуг графа модели ФПД (характерных точек, объектов окружения, потоков хладагента и взаимодействий) приведены в таблице.

Аналогичные модели ФПД были разработаны для газовой криогенной машины, работающей по обратному циклу Стирлинга, турбохолодильной машины, парожетторной и бромид-литиевой абсорбционной холодильных установок, а также многих других систем охлаждения, что продемонстрировало продуктивность инженерно-физического подхода для рассматриваемого класса технических систем.

Последовательность построения моделей ФПД показана в виде блок-схемы на рис. 8. Она состоит из двух последовательно выполняемых циклов с постусловием, охваченных структурой "альтернатива с одним действием". Наличие циклов с постусловием предполагает, что каждая процедура, входящая в тело данного цикла, должна быть выполнена хотя бы один раз. Данная структура методики полностью соответствует принципам структурного программирования, что обуславливает возможность ее реализации в виде интерактивной информационно-поисковой системы принятия решений в рамках САПР систем охлаждения.

Структурирование информации позволяет перейти к этапу автоматизации отдельных процедур при построении моделей ФПД систем охлаждения. При этом информация фиксируется в виде таблиц, что позволяет использовать хорошо зарекомендовавший себя инструментарий реляционных СУБД для поддержки программных приложений.

Для каждой характерной точки графа ФПД рассматриваемой системы охлаждения делаются записи в таблице A1 со структурой

$$S_{A1} = \{a_1^1, a_1^2, a_1^3, a_1^4, a_1^5\}, \quad (1)$$

где a_1^1 — порядковый номер кортежа таблицы; a_1^2 — обозначение характерной точки графа ФПД; a_1^3 — семантическое описание вершины; a_1^4 — описание состояний хладагента в характерной точке; a_1^5 — обозначение унитарной системы охлаждения.

Для каждого истока и стока рабочего тела создаются записи в таблице A2 со структурой

$$S_{A2} = \{a_2^1, a_2^2, a_2^3, a_2^4, a_2^5\}, \quad (2)$$

где a_2^1 — порядковый номер кортежа таблицы; a_2^2 — обозначение объекта окружения; a_2^3 — тип объекта окружения; a_2^4 — семантическое описание объекта окружения; a_2^5 — обозначение унитарной системы охлаждения.

Выявляются потоки хладагента, связывающие вершины графа ФПД, и для каждого потока делаются записи в таблице $A3$ со структурой

$$S_{A3} = \{a_3^1, a_3^2, a_3^3, a_3^4\}, \quad (3)$$

где a_3^1 — порядковый номер кортежа таблицы; a_3^2 — обозначение дуги на графе; a_3^3 — семантическое описание дуги; a_3^4 — обозначение унитарной системы охлаждения.

Описание взаимодействий, выполняемых характерных точках унитарной системы охлаждения, осуществляется в формуляре таблицы $A4$ со структурой

$$S_{A4} = \{a_4^1, a_4^2, a_4^3, a_4^4, a_4^5, a_4^6\}, \quad (4)$$

где a_4^1 — порядковый номер кортежа таблицы; a_4^2 — обозначение фактора экстенсивности; a_4^3 — обозначение объекта окружения или характерной точки, являющегося источником фактора экстенсивности; a_4^4 — обозначение объекта окружения или характерной точки, являющегося приемником фактора экстенсивности; a_4^5 — семантическое описание дуги; a_4^6 — обозначение унитарной системы охлаждения.

С каждым элементом графа модели ФПД связаны элементарные функции, смысловое значение которых вытекает из принципов феноменологической термодинамики. Перечень этих функций и их логическое обоснование подробно описаны в работах [1, 2, 5].

Последовательность выполнения процедур синтеза технических решений приведена в виде блок-схемы на рис. 9. Она состоит из восьми процедур, вызов которых определяется двумя условиями. После выполнения двух процедур проверяется условие продолжения (найден вариант хладагента). Далее выполняется тело цикла, которое состоит из

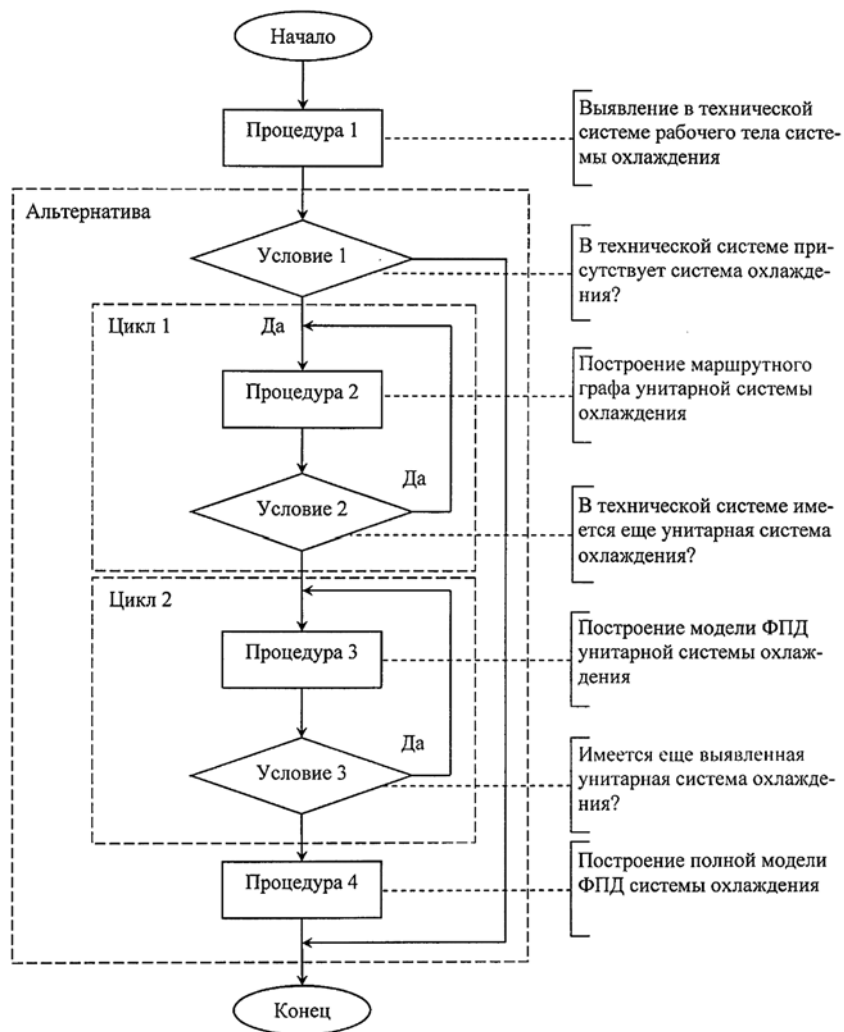


Рис. 8. Последовательность выполнения процедур методики построения модели ФПД

последовательного вызова процедур 3—8. Продолжение цикла определяется условием 2. Данная структура методики также соответствует принципам структурного программирования. Она включает два следования, альтернативу и цикл. Каждая из этих структур имеет одну точку входа и одну точку выхода.

В процессе выполнения процедуры определения элементарных функций, связанных с вершинами и дугами графа ФПД, заполняется таблица $M1$ со структурой

$$S_{M1} = \{m_1^1, m_1^2, m_1^3, m_1^4, m_1^5\}, \quad (5)$$

где m_1^1 — обозначение элемента графа ФПД (вершины и/или дуги); m_1^2 — тип элементарной функции; m_1^3 — вербальное описание элементарной функции для данного элемента графа; m_1^4 — иден-

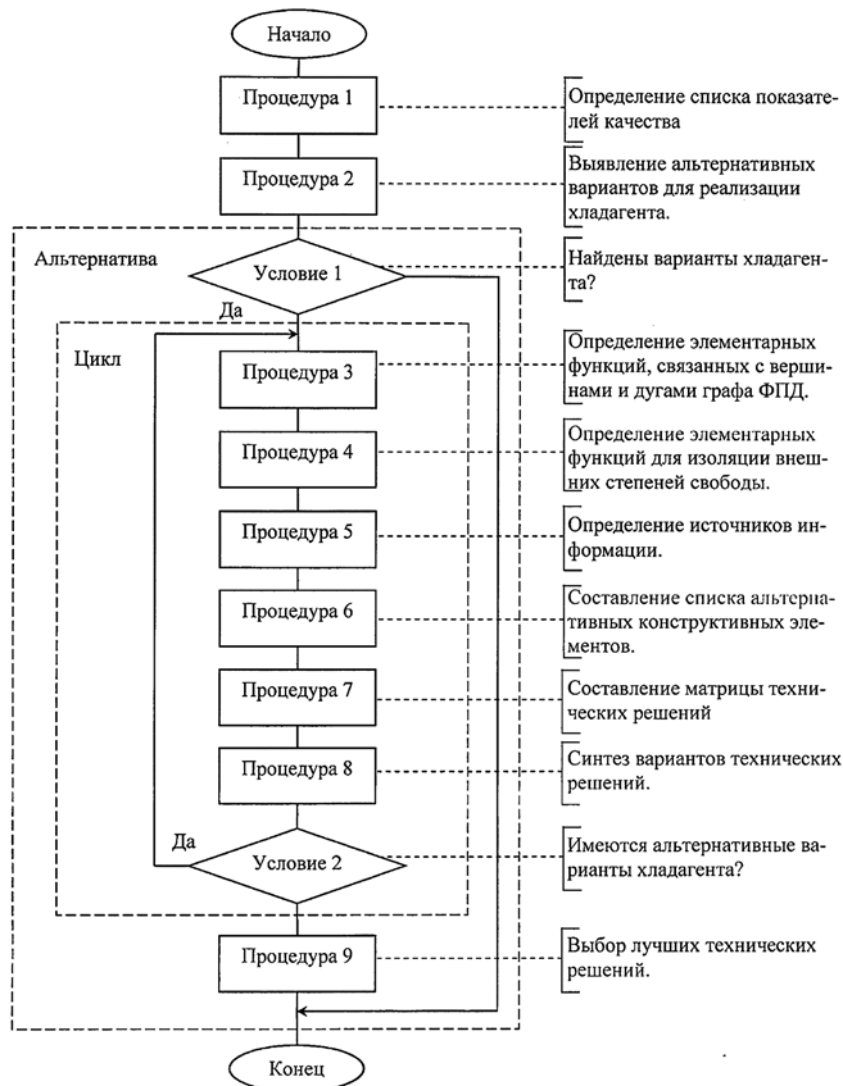


Рис. 9. Последовательность выполнения процедур методики синтеза технических решений

идентификатор элементарной функции; m_1^5 — необходимость конструктивной реализации.

Далее определяются нежелательные взаимодействия и определяются элементарные функции для их исключения. По результатам анализа заполняется таблица $M2$ со структурой

$$S_{M2} = \{m_2^1, m_2^2, m_2^3, m_2^4, m_2^5, m_2^6\}, \quad (6)$$

где m_2^1 — обозначение элемента графа ФПД (вершины или дуги), с которым связано нежелательное взаимодействие; m_2^2 — обозначение и вербальное описание нежелательного взаимодействия; m_2^3 — обозначение элементарной функции; m_2^4 — вербальное описание элементарной функции; m_2^5 —

идентификатор элементарной функции; m_2^6 — необходимость конструктивной реализации.

Для поиска альтернативных конструктивных элементов, реализующих элементарные функции, определяются источники информации, которые фиксируются в таблице $M3$ со структурой

$$S_{M3} = \{m_3^1, m_3^2, m_3^3, m_3^4, m_3^5\}, \quad (7)$$

где m_3^1 — идентификатор типа конструктивного элемента; m_3^2 — наименование конструктивного элемента; m_3^3 — перечень элементарных функций, выполняемых элементом; m_3^4 — классы Международной патентной классификации (МПК); m_3^5 — другие источники информации, в которых содержатся описания конструктивных элементов, выполняющих подобные функции.

По результатам поиска в источниках информации составляется список конструктивных элементов в таблице $M4$ со структурой

$$S_{M4} = \{m_4^1, m_4^2, m_4^3, m_4^4, m_4^5, m_4^6\}, \quad (8)$$

где m_4^1 — обозначение конструктивного элемента; m_4^2 — наименование конструктивного элемента; m_4^3 — источник информации о данном конструктивном элементе, где дается его описание; m_4^4 — набор элементарных функций, выполняемых данным элементом; m_4^5 — перечень показателей качества системы охлаждения, на которые влияет данный элемент; m_4^6 — экспертные оценки элемента по каждому показателю качества.

Техническое решение системы охлаждения представляет собой список конструктивных элементов, сформированный так, чтобы входящие в него элементы выполняли все элементарные функции, сформулированные в таблицах $M1$ и $M2$. Для этого составляется матрица технических решений, последовательность разработки и пример которой подробно описаны в работах [2, 3]. Практика составления и использования таких матриц показала, что во многих случаях конструктивные элементы выполняют сразу несколько элементарных функций,

которые накладываются друг на друга. Это обуславливает несовместимость таких конструктивных элементов в одном техническом решении. Составление полного списка возможных технических решений, состоящих из совместимых элементов, осуществляется с помощью программного приложения.

Алгоритм синтеза технических решений и методика экспертной оценки уровня качества получаемых технических решений описаны в работе [5].

Оценка качества проектируемого изделия осуществляется в следующих целях:

- обоснования прогнозируемого уровня качества изделия;
- обоснования технических требований к изделию;
- выбора оптимального варианта для запуска в производство;
- достижения и превышения технического уровня проектируемых изделий.

Базовые показатели качества определяются на основе анализа известных аналогов проектируемой системы охлаждения с учетом перспектив технического прогресса и требований потребителя. В качестве базовых показателей могут быть приняты следующие:

- показатели существующих лучших образцов;
- планируемые и прогнозируемые показатели (если они обоснованы);
- рекомендации международных организаций.

При проектировании новых образцов систем охлаждения, когда аналог разрабатываемого изделия отсутствует, для определения базовых показателей используют всю доступную информацию вплоть до перспективных показателей гипотетического образца, обоснованных результатами научных исследований. Показатели качества вырабатывает ведущий конструктор с привлечением необходимых специалистов, а затем их номенклатуру рассматривает экспертная комиссия, которая принимает окончательное решение. Значения принятых показателей качества определяются при сопоставлении аналогичных базовых и перспективных показателей.

Коэффициенты весомости показателей определяются экспертным методом, который предполагает создание экспертной комиссии из высококвалифицированных специалистов (не менее семи человек), уровень компетентности которых в оценке качества рассматриваемого объекта примерно одинаков.

Заключение и выводы

Существует взгляд, что новые методы нужно разрабатывать для получения новых более прогрессивных технических решений. Однако, на наш взгляд, не менее важен и аспект их использования для обучения нового поколения проектировщиков [6, 7]. Необходимы методы, обоснование которых строится на базе современных физических представлений и вытекающих из них законов строения

технических систем. Это позволяет проектировщику лучше понимать особенности проектирования новых систем и правильно выбирать пути их конструктивной трансформации для повышения потребительских характеристик проектируемых изделий. В результате проведенного исследования можно сделать следующие выводы.

1. Предлагаемая методика позволяет сформировать матрицу технических решений систем охлаждения с газообразным и жидким рабочим телом на основе модели физического принципа действия. Главным ее преимуществом по сравнению с морфологическими таблицами является то, что получаемые технические решения не содержат функционально несовместимых элементов.

2. Описанный подход позволяет не только существенно повысить производительность труда проектировщика и открыть широкие перспективы автоматизации данного, наиболее творческого этапа проектирования, но и повышает вероятность получения нового технического решения с форсированным уровнем качества.

3. Новая методика поискового конструирования систем охлаждения через построение моделей физического принципа действия позволяет по-новому обучать процессу проектирования. С ее помощью можно не только изучать уже существующие схемы систем охлаждения, но и понимать принципы их создания и способы трансформации технических систем для получения новых, более прогрессивных технических решений.

Работа выполнена при финансовой поддержке со стороны Минобрнауки России в рамках базовой части проекта 2586 задания № 2014/16, проекта № 2.1917.2014К_2014, а также при финансовой поддержке со стороны РФФИ, проект № 16-47-340229 p_a.

Список литературы

1. Яковлев А. А. Разработка множеств технических решений установок для преобразования энергии. М.: Машиностроение-1, 2007. 128 с.
2. Yakovlev A. A., Chursina S. V., Pozdnyakova P. E., Sorokin V. S. Energy converters with gaseous and liquid working body searching design // World Applied Sciences Journal (WASJ). 2013. Vol. 24, Spec. iss. 24: Information Technologies in Modern Industry, Education & Society. P. 213–219.
3. Камаев В. А., Яковлев А. А. Моделирование физических принципов действия и формирование множеств технических решений преобразователей энергии // Информационные технологии. 2006. № 1. С. 2–8.
4. Коробкин Д. М., Фоменков С. А., Колесников С. Г. Автоматизация процесса формирования информационного обеспечения базы данных физических эффектов // Вестник компьютерных и информационных технологий. 2005. № 3. С. 22–25.
5. Фоменков С. А., Колесников С. Г. Информационное наполнение баз данных по физическим эффектам // Информационные технологии. 2004. № 6. С. 60–62.
6. Воинов Б. С., Бугров В. Н., Воинов Б. Б. Информационные технологии и системы: поиск оптимальных, оригинальных и рациональных решений. Второе доп. науч. изд. В 2 томах. Т. 1. Методология синтеза новых решений. М.: Наука, 2007. 480 с.
7. Kravets A. G., Gurtjakov A. S., Darmanian A. P. Enterprise intellectual capital management by social learning environment implementation // World Applied Sciences Journal. 2013. Vol. 23, N 7. P. 956–964.

Search Engine Design Systems Cooling Based Engineering and Physical Approach

The technique of search engine design that allows by constructing a model of the physical principle of action to get a lot of cart-sible technical solutions cooling systems and identify the most promising options for constructive realization. Obtained data domain techniques and refined model of the physical principle of action for the cooling systems. The algorithms and the basic data structures for information support search design.

Keywords: tags, search engine design, the physical principle of action, the technical solution, the cooling system, the working fluid

References

1. **Yakovlev A. A.** *Razrabotka mnozhestv tehniceskikh reshenij ustanovok dlja preobrazovanija jenerгии* (Development of technical solutions sets of installations for energy conversion), Moscow, Mashinostroenie-1, 2007, 128 p.
2. **Yakovlev A. A., Chursina S. V., Pozdnyakova P. E., Sorokin V. S.** Energy converters with gaseous and liquid working body searching design, *World Applied Sciences Journal (WASJ)*, 2013, vol. 24, spec. iss. 24: Information Technologies in Modern Industry, Education & Society, pp. 213–219.
3. **Kamaev V. A., Yakovlev A. A.** Modelirovanie fizicheskikh principov dejstvija i formirovanie mnozhestv tehniceskikh reshenij preobrazovatelej jenerгии (Modelling of physical principles of operation and formation of sets of technical solutions energy converters), *Informacionnye tehnologii*, 2006, no. 1, pp. 2–8.
4. **Korobkin D. M., Fomenkov S. A., Kolesnikov S. G.** Avtomatizacija processa formirovanija informacionnogo obespechenija bazy dannyh fizicheskikh jeffektov (The formative process automation is information provision database of physical effects), *Vestnik komp'uternyh i informacionnyh tehnologij*, 2005, no. 3, pp. 22–25.
5. **Fomenkov S. A., Kolesnikov S. G.** Infoemacionnoe napolnenie baz dannyh po fizicheskim jeffektam (Information content database of physical effects), *Informacionnye tehnologii*, 2004, no. 6, pp. 60–62.
6. **Voinov B. S., Bugrov V. N., Voinov B. B.** Informacionnye tehnologii i sistemy: poisk optimal'nyh, original'nyh i racional'nyh reshenij (Information technology and system: the search for optimal, original and rational solutions). *Vtoroe dopolnennoe nauchnoe izdanie*, vol. 1, Metodologija sinteza novyh reshenij, Moscow, Nauka, 2007, 480 p.
7. **Kravets A. G., Gurtjakov A. S., Darmanian A. P.** Enterprise intellectual capital management by social learning environment implementation, *World Applied Sciences Journal*, 2013, vol. 23, no. 7, pp. 956–964.

4-я Международная Научно-практическая конференция ИНСТРУМЕНТЫ И МЕТОДЫ АНАЛИЗА ПРОГРАММ, ТМРА—2017, Москва, 3–4 Марта 2017

Приглашаем студентов старших курсов, аспирантов, молодых ученых, преподавателей, научных сотрудников и сотрудников технологических компаний, чьи научные и практические интересы связаны с разработкой и анализом программного обеспечения, принять участие в четвертой научно-практической конференции Tools & Methods of Program Analysis/Инструменты и методы анализа программ.

Вопросы эффективности и корректности функционирования программного обеспечения являются ключевыми для большинства наукоемких отраслей современной экономики. Четвертая международная научно-практическая конференция Tools & Methods of Program Analysis нацелена на развитие индустрии разработки программного обеспечения и внедрение новейших разработок в области тестирования, анализа и верификации программ.

На конференции будут представлены приглашенные доклады признанных специалистов в области программной инженерии и конкурсные доклады, прошедшие рецензирование несколькими независимыми экспертами.

Темы, рассматриваемые на конференции, включают (но не ограничиваются):

- автоматизация тестирования программного обеспечения
- статический анализ программ
- верификация
- динамические методы анализа программ
- тестирование и анализ параллельных и распределенных систем
- тестирование и анализ высоконагруженных систем и систем высокой доступности
- анализ и верификация программно-аппаратных систем
- методы создания качественного программного обеспечения
- инструментальные средства анализа, тестирования и верификации

Ключевые даты

- 9 декабря — окончание приёма аннотаций к докладам;
- 16 декабря — окончание приёма статей;
- 27 января — публикация списка принятых работ;
- 3–4 марта — конференция ТМРА—2017.

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ИЗОБРАЖЕНИЙ DIGITAL PROCESSING OF SIGNALS AND IMAGES

УДК 004.272.26, 004.932.72'1

Е. В. Бимаков, ведущий инженер, robint@mail.ru,
ООО "Воксельная техника", г. Ижевск

Метод виртуального зондирования сцены на базе воксельной технологии обработки сцен

Описан скоростной метод виртуального зондирования сцены на базе специальной воксельной аппаратно-программной технологии обработки сцен, предназначенный для решения задач навигации автономного мобильного робота (автоматический обход препятствий и распознавание 2D- и 3D-объектов в режиме реального времени). Изложена общая концепция метода, обоснован выбор машинных моделей исследуемой сцены и исследовательского зонда, тип применяемого вычислителя для реализации описываемого метода зондирования с предельной степенью параллелизации.

Ключевые слова: обработка сцен и изображений, навигация, автономный наземный робот, обход препятствий, распознавание объектов, воксельные геометрические модели, воксельные вычислители

Введение

Использование физического тела (исследовательского зонда), обладающего известными геометрическими и физическими свойствами, для исследования 3D-сцен хорошо известно. Переход от реального зондирования сцены к виртуальному зондированию сцены (ВЗС) сопровождается заменой реального зонда и реальной зондируемой сцены их виртуальными аналогами. Уже сейчас ВЗС широко применяется для поиска обхода препятствий в компьютерных играх и для управления мобильными роботами [1, 2]. При этом эффективность машинной реализации этого метода определяется, главным образом, скоростью перемещения зонда (геометрической модели робота) в виртуальном пространстве исследуемой сцены, а также скоростью определения факта столкновения виртуального зонда с виртуальными моделями препятствий (т.е. факта пересечения геометрической модели зонда с геометрическими моделями препятствий сцены).

В безлюдных технологиях будущего время анализа сложных реальных сцен автономным мобильным роботом (АМР) должно быть сопоставимо со временем анализа этих сцен человеком (доли секунды) или даже быть меньше. Таких бортовых навигационных видеоанализаторов сейчас нет. Попытки их создания [3–6] на основе существующих программируемых вычислителей натолкнулись на сильнейшую зависимость времени автоматического анализа сцен от их геометрической сложности.

Попытки применения современных искусственных нейронных сетей (т.е. непрограммируемых самообучающихся устройств) [7, 8] для управления АМР внутри сложных реальных сцен крайне далеки от успеха.

В данной работе предлагается метод ВЗС, основанный на применении нового типа программируемых вычислителей, названных нами воксельными вычислителями (ВВ) [9–14]. Для ВВ характерно сочетание преимуществ как нейронных сетей (высокое распараллеливание процесса обработки), так и традиционных программируемых устройств (возможность использования богатого, наработанного годами математического и алгоритмического аппарата).

К настоящему моменту разработана теория проектирования ВВ, базовое программное обеспечение ВВ, методы и алгоритмы его использования, изготовлен и испытан макет ВВ, начаты работы по созданию макета системы управления роботом на базе ВВ [15], которые вследствие отсутствия финансирования приостановлены.

1. Виртуальное зондирование сцены. Основные определения

Определим ряд терминов, используемых в дальнейшем при изложении предлагаемого метода виртуального зондирования сцены.

Определение 1.1. Назовем n -мерной сценой S формула $F_S^{(n)}$ семейство $S = \{E_q\}_{q \in I}$ (где $I = 1, 2, 3, \dots, q_s$)

точечных множеств E_q евклидового пространства R^n , удовлетворяющее выражению:

$$(\forall E_q \in S) [(E_q \subseteq F_S^{(n)}) \wedge (\mu^{(n)}(E_q) \neq 0)],$$

где $\mu^{(n)}(E_q)$ — n -мерная мера множества E_q . Понятие n -мерной меры (n -мерного объема) $\mu^{(n)}(E^{(n)})$ является естественным обобщением понятий: длины $\mu^{(1)}(E^{(1)})$ отрезка $E^{(1)}$ — для R^1 ; площади $\mu^{(2)}(E^{(2)})$ плоской фигуры $E^{(2)}$ — для R^2 ; объема $\mu^{(3)}(E^{(3)})$ трехмерной фигуры $E^{(3)}$ — для R^3 .

С множеством $F_S^{(n)}$ свяжем систему координат исследуемой сцены S , например, декартову систему координат $O_S X_S Y_S Z_S$ (рис. 1, а).

Определение 1.2. Множество $E_q \in S$ назовем q -м объектом n -мерной сцены S .

Определение 1.3. Обобщенным объектом сцены $S = \{E_q\}_{q \in I}$ назовем объединение всех объектов сцены S , т.е. объект $E_S = \bigcup_{q=1}^{q_S} E_q$.

Определение 1.4. Сцену P формата $F_P^{(n)}$ (рис. 1, б), у которой геометрическая форма обобщенного объекта E_P и положение центра O_P системы координат $O_P X_P Y_P Z_P$ удовлетворяют условиям $f(E_P) \subseteq F_P^{(n)} \subset F_S^{(n)}$; $O_P \equiv T_m \in F_S^{(n)}$, где $f(E_P)$ — образ обобщенного объекта E_P при произвольном пространственном повороте f (т.е. биективном отображении $f: R^n \rightarrow R^n$) множества E_P вокруг O_P , назовем сценой P , зондирующей сцену S в точке T .

Определение 1.5. Зондом исследуемой сцены S назовем обобщенный объект E_P сцены P , зондирующей сцену S .

Определение 1.6. Исследуемым (зондируемым) объектом E_S назовем обобщенный объект исследуемой (зондируемой) сцены S .

Определение 1.7. Виртуальным зондированием n -мерной сцены S назовем вычислительную процедуру, заключающуюся в последовательном выполнении одного или нескольких элементарных актов зондирования (ЭАЗ). Каждый m -й ЭАЗ (где m — номер текущего акта зондирования) состоит в общем случае из трех последовательно выполняемых операций:

- 1) поворота f_m зонда E_P вокруг центра O_P формата $F_P^{(n)}$ зондирующей сцены P (рис. 1, б) на m -й n -мерный пространственный угол ($E_P \rightarrow f_m(E_P)$);
- 2) параллельного переноса c_m зонда $f_m(E_P)$ и F_P (т.е. зондирующей сцены) внутрь зондируемой сцены S ($f_m(E_P) \rightarrow c_m(f_m(E_P))$, $F_P^{(n)} \rightarrow c_m(F_P^{(n)})$) в точку m -го центра зондирования ($T_m \in F_S^{(n)}$) так, чтобы центр $c_m(O_P)$ зонда $c_m(f_m(E_P))$ совпал с точкой T_m формата $F_S^{(n)}$ исследуемой сцены S (рис. 1, з). После выполнения переноса будет установлено взаимнооднозначное соответствие между точками форматов $c_m(F_P^{(n)})$ и $F_S^{(n)}$ (а также между точками $c_m(f_m(E_P))$ и E_S), имеющими одинаковые значения координат в системе $O_S X_S Y_S Z_S$;

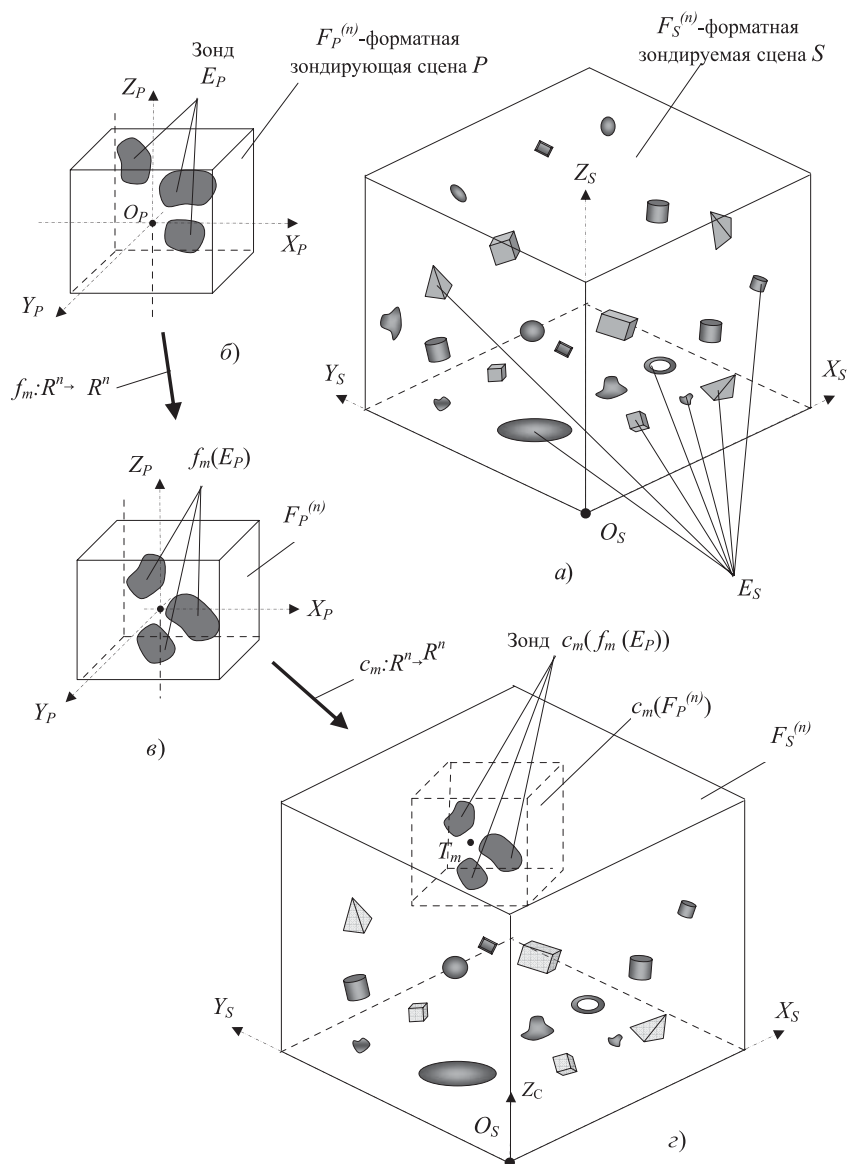


Рис. 1. Виртуальное зондирование сцены S зондирующей сценой P

3) определения значения функции зондирования $F_m(c_m(f_m(E_P)), F_S^{(n)})$ между точками зонда $c_m(f_m(E_P))$ и точками $F_S^{(n)}$ исследуемой сцены S , соответствующими точкам зонда $c_m(f_m(E_P))$.

Физический смысл виртуального зонда E_P , виртуальной сцены S и точечного множества $G \subset F_S^{(n)}$, образованного центрами зондирования сцены S , определяется характером решаемой прикладной задачи. Так, применительно к задачам поиска обхода препятствий E_P может выступать в качестве геометрической модели АМР, а объект E_S — в качестве геометрической модели обобщенного объекта-препятствия сцены S , который предстоит обойти АМР. Множество G (траектория зондирования) в этом случае — непрерывная 3D-линия, воспроизводящая, в частности, траекторию движения АМР, у которой соседние центры зондирования бесконечно близко расположены друг к другу. При решении задач обхода препятствий каждый ЭАЗ служит для выявления факта пересечения (т.е. столкновения) зонда $c_m(f_m(E_P))$ с препятствием E_S . Для этого в каждом m -м ЭАЗ выполняется теоретико-множественная операция $M_m = c_m(f_m(E_P)) \cap E_S$, после чего определяется значение функции зондирования — предиката " $M_m = \emptyset$ ". Если значение предиката равняется лог. 1, зонд $c_m(f_m(E_P))$ и объект-препятствие E_S объявляются виртуально нестолкнувшимися, в противном случае (лог. 0) — зонд $c_m(f_m(E_P))$ и препятствие E_S сцены объявляются виртуально столкнувшимися.

При решении задач распознавания виртуальный зонд E_P выступает уже в качестве n -мерного геометрического эталона распознаваемого объекта. Исследуемая $F_S^{(n)}$ -форматная сцена выступает в роли n -мерного изображения-препарата, внутри которого ищется заданный распознаваемый n -мерный объект. Виртуальное зондирование служит в этом случае для выявления факта наличия (или отсутствия) совпадения определенных геометрических свойств эталона с содержимым анализируемой сцены, а также для определения местоположения обнаруженного совпадения на исследуемой сцене. В идеале область G здесь также состоит из бесконечного числа центров зондирования, хотя порядок зондирования в данном случае несущественен. Функция зондирования, по значению которой определяется факт совпадения геометрических характеристик зонда и сцены, вычисляется здесь более сложно, чем при решении задачи обхода препятствий, и не ограничивается лишь анализом $c_m(f_m(E_P))$ и E_S на их пересекательность (в данной статье этот аспект подробно не рассматривается).

2. Машинная реализация виртуального зондирования сцены

Вышеописанная процедура ВЗС — это всего лишь идеальная версия данной процедуры, оперирующая бесконечными точечными множествами пространства R^n , в которой виртуальный зонд перемещается по непрерывной траектории зондирования, состоящей из бесконечного числа центров зондирования. Эта идеальная версия ВЗС дает самое общее (концептуальное) представление о процедуре виртуального зондирования сцены, но она не предназначена для машинной реализации. Для выполнения ВЗС с помощью программируемого устройства цифровой обработки (УЦО) идеальная версия ВЗС должна быть модифицирована следующим образом:

- точечные множества E_S и E_P должны быть представлены в УЦО в виде упорядоченного массива адресуемых информационных единиц памяти УЦО;
 - бесконечное множество G центров зондирования заменено дискретной траекторией зондирования (ДТЗ), состоящей из конечного числа центров зондирования ($T_0, T_1, T_2, \dots, T_m, T_{m+1}, \dots, T_k$).
- Переход от идеальной версии ВЗС к машинной версии ВЗС неизбежно приведет к потере точности результата зондирования сцены. Существенное влияние на эту погрешность зондирования окажет выбор способа машинного моделирования содержимого исследуемой сцены (E_S).

Предлагаемая машинная версия ВЗС использует новую аппаратно-программную технологию обработки сцен (АПТОС), опирающуюся на применение:

- специальных воксельных вычислителей (ВВ);
- дискретных (пиксельных и воксельных) геометрических моделей, представленных в ВВ в виде адресуемых 2D- и 3D-бинарных изображений низкого разрешения;
- высокопараллельных методов обработки 2D- и 3D-изображений среднего и высокого разрешения, ориентированных на возможности ВВ.

3. Воксельная n -мерная модель содержимого исследуемой сцены

После проведения АМР видеомониторинга сцены, окружающей АМР, форма E_S может иметь бесконечное число вариаций (от совершенно пустого пространства сцены, когда $E_S = \emptyset$, до ее полного заполнения, когда $E_S = F_S^{(n)}$). Отобразить в памяти УЦО робота все это бесконечное разнообразие форм содержимого сцены, окружающей робота, посредством конечного числа двоичных накопителей УЦО можно только за счет снижения точности представления формы E_S в УЦО. Например, когда каждая точка E_S будет представлена в машинной модели E_S не одной точкой, а целым ограничен-

ным односвязным бесконечным множеством точек R^n — nD -дискретом.

Разобьем пространство R^n на классы точечных множеств $v_l^{(n)} \subset R^n$ (где $l \in L, L = 1, 2, \dots, \infty$), для которых одновременно справедливы выражения

$$(\forall l \in L) (\forall a, b \in v_l^{(n)}) [\rho(a, b) \leq \xi_p],$$

где $\rho(a, b)$ — расстояние между точками a и b множества $v_l^{(n)}$, ξ_p — постоянная, и

$$\begin{aligned} & (\forall l \in L) [\mu^{(n)}(v_l^{(n)}) \neq 0]; (\forall i, j \in L) [(i \neq j) \Rightarrow \\ & \Rightarrow (v_i^{(n)} \cap v_j^{(n)} = \emptyset)]; \bigcup_{l \in L} v_l^{(n)} = R^n. \end{aligned}$$

Определение 3.1. Множества $v_l^{(n)} \subset R^n$, одновременно удовлетворяющие всем перечисленным требованиям, назовем *n -мерными вокселями $v^{(n)}$ -воксельного разбиения R^n (или $v^{(n)}$ -вокселями)*, а ξ_p — постоянной этого разбиения. Двумерные воксели будем называть пикселями.

Определение 3.2. Точечное множество $\bigcup_{l \in L_k} v_l^{(n)}$, составленное из конечного ($L_k \subset L$) числа вокселей $v_l^{(n)}$, назовем *$v^{(n)}$ -воксельным объектом*.

Определение 3.3. $v^{(n)}$ -воксельной моделью множества $\dot{E} \subset R^n$ назовем $v^{(n)}$ -воксельный объект $E = \bigcup_{l \in L_E} v_l^{(n)}$ (где $L_E \subset L$), составленный из всех вокселей $v_l^{(n)}$, для которых выполняется: $(\forall l \in L_E) [\mu^{(n)}(\dot{E} \cap v_l^{(n)}) / \mu^{(n)}(v_l^{(n)}) \geq \delta_l]$, где δ_l — постоянная ($0 < \delta_l \leq 1$). Двумерные воксельные модели будем называть пиксельными моделями.

Наглядно $v^{(n)}$ -воксельную модель E_S исследуемой сцены S можно представить в виде физической "песочной" модели, образованной из отдельных кусков-фрагментов, состоящих либо из одной "песчинки" (аналога $v^{(n)}$ -вокселя), либо из множества таких "песчинок", склеенных между собой и расположенных внутри "песчицы" $F_S^{(n)}$. Максимально возможный "объем песка" в $F_S^{(n)}$, из которого строится $v^{(n)}$ -воксельная модель E_S , характеризуется максимальным объемом памяти УЦО, отведенный для хранения самых геометрически сложных $v^{(n)}$ -воксельных моделей E_S . За критерий степени сложности геометрической формы $v^{(n)}$ -воксельной модели можно принять число *поверхностных* вокселей $v_l^{(n)}$ этой модели (т.е. вокселей, которые сами принадлежат данной модели, а их соседние воксели

в $v^{(n)}$ -воксельной модели $F_S^{(n)}$ не принадлежат). Другими словами, чем более развита внешняя поверхность $v^{(n)}$ -воксельной модели, тем геометрическая форма данной $v^{(n)}$ -воксельной модели E_S будет более сложной. У воксельных моделей E_S со сложной геометрической формой число вокселей равняется примерно половине вокселей $v^{(n)}$ -воксельной модели $F_S^{(n)}$.

Для ускорения машинной реализации процедуры зондирования $v^{(n)}$ -воксельной модели содержимого сцены S примем следующие упрощения:

а) все $v^{(n)}$ -воксели $v^{(n)}$ -воксельных моделей содержимого сцен P и S будут иметь одинаковую геометрическую форму, одинаковую угловую ориентацию в R^n и центр, расположенный внутри каждого $v^{(n)}$ -вокселя одинаковым образом;

б) центры всех $v^{(n)}$ -вокселей воксельных моделей P и S расположим в узлах регулярной n -мерной решетки с шагом d ;

в) переносы зондирующей сцены P и зондируемой сцены S будем выполнять только кратными d ;

г) $v^{(n)}$ -воксель зондирующей сцены P и $v^{(n)}$ -воксель зондируемой сцены S будем считать взаимно пересекающимися, если их центры совпадают;

д) поворот ($f: R^n \rightarrow R^n$) $v^{(n)}$ -вокселя модели E_P на произвольный пространственный угол будем моделировать путем последовательного выполнения: 1) поворота $v_P^{(n)} \rightarrow f(v_P^{(n)})$; 2) выделения на неподвижной сцене $v^{(n)}$ -воксельной модели $F_S^{(n)}$ тех ее $v^{(n)}$ -вокселей, $\mu^{(n)}$ -мерная мера области пересечения которых с $f(v_P^{(n)})$ не меньше некоторого заданного значения σ .

Машинное описание $v^{(n)}$ -воксельной модели E_S (или E_P) будем осуществлять с помощью N_S (N_P) бит памяти УЦО, где N_S (N_P) — число вокселей в $v^{(n)}$ -воксельной модели $F_S^{(n)}$ ($F_P^{(n)}$). При совпадении центра вокселя описываемой модели E_S с i -м центром вокселя $F_S^{(n)}$ i -й бит N_S -разрядного двоичного слова будет равен 1, при несовпадении — 0.

4. Факторы, влияющие на скорость виртуального зондирования сцены

В мобильных робототехнических приложениях время выполнения анализа сцены является чрезвычайно критичным параметром, часто превышающим по своей значимости даже точность выполненного анализа (в отличие, например, от задач САПР).

Скорость выполнения ВЗС в УЦО зависит от многих факторов:

- 1) точности применяемых геометрических моделей E_S и E_P ;
- 2) числа центров зондирования;
- 3) времени машинного выполнения каждого ЭАЗ;
- 4) скорости преобразования модели содержимого сцены из одной системы координат в другую;

5) правильного выбора тактики поиска разведывательных перемещений виртуального зонда внутри зондируемой сцены.

Снижение точности $v^{(n)}$ -воксельных моделей E_S и E_p сокращает время выполнения ВЗС, но увеличивает погрешность найденных решений. Для повышения эффективности зондирования сцены целесообразно использовать многоэтапный алгоритм выполнения ВЗС, когда на начальных этапах ищутся только самые быстрые и грубые решения с использованием $v^{(n)}$ -воксельных моделей низкой точности и небольшим числом центров зондирования. Найденные траектории АМР затем уточняются путем их повторного зондирования с применением уже более точных машинных моделей E_S , E_p и G .

У АМР бортовая система технического зрения (СТЗ) имеет ограниченный видеобзор. Поэтому 3D-сцена, сформированная АМР по результатам одного видеомониторинга, в общем случае не может охватить всю территорию маршрута АМР от старта до финиша. В этой связи 3D-модель области, охватывающей весь маршрут движения АМР, может быть сформирована только путем объединения $v^{(n)}$ -воксельных моделей E_S , синтезированных по результатам нескольких видеомониторингов, выполненных бортовой СТЗ из разных положений АМР на маршруте. Для построения единой 3D-карты всей области маршрута содержимое всех исходных сцен должно быть преобразовано из систем координат отдельных сцен в единую абсолютную систему координат всего маршрута АМР.

Неправильный выбор тактики перемещения виртуального зонда внутри исследуемой сцены может привести не только к неоправданному увеличению времени поиска решения, но даже к невозможности решения задачи в принципе.

5. Распараллеливание процедуры зондирования сцены

Повысить скорость выполнения ВЗС можно следующим путем:

- 1) повышением тактовой частоты управления;
- 2) распараллеливанием процедур обработки, наиболее часто используемых при выполнении ВЗС (геометрических преобразований поворота и переноса, теоретико-множественных операций, анализа объектов на пересекаемость и др.).

Рассмотрим подробно второй путь увеличения скорости выполнения ВЗС.

При обработке сцены функцию операндов выполняют геометрические модели объектов сцены. Описание моделей зонда и сцены, хранящееся в памяти УЦО и участвующее в ВЗС, состоит из отдельных компонентов. Обработка этих компонентов может осуществляться в УЦО одновременно и независимо друг от друга, т.е. параллельным обра-

зом. В устройстве цифровой обработки, использующем двоичную систему счисления, самыми низшими структурными компонентами описания сцены являются биты. Глубина распараллеливания ВЗС определяется числом одновременно обрабатываемых компонентов описания: чем выше их число, тем глубже степень распараллеливания ВЗС. Поскольку число простых компонентов геометрического описания всегда превышает число сложных компонентов, то в УЦО наивысшая степень распараллеливания ВЗС может быть достигнута только на уровне битовых элементов описания геометрических моделей E_S и E_p .

В УЦО $v^{(n)}$ -воксельные модели могут храниться и обрабатываться только в виде отдельных адресуемых двоичных фрагментов (например, адресуемых фрагментов зондируемой сцены (АФС) и адресуемых фрагментов зондирующей сцены (АФЗ)), роль которых могут выполнять:

- адресуемый бит;
- адресуемое двоичное слово (по-другому — одномерное бинарное изображение — 1D-БИ);
- 2D-бинарное изображение (2D-БИ);
- 3D-бинарное изображение (3D-БИ).

Установим между пространственным положением $v^{(n)}$ -вокселя $v^{(n)}$ -воксельной модели формата обрабатываемой сцены и адресом адресуемого фрагмента, где этот $v^{(n)}$ -воксель хранится в памяти УЦО, взаимно однозначное соответствие. Установление такого соответствия упорядочит биты, описывающие $v^{(n)}$ -воксельные модели зонда и исследуемой сцены в *RAM*-памяти (т.е. памяти с произвольной выборкой) УЦО. Это ускорит алгоритмы ВЗС, поскольку они будут в этом случае напоминать алгоритмы сортировки.

6. Реализация ВЗС в УЦО с различным типом *RAM*-адресации

При ВЗС наиболее часто выполняемой процедурой является ЭАЗ, поэтому время выполнения ЭАЗ оказывает решающее влияние на скорость выполнения ВЗС. В свою очередь, вследствие хранения $v^{(n)}$ -воксельных моделей зонда и зондируемой сцены в УЦО в виде адресуемых фрагментов каждый ЭАЗ разбивается на множество элементарных актов микрозондирования (ЭАМЗ), где в качестве микрозонда выступает текущий АФЗ $v^{(n)}$ -воксельной модели E_p , а исследуемыми участками микрозондирования являются АФС $v^{(n)}$ -воксельной модели E_S , которые пересекают текущий микрозонд. При этом каждый ЭАМЗ представляет собой последовательность следующих элементарных операций:

- 1) расчет адреса текущего микрозонда в памяти УЦО;
- 2) чтение текущего микрозонда из *RAM*-памяти УЦО в блок анализа микрозондирования (БАМЗ) УЦО в качестве 1-го операнда;

- 3) расчет адресов АФС, биты которых пересекают текущий микрозонд, по текущему координатному положению микрозонда внутри сцены;
- 4) чтения всех АФС, биты которых пересекают текущий микрозонд;
- 5) сборка 2-го операнда для БАМЗ из всех прочитанных АФС;
- 6) проведение анализа на пересекаемость 1-го и 2-го операндов в БАМЗ.

Если хотя бы в одном ЭАМЗ будет обнаружено пересечение, то пересечение сцены с зондом объявляется выявленным.

Наибольшая степень распараллеливания ВЗС достигается в случае увеличения разрядности двоичных фрагментов АФС и АФЗ. Кроме того, для повышения скорости выполнения ЭАМЗ желательно, чтобы в УЦО размерность операндов БАМЗ совпадала с размерностью n зондируемых сцен S , а разрядность АБИ была предельно большой, с точки зрения применяемой микроэлектронной технологической нормы.

Важно отметить, что реализация пространственных поворотов $v^{(n)}$ -воксельных моделей E_p и E_S с помощью традиционных УЦО (где адресуемыми фрагментами являются 1D-БИ) уменьшает скорость поворотов и приводит к разупорядочиванию описания $v^{(n)}$ -вокселей в RAM-памяти УЦО, что дополнительно увеличивает затраты машинного времени на выполнение ВЗС. Для предотвращения этого наиболее оптимальным является применение ВВ.

7. Отличительные признаки и особенности функционирования воксельного вычислителя

Дискретный метод геометрического моделирования имеет исключительные преимущества по глубокому распараллеливанию следующих основных процедур обработки сцен (ОПОС) [9]:

- теоретико-множественных операций над геометрическими моделями;
- анализа геометрических моделей на пересекаемость;
- геометрических преобразований (поворот, перенос, сжатие и др.);
- операций вычисления объема (площади) воксельной (пиксельной) модели;
- определения положения геометрической модели в пространстве.

В полной мере реализовать эти преимущества позволяют специализированные вычислительные устройства, так называемые устройства параллельной обработки n -мерных пиксельных моделей [9—12]. Воксельный вычислитель представляет собой частный случай такого устройства и предназначен для параллельной обработки 2D (пиксельных) и 3D (воксельных) дискретных геометрических моделей. В воксельном вычислителе под пикселем понимается квадратик 1×1 , под вокселем —

кубик $1 \times 1 \times 1$. Отличительные признаки воксельного вычислителя:

- 1) роль адресуемых информационных единиц в них выполняют:
 - традиционные двоичные переменные (1D-АБИ);
 - двумерные адресуемые бинарные изображения (2D-АБИ) формата $A\sqrt{A} \times A\sqrt{A}$;
 - трехмерные бинарные изображения (3D-АБИ) формата $A \times A \times A$ (где $A = q^2$, $q \geq 2$ — число целое);
- 2) в его состав помимо обычного арифметико-логического узла (АЛУ), предназначенного для обработки 1D-АБИ, входят также аппаратные средства, позволяющие выполнять все перечисленные выше ОПОС над 2D-АБИ и 3D-АБИ за время, сравнимое с элементарным тактом управления блока управления ВВ;
- 3) выполнение геометрических преобразований заменено выполнением так называемых дискретизирующих отображений [9, 10] над пиксельными и воксельными моделями;
- 4) система команд ВВ помимо традиционных арифметико-логических команд, используемых для обработки 1D-АБИ, содержит набор команд, оперирующих 2D-АБИ и 3D-АБИ, каждая из таких команд реализует одну ОПОС над одноцветным содержимым АБИ, этот набор является *функционально полным*, позволяющим выполнять весь цикл обработки сцены на основе применения только пиксельных и воксельных моделей.

8. Сверхоперативное хранение воксельных и пиксельных моделей в ВВ

Предельно высокая степень распараллеливания процедуры обработки сцены (т.е. выполнение процедуры за время, сравнимое с тактом ВВ) достигается, если разрешение АБИ будет, как минимум, не меньше разрешения обрабатываемой воксельной модели сцены. К сожалению, существующие микроэлектронные технологии не позволяют создавать процессорные матрицы ВВ с разрешением, сопоставимым с разрешением реальных обрабатываемых 2D- и 3D-сцен. В этой связи обработка сцен среднего и высокого разрешения в ВВ может осуществляться только предельно крупными фрагментами формата $A \times A \times A$ или $A\sqrt{A} \times A\sqrt{A}$ и выполняемых последовательно — фрагмент за фрагментом. Сама сцена или ее части представляются в ВВ в виде бинарных составных 2D- и 3D-изображений (БСИ), состоящих из 2D-АБИ и 3D-АБИ.

Если форматы АБИ-фрагментов и положение вокселей в них жестко определены еще на этапе проектирования ВВ, то формат БСИ, положение и ориентация АБИ-фрагментов в этом БСИ требуют дополнительного описания в программе обработки содержимого БСИ. При этом пространственное положение каждого АБИ-фрагмента $A \times A \times A$ внутри описываемого БСИ должно быть строго увязано с адресом этого АБИ-фрагмента (СПИ-ад-

ресом) в сверхоперативной памяти изображений ВВ (т.е. RAM-памяти УЦО). На рис. 2 показан вариант пространственной интерпретации содержимого участка СПИ, состоящей из $M_x \cdot M_y \cdot M_z = 729$ АБИ-фрагментов $A \times A \times A$ одного БСИ, размещенных в 729 последовательных адресах СПИ. Как следует из рисунка, одноцветное содержимое всех указанных АБИ-фрагментов образует единую воксельную модель с разрешением, не превышающим $M_x \times M_y \times M_z$, в котором каждый АБИ-фрагмент имеет строго определенное пространственное положение.

9. Анализ на пересекаемость двух БСИ средствами воксельного вычислителя

Рассмотрим способ реализации в ВВ процедуры анализа на пересекаемость содержимого двух БСИ с форматами $M_x \cdot A \times M_y \cdot A \times M_z \cdot A$ и $K_x \cdot A \times K_y \cdot A \times K_z \cdot A$ соответственно, где $K_y < M_y$, $K_x < M_x$, $K_z < M_z$ (рис. 3, а).

Единичное содержимое БСИ $K_x \cdot A \times K_y \cdot A \times K_z \cdot A$ может быть интерпретировано как воксельная модель зонда (ВМЗ). Единичное содержимое БСИ $M_x \cdot A \times M_y \cdot A \times M_z \cdot A$ может быть интерпретировано как воксельная модель содержимого исследуемой сцены (ВМС). Повороты и переносы ВМЗ моделируются в ВВ путем выполнения соответствующих дискретизирующих отображений (ДО) над содержимым БСИ $K_x \cdot A \times K_y \cdot A \times K_z \cdot A$ [9, 10]. Параллельный перенос ВМЗ внутри ВМС моделируется путем переноса БСИ $K_x \cdot A \times K_y \cdot A \times K_z \cdot A$ относительно БСИ $M_x \cdot A \times M_y \cdot A \times M_z \cdot A$ следующим образом.

Каждый i -й АБИ-фрагмент $A \times A \times A$ (АБИ i) переносимого БСИ $K_x \cdot A \times K_y \cdot A \times K_z \cdot A$ (где $0 \leq i \leq (K_x \cdot K_y \cdot K_z - 1)$) пересекает, в общем случае, восемь соседних АБИ-фрагментов, принадлежащих БСИ $M_x \cdot A \times M_y \cdot A \times M_z \cdot A$. Эти восемь АБИ-фрагментов отсекают от АБИ i восемь областей: $Gi0...Gi7$ соответственно (рис. 3, б). Анализ на пересекаемость содержимого БСИ $K_x \cdot A \times K_y \cdot A \times K_z \cdot A$ и содержимого БСИ $M_x \cdot A \times M_y \cdot A \times M_z \cdot A$ состоит в выполнении восьми $\cdot K_x \cdot K_y \cdot K_z$ теоретико-множественных операций "пересечение" над содержимым $Gi0...Gi7$ всех $K_x \cdot K_y \cdot K_z$ АБИ-фрагментов $A \times A \times A$ БСИ $K_x \cdot A \times K_y \cdot A \times K_z \cdot A$ с содержимым соответствующего АБИ-фрагмента БСИ $M_x \cdot A \times M_y \cdot A \times M_z \cdot A$. Если результатом хотя бы одной из этих операций будет не пустое единичное

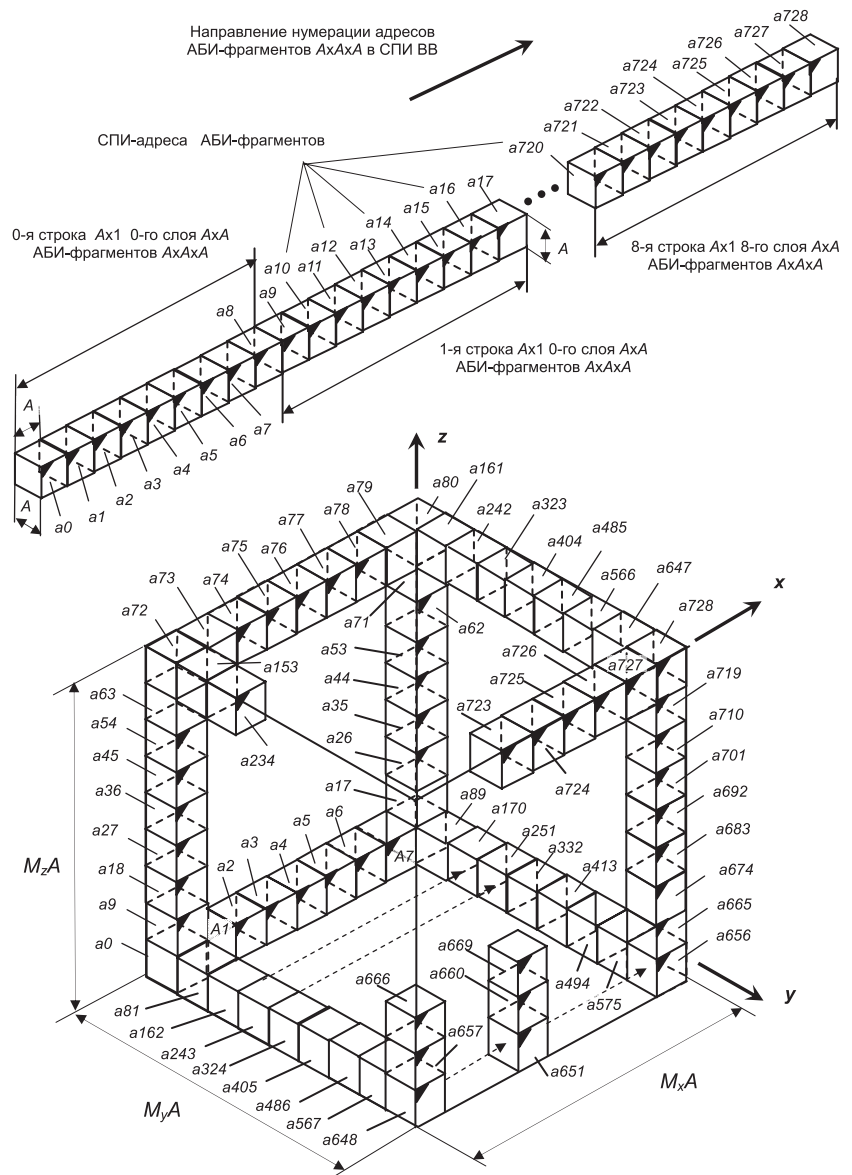


Рис. 2. Вариант пространственной интерпретации содержимого участка памяти ВВ в виде БСИ формата $9A \times 9A \times 9A$, состоящего из 729 АБИ-фрагментов $A \times A \times A$

содержимое результирующего АБИ-фрагмента, то содержимое БСИ $K_x \cdot A \times K_y \cdot A \times K_z \cdot A$ и БСИ $M_x \cdot A \times M_y \cdot A \times M_z \cdot A$ объявляются взаимно пересекающимися. ВВ за время, сравнимое с тактом управления, позволяет:

- выделить каждую область $Gi0...Gi7$ в виде отдельного АБИ со СПИ-адресом r ;
- сформировать АБИ со СПИ-адресом d , представляющее собой результат повоксельного пересечения содержимого АБИ (со СПИ-адресом r) с содержимым своего АБИ-фрагмента $A \times A \times A$, принадлежащего БСИ $M_x \cdot A \times M_y \cdot A \times M_z \cdot A$;
- осуществить анализ единичного содержимого АБИ со СПИ-адресом d на "пусто/не пусто".

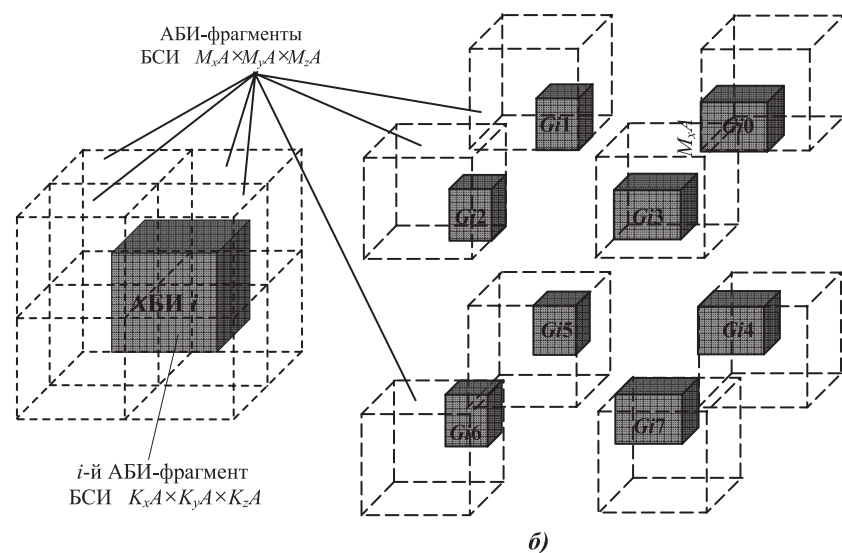
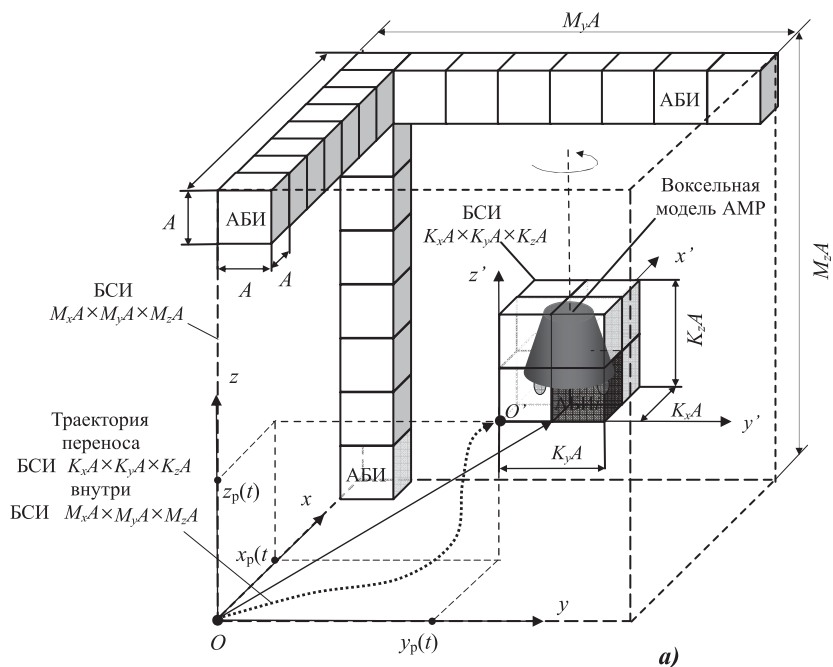


Рис. 3. Зондирование БСИ $M_x A \times M_y A \times M_z A$ с помощью зондирующего БСИ $K_x A \times K_y A \times K_z A$ средствами воксельного вычислителя

10. Преимущества применения воксельной АПТОС для организации виртуального зондирования сцены

Скоростное выполнение ВЗС требует наличия в УЦО специальных аппаратно-программных средств для скоростной обработки $v^{(n)}$ -воксельных моделей. В традиционных компьютерах аппаратно не реализованы многие операции и процедуры, используемые при организации ВЗС, которые необходимо выполнять с высокой степенью параллелизации над пиксельными и воксельными моделями, например, анализ моделей на пересекаемость, теоретико-множественные операции, геометрические преобразования поворота и переноса и другие. В традиционных программируемых компьютерах

скоростное выполнение геометрических преобразований проводится только с применением многоядерных центральных процессоров (CPU) или графических конвейеров (GPU) над полигональными числовыми геометрическими моделями или дискретными точечными моделями, описываемыми с помощью числовых координат.

Наоборот, в системе машинных команд ВВ присутствует целая группа команд (так называемых команд блока обработки БИ), существенно повышающих скорость процедуры зондирования сцены. Поскольку блок управления в ВВ имеет архитектуру традиционного УЦО, возможности обработки числовой информации в ВВ, как минимум, не будут уступать возможностям традиционных компьютеров и микроконтроллеров.

Заключение

На основании изложенного выше можно сделать вывод, что метод ВЗС на базе воксельной АПТОС позволяет создать сверхскоростные навигационные видеоанализаторы, которые, будучи встроенными в систему управления АМР, могут значительно повысить их адаптивность, маневренность и автономность, что, в свою очередь, расширит области и масштабы применения АМР в военном деле, сельском хозяйстве, промышленности, сфере услуг.

Список литературы

1. Stout B. Smart Moves: Intelligent Path-Finding. URL: http://www.gamasutra.com/view/feature/131724/smart_move_intelligent_php.
2. Плотников В. А. Анализ эффективности существующих методов уклонения от столкновения для мобильного робота. Донецк: Штучный интеллект, 2010.
3. Seder M., Macek K., Petrovic I. An integrated approach to real-time mobile robot control in partially known indoor environments // Industrial Electronics Society, 2005. IECON 2005. 31st Annual Conference of IEEE. 6 p.
4. Ferguson D., Darms M., Urmson C., Kolski S. Detection, Prediction and Avoidance of Dynamic Obstacles in Urban Environments // Proceedings of the 2008 IEEE Intelligent Vehicles Symposium. 2008. June. P. 1149–1154.
5. Герасимов В. Н., Михайлов Б. Б. Решение задачи управления движением мобильного робота при наличии динамических препятствий // Вестник МГТУ им. Н. Э. Баумана. Сер. "Приборостроение". 2012. № 6. С. 83–92.
6. Андрианов О. Г., Королькова М. А., Кочетков С. А., Краснова С. А. Кинематическое управление мобильным роботом при движении по полигону с обходом препятствий // Материалы конференции "Управление в технических системах" (УТС–2010, Санкт-Петербург). СПб.: ЦНИИ "Электроприбор", 2010. С. 356–359.

7. Зозуля Ю. И. Интеллектуальные системы обработки информации на основе нейросетевых технологий: учеб. пособие. Уфа: Изд. Уфимск. гос. авиац. техн. ун-т, 2000. 138 с.

8. Юдинцев Б. С., Даринцев О. В. Интеллектуальная система планирования траекторий мобильных роботов, построенная на сети Хопфилда // Современные проблемы науки и образования. 2014. № 4. URL: <http://www.science-education.ru/ru/article/view?id=14131>.

9. Бимаков В. А. Высокоскоростная обработка n -мерных пиксельных геометрических моделей // Информационные технологии. 2009. № 3. С. 29—36.

10. Бимаков В. А., Бимаков Е. В. Устройство обработки двумерных и трехмерных изображений // Патент РФ № 2376637. 2009. Бюл. № 35.

11. Бимаков В. А., Бимаков Е. В. Устройство управления автономным роботом. Патент РФ № 2424105. 2011. Бюл. № 20.

12. Бимаков В. А., Бимаков Е. В. Принципы построения и функционирования устройств для высокоскоростной обработки n -мерных пиксельных геометрических моделей // Информационные технологии. 2012. № 6. С. 44—49.

13. Бимаков В. А., Касимов Т. Р. Использование воксельных вычислителей для скоростного преобразования 2D- и 3D-изображений высокого разрешения // Информационные технологии. 2013. № 3. С. 42—47.

14. Бимаков В. А., Бимаков Е. В., Касимов Т. Р. О применении воксельных вычислителей для решения задач стереозрения и навигации автономного робота // Мехатроника, автоматизация, управление. 2013. № 2. С. 47—54.

15. Бимаков Е. В., Касимов Т. Р. Разработка системы управления автономным мобильным роботом на базе воксельного вычислителя // Электронный научный журнал "Труды МАИ". 2014. № 74. URL: <http://www.mai.ru/publications/index.php?ID=49313>.

E. V. Bimakov, Leading Engineer, LLC "Voxel equipment",
Graduate Student, Izhevsk State Technical University, robint@mail.ru

The Method of Virtual Probing of Scene Based on the Voxel Technology of Scene Processing

The high-speed method of virtual probing of a scene based on voxel technology of scene processing is explained. The method is intended for the solution of tasks of navigation of the autonomous ground mobile robot (obstacle avoidance and object recognition in real time). The general concept of this method, the choice of the type of digital device, the choice of probe model and scene model for the high-speed virtual scene probing are described. High speed of the automatic analysis of scene is reached thanks to deep multisequencing of operations and procedures which are most often used when scene is probing.

Keywords: scene and image processing, navigation, autonomous ground robot, obstacle avoidance, object recognition, voxel geometrical model, voxel computers

References

1. Bryan Stout. *Smart Moves: Intelligent Path-Finding*, URL: http://www.gamasutra.com/view/feature/131724/smart_move_intelligent_.php.

2. Plotnikov V. A. *Analiz jeffektivnosti sushhestvujushhikh metodov uklonenija ot stolknovenija dlja mobil'nogo robota*. Doneck: Shtuchnij intelekt, 2010.

3. Seder M., Macek K., Petrovic I. An integrated approach to real-time mobile robot control in partially known indoor environments, *Industrial Electronics Society, 2005. IECON 2005. 31st Annual Conference of IEEE*. 6 p.

4. Ferguson D., Darms M., Urmsen C., Kolski S. Detection, Prediction and Avoidance of Dynamic Obstacles in Urban Environments, *Proceedings of the 2008 IEEE Intelligent Vehicles Symposium*, 2008. June. P. 1149—1154.

5. Gerasimov V. N., Mihajlov B. B. Reshenie zadachi upravlenija dvizheniem mobil'nogo robota pri nalichii dinamicheskikh prepjatstvij, *Vestnik MGTU im. N. Je. Bauman. Ser. Priborostroenie*, 2012, no. 6, pp. 83—92.

6. Andrianova O. G., Korol'kova M. A., Kochetkov S. A., Krasnova S. A. Kinematischeskoe upravlenie mobil'nym robotom pri dvizhenii po poligonu s obhodom prepjatstvij, *Materialy konferencii "Upravlenie v tehniceskikh sistemah" (UTS—2010, Sankt-Peterburg)*, SPb.: JSC "Koncern "CNII "Jelektropribor". 2010. P. 356—359.

7. Zozulja Ju. I. *Intellektual'nye sistemy obrabotki informacii na osnove nejrosetevyh tehnologij*: Ucheb. Posobie. Ufa: Ufimsk. gos. aviac. tehn. un-t, 2000. 138 p.

8. Judincev B. S., Darincev O. V. Intellektual'naja sistema planirovanija traektorij mobil'nyh robotov, postroennaja na seti Hopfilda, *Sovremennye problemy nauki i obrazovanija*, 2014, no. 4. URL: <http://www.science-education.ru/ru/article/view?id=14131>.

9. Бимаков В. А. Высокоскоростная обработка n -мерных пиксельных геометрических моделей, *Информационные технологии*, 2009, no. 3, pp. 29—36.

10. Бимаков В. А., Бимаков Е. В. Устройство обработки двумерных и трехмерных изображений, *Patent RF № 2376637, 2009, Bjul*, no. 35.

11. Бимаков В. А., Бимаков Е. В. Устройство управления автономным роботом, *Patent RF № 2424105. 2011. Bjul*. no. 20.

12. Бимаков В. А., Бимаков Е. В. Принципы построения и функционирования устройств для высокоскоростной обработки n -мерных пиксельных геометрических моделей, *Информационные технологии*, 2012, no. 6, pp. 44—49.

13. Бимаков В. А., Касимов Т. Р. Использование воксельных вычислителей для скоростного преобразования 2D и 3D изображений высокого разрешения, *Информационные технологии*, 2013, no. 3, pp. 42—47.

14. Бимаков В. А., Бимаков Е. В., Касимов Т. Р. О применении воксельных вычислителей для решения задач стереозрения и навигации автономного робота, *Мехатроника, автоматизация, управление*, 2013, no. 2, pp. 47—54.

15. Бимаков Е. В., Касимов Т. Р. Разработка системы управления автономным мобильным роботом на базе воксельного вычислителя, *Электронный научный журнал "Труды МАИ"*, 2014, no. 74. URL: <http://www.mai.ru/publications/index.php?ID=49313>.

УДК 004.6; 528; 004.8

С. Л. Беляков, д-р техн. наук,
проф. каф. информационно-аналитических систем безопасности, beliacov@yandex.ru;
М. Л. Белякова, канд. техн. наук,
доц. каф. информационно-измерительных систем и технологий, mbelyakova@sfnu.ru;
А. А. Глушков, аспирант, andrey@glushkov.net,
Инженерно-технологическая академия Южного федерального университета, г. Таганрог

Метапреобразования образов при поиске достоверных решений в интеллектуальных геоинформационных системах

Предлагается метод построения решений на основе опыта, представленного картографическими образами — центрами и преобразованиями ситуаций, не изменяющими их смысл. Метод базируется на метапреобразованиях образов. Цель метапреобразования заключается в конструировании новых ситуаций по известным прецедентам. Использование метапреобразований позволяет особым образом оценивать смысловую близость ситуаций и адаптировать ранее принятые решения. Проанализированы факторы, определяющие достоверность сформированных решений.

Ключевые слова: принятие решений, знания, прецедентный анализ, интеллектуальные системы, картографическая визуализация, геоинформационные системы

Введение

Принятие достоверных решений с использованием геоинформационных систем имеет важное значение во многих областях бизнеса и управления. Представление картами и схемами планируемых или реализуемых проектов транспортировки, строительства, защиты от чрезвычайных ситуаций и многих других задач не только несет важную для принятия решений информацию, но и способно сохранить опыт решения проблем для последующего использования. Эта возможность реализуется при разработке интеллектуальных геоинформационных систем и сетевых сервисов [1–3]. Отличительной особенностью систем подобного типа является использование знаний. За счет знаний стремятся добиться высокого уровня достоверности — соответствия решений реальному состоянию той области внешнего мира, в которой будет исполнено принятое решение. При несоответствии возникает ущерб как следствие применения неадекватных решений [4].

Недостаточный уровень достоверности решений, которые строят интеллектуальные системы, обусловлен несколькими причинами:

- недостатком и неопределенностью описания внешнего мира, в котором будет реализовано решение;

- неполнотой описания ранее принятых решений, на основе которых строится новое;
- неоднозначностью влияния параметров решений на их практическую реализацию.

Одним из перспективных путей повышения уровня достоверности может стать накопление высококачественных знаний и применение механизмов рассуждений, компенсирующих неполноту, неоднозначность и неопределенность знаний [3]. В каждом классе интеллектуальных систем такой подход имеет специфику. В случае применения геоинформационных систем (ГИС) можно указать несколько предпосылок повышения достоверности:

— в составе ГИС имеется картографическая основа с богатым по содержанию описанием среды реализации решений;

— картографическая основа системы непрерывно обновляется, что позволяет оценивать качество адаптации ранее принятых решений к текущим условиям;

— система визуализации дает особые возможности представления и использования опыта принятия решений в виде картографических объектов [5].

Реализовать перечисленные потенциальные возможности повышения достоверности в ГИС можно за счет введения новых моделей использования опыта.

Прецедентный анализ в ГИС

В основе получения решений на основе опыта лежит прецедентный анализ [6]. При анализе наблюдавшихся ранее прецедентов в соответствии с гипотезой компактности предполагают, что близкие по смыслу ситуации имеют близкие по смыслу решения. Следовательно, задача оценки близости прецедентов играет важнейшую роль в отборе потенциально применимых решений. Трудностью решения данной задачи является подбор метрики оценки близости прецедентов [7]. Метрика сложным образом отражает свойства пространства наблюдаемых ситуаций и картины мира эксперта, фиксирующего прецеденты. Ее вид необходимо подбирать и оценивать соответственно предметной области, модели ее представления и логике вывода заключений.

В данной работе рассматривается прецедентный анализ при использовании опыта управления материальными потоками [8]. Задача заключается в том, чтобы при заданных ресурсах обеспечить перемещение материальных объектов из одной области пространства в другую с заданным уровнем показателей качества. Перемещение товаров, продуктов, материалов, сырья предполагает конструирование логистических проектов и мониторинг их исполнения с помощью ГИС. Под логистическим проектом понимается информационная модель конкретной задачи доставки груза. Материальный поток в логистическом проекте может быть описан моделью

$$\langle S_L, P_S, T \rangle,$$

в которой S_L — множество пространственно локализованных логистических центров; P_S — множество транспортных маршрутов между логистическими центрами; T — расписание доставки груза. Под логистическим центром здесь обобщенно понимается место выполнения логистических операций над грузом (упаковка, распаковка, погрузка, разгрузка, сортировка, переработка, и т.д.). Картографическое описание потока отражает привязку к условиям реального мира и использует пространственные данные для стратегического планирования и оперативного управления потоком.

Картографическое описание логистического проекта может быть построено в терминах ситуаций [9]. Тогда принятие управляющих решений следует рассматривать как результат картографического анализа ситуаций, маршрутов транспортировки и ограничений расписания. Картографический анализ может быть выполнен традиционным образом [4, 10, 11]. Однако практика показала, что это не дает ожидаемого результата по следующим причинам.

1. Инструментарий ГИС предназначен для обработки карт, схем и планов как образно-знаковых моделей реального мира. Смысловое пространство

ситуаций и решений непосредственным образом не отображается и требует разработки специальных моделей представления смысла в определенной трактовке, в частности, для задач управления материальными потоками.

2. Понятие "смысл ситуации" трактуется ГИС-аналитиками субъективно. Вследствие этого невозможно создавать единое смысловое пространство ситуаций и обеспечить повторное использование опыта принятия решений. Требуется унифицировать понятие смысла ситуации в контексте принятия решений при управлении материальными потоками.

3. Понятие достоверности связывается исключительно со свойствами пространственных данных. На практике этого недостаточно, поскольку пространственные данные отражают действительность фактографически. Необходимы знания, которые позволяют прогнозировать вид и последствия принимаемых решений.

Указанные проблемы можно решать введением новой концепции представления и использования карт, схем и планов для описания смысла ситуаций.

Специфический подход к представлению прецедентов в геоинформационных системах заключается в представлении ситуаций образами [12].
Образ ситуации

$$I_s = \langle c, H(c) \rangle$$

описывается центром образа (c) и набором его преобразований $H(c) = (h_1(c), h_2(c), \dots, h_M(c))$. Содержательно центр — это реально наблюдавшаяся ранее ситуация, а преобразования описывают возможные, с точки зрения эксперта, трансформации параметров ситуации, не изменяющих ее смысл. Преобразования касаются пространственных, временных и семантических характеристик ситуации. Сравнение ситуаций (оценка их близости) строится на сравнении взаимного расположения центров образов и их преобразований. Тем самым моделируется образное мышление человека [13, 14]. Логика образного сравнения описана в работе [12]. Там же предложена метрика оценки близости ситуаций как степень общности преобразований h_i двух образов I_1 и I_2 :

$$\alpha = \frac{2S(h_i^{(I_1)} \cap h_i^{(I_2)})}{Sh_i^{(I_1)} + Sh_i^{(I_2)}}, \quad i = \overline{1, M},$$

где $S(x)$ — площадь области x . Значение $\alpha = 1$ имеет место при совпадении возможных преобразований, $\alpha = 0$ — в противном случае. Можно заметить, что в данном выражении заложена возможность сравнивать образы ситуаций исключительно в общей области карты. Если это не так, то

$$S(h_i^{(I_1)} \cap h_i^{(I_2)}) = 0$$

и ситуации рассматриваются как различающиеся по смыслу. Данная особенность неслучайна и позволяет гарантировать достоверность решений, основываясь на дедукции [15]. Однако подобное ограничение делает невозможным использование опыта принятия решений в территориально удаленных ситуациях. В то же время практика работы экспертов показывает необходимость и возможность сравнивать "одно и то же в разных местах" [16]. С точки зрения логики рассуждений в рассматриваемом случае дедукция должна заменяться абдукцией, направленной на получение дополнительных фактов и утверждений для достоверного вывода. Поэтому необходимо усовершенствовать принцип сравнения преобразований ситуаций.

Целью настоящей работы является повышение достоверности генерируемых ГИС решений на основе применения метода преобразования образов ситуаций.

Постановка задачи поиска достоверных решений

В общем виде задачу нахождения максимально достоверных решений с помощью ГИС сформулируем следующим образом: имеется набор прецедентов принятия решения в виде образов $I = \{I_n\}$, $n = \overline{1, |I|}$, каждый из которых включает компоненту описания ситуации и принятого решения $I_n = \langle I_{s_n}, I_{d_n} \rangle$.

Пусть задан образ проблемной ситуации $\tilde{I} = \langle \tilde{I}_s, \tilde{I}_d \rangle$, для которой требуется найти решение. Обозначим через I^* образ решения, который соответствует проблемной ситуации и имеет достоверное решение. То, что такое решение действительно существует, можно проверить только экспериментально, поэтому будем считать, что решение соответствует реальности, если оно согласовано с опытом реализации решений и непротиворечиво изображено на карте, состоящей из множества элементов Ω , т.е. при его построении не нарушены правила, отражающие связи и отношения между картографическими объектами:

$$R(I_d^*, \Omega) = true.$$

Тогда поиск достоверного решения проблемной ситуации заключается в том, чтобы построить образ решения \tilde{I}_d , который наиболее близок к реальности:

$$\begin{aligned} \|I_d^* - \tilde{I}_d\| &\rightarrow \min, \\ \|I_{s_n} - I_s^*\| &< d, n = \overline{1, |I|}, \\ R(\tilde{I}_d, \Omega) &= true, \end{aligned} \quad (1)$$

где $\|\cdot\|$ — метрика расстояния между образами; d — ограничение на значение расстояния. Анализируя

приведенную постановку задачи и пути нахождения ее решения, следует отметить следующее.

1. Поскольку образ I_d^* не известен, задачу следует решать конструктивно: перебирая известные прецеденты, строить на их основе возможные решения и оценивать их реальность через корректность картографического отображения.

2. Процедура поиска решения поставленной задачи неустойчива. Всякий прецедент есть совокупность объектов и отношений, вероятность точного повторения которых в других областях пространства и времени практически равна нулю. Снижение точности сравнения не даст ожидаемого эффекта, так как непредсказуемо влияет на достоверность решения. Поэтому при использовании описания прецедентов образами, в которых указаны не только ситуации, но и их допустимые изменения, возрастает вероятность нахождения близкой ситуации, процесс решения приобретает устойчивость.

3. Метрика близости ситуаций должна быть определена через преобразования картографических объектов, поскольку конструктивный характер поиска решения задачи (1) требует отображения образов прецедентов в проблемную пространственно-временную и семантическую область.

4. Правила непротиворечивости $R(\tilde{I}_d^*, \Omega)$ в постановке задачи (1) выделены явно, что создает определенные трудности поиска решений: часть сгенерированных альтернатив окажется бесполезной после проверки правил. Представляется целесообразным включить ограничения непротиворечивости в процедуры картографического отображения.

Метапреобразование образов

Постановка задачи (1) имеет недостатки, которые можно преодолеть введением метапреобразований, обеспечивающих отображение образов ситуаций и решений в произвольные области карты. Через $I_2 = H(I_1, c_2)$ обозначим метапреобразование образа I_1 в образ I_2 с центром c_2 . Результат метапреобразования представляет собой набор картографических объектов, который должен быть построен в области карты, заданной центром c_2 . Требование к метапреобразованию состоит в том, чтобы все объекты и отношения образа I_1 были сохранены в образе I_2 . Этим гарантируется сохранение смысла и непротиворечивость.

В качестве примера рассмотрим метапреобразование места парковки транспортного средства. Парковка транспорта рассматривается как простейший логистический центр, который является элементом логистической цепи. Образ парковки описывается следующим образом: центр образа — это точка на карте, указывающая место припаркованного транспортного средства, преобразования — зона его возможного расположения, не влияющая на смысл операции парковки. Обозначим через \hat{P}_1 образ прецедента парковки, через \hat{P}_2 — образ проб-

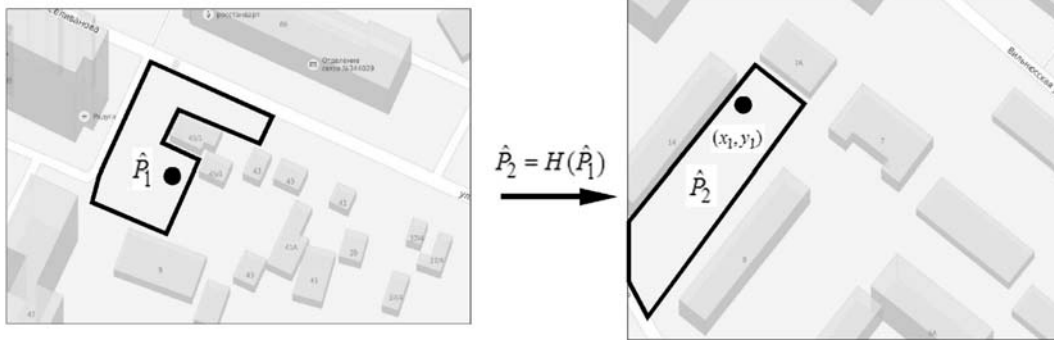


Рис. 1. Пример метапреобразования "Парковка легкового автомобиля"

лемной ситуации. На рис. 1 образ \hat{P}_1 отображен многоугольником, центр образа отображен точкой. Центр образа \hat{P}_2 задается координатами точки (x_1, y_1) . Метапреобразование "Парковка легкового автомобиля" $\hat{P}_2 = (\hat{P}_1, c_2)$, которое отображает образ \hat{P}_1 в область, заданную центром образа c_2 , в точке (x, y) строит площадной объект \hat{P}_2 класса "Место парковки" в том случае, если:

- точка (x, y) находится в отношении "Находиться Внутри" с объектом одного из классов: "Место парковки", "Платная парковка", "Бесплатная парковка", "Асфальтированная площадка", "Дорожная обочина", "Внутриквартальный проезд", а также с объектом \hat{P}_2 ;

- объект не находится в отношении "Пересекает" с объектами одного из классов: "Здания", "Соружения", "Инженерные коммуникации", "Водоём", "Озеленение";

- объект \hat{P}_2 имеет площадь, не отличающуюся от площади \hat{P}_1 более, чем на 30 %, габаритные размеры \hat{P}_2 не должны отличаться от габаритных размеров \hat{P}_1 более, чем на 10 %;

- объект \hat{P}_2 должен находиться в отношении "Примыкает" к объекту "Автомаршрута", либо "Проезд", либо "Переулок", либо "Автомаршрут";

- объект \hat{P}_2 не должен находиться ближе 200 м от объекта "Железнодорожный переезд";

- объект \hat{P}_2 не должен находиться в зоне действия дорожного знака "Остановка запрещена".

На рис. 1 показан результат отображения \hat{P}_1 в заданную точку. Можно видеть, что метапреобразование заключалось в геометрическом построении выпуклого многоугольника и проверке соблюдения перечисленных выше ограничений.

Внутренняя структура метапреобразования имеет следующий вид:

$$H: \langle I_{s_1}, I_{d_1} \rangle \rightarrow \langle I_{s_2}, I_{d_2} \rangle,$$

$$H = \langle H_s, H_d \rangle,$$

$$H_s: I_{s_1} \rightarrow I_{s_2},$$

$$H_d: I_{d_1} \rightarrow I_{d_2}.$$

Подчеркнем, что метапреобразование образа включает в себя два компонента: метапреобразования ситуации (H_s) и метапреобразования решения (H_d). Такой вид соответствует предположению о логической независимости ситуаций и принятых в них решений. Это позволяет учесть в процессе поиска решений четыре случая:

1) $H_s(I_{s_1}, c_{s_2}) \neq \emptyset, H_d(I_{d_1}, c_{d_2}) \neq \emptyset$ означает, что прецедент с образом I_1 может происходить в другом месте и ранее принятое решение может использоваться повторно;

2) $H_s(I_{s_1}, c_{s_2}) = \emptyset, H_d(I_{d_1}, c_{d_2}) \neq \emptyset$ означает, что прецедент с образом I_1 не может возникнуть в другом месте, хотя принятое ранее решение вполне применимо в рассматриваемом случае;

3) $H_s(I_{s_1}, c_{s_2}) \neq \emptyset, H_d(I_{d_1}, c_{d_2}) = \emptyset$ означает, что прецедент с образом I_1 может произойти в другом месте, но его решение не удастся использовать повторно;

4) $H_s(I_{s_1}, c_{s_2}) = \emptyset, H_d(I_{d_1}, c_{d_2}) = \emptyset$ означает, что прецедент с образом I_1 уникален, т.е. не может возникнуть в другом месте, а принятое решение использоваться повторно.

Перечисленные случаи охватывают практически полезные варианты повторного использования опыта.

Будем считать, что метапреобразование существует, если может быть построено нетривиальное отображение из картографических объектов:

$$H(I_1, c_2) \neq \emptyset \Leftrightarrow I_1 \neq \emptyset, I_1 \in \Omega, I_2 \neq \emptyset, I_2 \in \Omega.$$

Существенную роль в оценке результата метапреобразования играет то, что исходные преобразования центра образа описываются вектором преобразований $(h_1(c), h_2(c), \dots, h_M(c))$. Каждое отдельное преобразование $h_i(c)$ является картографическим объектом. Результат метапреобразования может рассматриваться как сужение (например, площадной объект отображен в линейный либо линейный в точечный) либо как расширение (линейный объект отображен в площадной либо точечный отображен в линейный). Нейтральным следует считать полное соответствие типов картографиче-

ских объектов. Тогда норму метапреобразования можно определить как

$$\|H(\bar{h}(c), c)\| = \sum_i f(h_i(c)),$$

$$f(h_i(c)) = \begin{cases} f_0, & \text{если объекты отображения отсутствуют,} \\ f_1, & \text{если объекты сужают отображение,} \\ f_2, & \text{если отображение нейтрально,} \\ f_3, & \text{если объекты расширяют отображение.} \end{cases}$$

Поиск достоверных решений, представленный задачей (1), можно переформулировать как нахождение образа $\tilde{I} = \langle\langle \tilde{c}_s, H_s(\tilde{c}_s) \rangle\rangle, \langle\langle \tilde{c}_d, H_d(\tilde{c}_d) \rangle\rangle$ такого, что

$$\|H_d(I_{d_n}, \tilde{c}_d)\| \rightarrow \max,$$

$$\|H_s(\tilde{I}_{s_n}, \tilde{c}_s)\| < d, \quad (2)$$

$$n = \overline{1, |\tilde{I}|}.$$

Постановка задачи (2) отличается тем, что мерой достоверности является число преобразований, которые отображены метапреобразованиями образов. Преимущество такого подхода проявляется в следующем.

Во-первых, условия реального мира, в которых будет применяться сформированное решение, непрерывно изменяются. Реализация логистических проектов при одних и тех же исходных данных никогда в точности не повторяется [8]. Как следствие, не повторяются решения и не повторяются способы их модификации. Чем динамичнее внешние условия и выше чувствительность решений к изменению этих условий, тем выше вероятность сконструировать недостоверное решение. Логика прецедентного анализа в этом случае не нарушена. Причина недостоверности решения — в отсутствии прогноза его использования. После включения в механизм прецедентного анализа оценки предполагаемого эффекта достоверность принятых решений возрастает.

Во-вторых, геоинформационные системы естественным образом поддерживают прогнозирование состояния реального мира. Географические карты никогда не являются мгновенным снимком реальности [17]. В них заложены обобщения, являющиеся результатом картографической генерализации. По этой причине достоверность найденных решений непрерывно поддерживается за счет обновления картографической основы ГИС.

В-третьих, использование метапреобразований повышает робастность решений. Знания об устойчивости решений, инвариантах их применения имеют большое практическое значение для логистиче-

ских систем, позволяя избежать больших потерь при малых изменениях условий транспортировки.

Процедура поиска достоверных решений в соответствии с постановкой задачи (2) состоит из следующих действий:

1. Задаются категория образов и метапреобразования, которые будут использоваться для поиска решения.

2. Задается центр образа проблемной ситуации \tilde{I} , представляющий собой набор картографических объектов. Для задач управления материальными потоками центром является начальная и конечная точки транспортировки.

3. Из базы знаний ГИС отбираются прецеденты заданной категории. Если таковых не имеется, следует изменить категорию и предпринять следующую попытку отбора. При отсутствии прецедентов процедура поиска завершается без результата, что вызвано отсутствием опыта.

4. Формируется множество возможных решений задачи. Для этого выполняются метапреобразования образов ситуации в центр образа проблемной ситуации и контролируется значение нормы метапреобразования. Если отсутствуют близкие по смыслу образы, процедура завершается без результата.

5. Выполняется поиск наиболее достоверного решения. С этой целью выполняются метапреобразования образов решения в центр образа решений проблемной ситуации и отбирается вариант с максимальным значением нормы метапреобразования.

Пример реализации

Рассмотрим пример, иллюстрирующий предложенную процедуру. Пусть требуется выполнить транспортировку груза из точки **М** в точку **Н** (рис. 2, см. третью сторону обложки). В базе знаний ГИС имеются прецеденты транспортировки $I = \{I_{AB}, I_{CD}, I_{EF}\}$. На карте показаны центры образов прецедентов, имеющие вид пары точек и стрелки, соответствующей направлению транспортировки. Центр образа показывает положение на карте исходного и конечного пунктов транспортировки.

На рис. 3 (см. третью сторону обложки) показаны преобразования образов ситуаций I_{AB} , I_{MN} , находящиеся на соответствующих фрагментах карты. Преобразование для I_{MN} получено метапреобразованием "Площадка для погрузки" точки **А** в точку **М**, метапреобразованием "Площадка для разгрузки" точки **В** в точку **Н**. Каждое метапреобразование сохраняет габаритные размеры и площадь картографических объектов, отношения с прилегающими объектами, аналогичные приведенным в предыдущем примере. Преобразования показаны заштрихованными зонами. Предположим,

что метрика вектора преобразования вычисляется по формуле

$$\|H(\tilde{h}(c), c)\| = \sum_i f(h_i(c)),$$

$$f(h_i(c)) =$$

$$= \begin{cases} 0, & \text{если объекты отображения отсутствуют,} \\ 1, & \text{во всех остальных случаях.} \end{cases}$$

Тогда число отображенных объектов равно 2 и при $d = 1$ образ I_{AB} является близким по смыслу образу I_{MN} .

Решение для прецедента с образом I_{AB} включает в себя четыре логистических центра $\{S_{L_1}, S_{L_2}, S_{L_3}, S_{L_4}\}$ и три участка транспортировки $P_{S_1}, P_{S_2}, P_{S_3}$. На рис. 4 все они показаны в виде схемы преобразования элементов образа. Преобразования логистических центров имеют вид заштрихованных многоугольников, преобразования каждого участка транспортировки представлены набором ломаных линий, каждая из которых указывает возможную траекторию перемещения груза.

Метапреобразование логистического центра выполняется оговоренным выше способом, метапреобразование траектории пути строит самый короткий путь между заданными логистическими центрами. Схема результата метапреобразования решения представлена на рис. 5. Штриховой линией показаны траектории транспортировки, за-

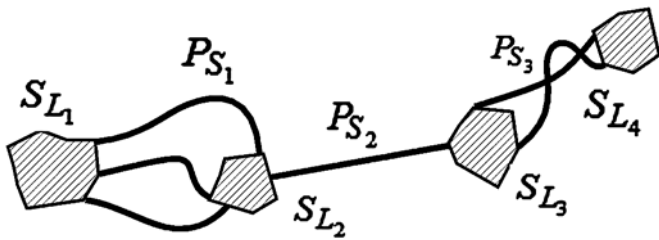


Рис. 4. Преобразования решения образа I_{AB}

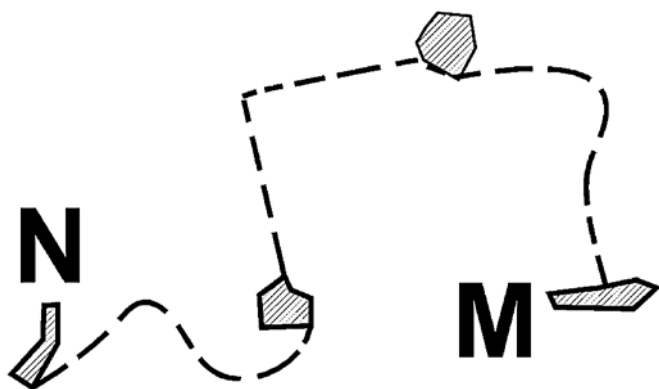


Рис. 5. Метапреобразование решения образа I_{AB}

штрихованными многоугольниками — логистические центры.

Аналогичным образом строят решения для оставшихся образов I_{CD} и I_{EF} . Возможно, что каждый из образов даст достоверное решение. Выбор наилучшего решения является задачей многокритериального сравнения альтернатив и определяется предпочтениями пользователя ГИС.

Заключение

Изложенный в настоящей работе подход обобщает механизм использования опыта принятия решений в среде геоинформационных систем. Использование подхода на практике потребует разработки картографических концептуальных моделей знаний более высокого уровня. На этом уровне аккумулируются результаты критического осмысления принятых экспертом решений. В перспективе это должно привести к распространению опыта на области, смежные с управлением материальными потоками. Дальнейшие исследования могут быть направлены на эффективную реализацию метапреобразований и изучение их поведения в условиях динамики картографической основы ГИС.

Работа выполнена при финансовой поддержке грантов РФФИ № 15-01-00149а и 16-01-00090а.

Список литературы

1. Розенберг И. Н. Спутниковые и геоинформационные технологии в интеллектуальных системах управления // Железнодорожный транспорт. 2013. № 3. С. 28—32.
2. Fang Y., Dhandas V., Arriaga E. Spatial Thinking in Planning Practice. Portland: Portland State University, 2014.
3. Беляков С. Л., Розенберг И. Н. Программные интеллектуальные оболочки геоинформационных систем. М.: Научный мир, 2010. 132 с.
4. Pangbourne K., Alvanides S. Towards intelligent transport geography // Journal of Transport Geography. 2015. N 34. P. 231—232.
5. Belyakov S. L., Bozhenyuk A. V., Belykova M. L., Rozenberg I. N. Model Of Intellectual Visualization Of Geoinformation Service // Proc. 28th European Conference on Modelling and Simulation ECMS 2014. 2014. P. 326—333.
6. Варшавский П. Р., Еремеев А. П. Моделирование рассуждений на основе прецедентов в интеллектуальных системах поддержки принятия решений. // Искусственный интеллект и принятие решений. 2009. № 2. С. 45—47.
7. Визильтер Ю. В., Желтов С. Ю. Меры сходства и метрики сравнения форм мозаичных изображений // Известия Российской академии наук. Теория и системы управления. 2014. № 4. С. 76—89.
8. Иванов Д. А. Управление цепями поставок. СПб.: Изд-во Политехн. ун-та, 2009. 660 с.
9. Цветков В. Я. Ситуационное моделирование в геоинформатике // Информационные технологии. 2014. № 6. С. 64—69.
10. Malczewski J. GIS-based land-use suitability analysis: a critical overview // Progress in Planning. 2004. Vol. 62. P. 3—65.
11. Galton A. Spatial and temporal knowledge representation // Earth Science Informatics. September. 2009. Vol. 2, N 3. P. 169—187.
12. Беляков С. Л., Белякова М. Л., Савельева М. Н. Прецедентный анализ образов в интеллектуальных геоинформационных системах // Информационные технологии. 2013. № 7. С. 22—25.
13. Кузнецов О. П. О концептуальной семантике // Искусственный интеллект и принятие решений. 2012. № 4. С. 32—42.

14. **Осипов Г. С., Панов А. И., Чудова Н. В.** Управление поведением как функция сознания. I. Картина мира и целеполагание // Известия Российской академии наук. Теория и системы управления. 2014. № 4. С. 49.
15. **Вагин В. Н., Головина Е. Ю., Загорянская А. А., Фомина М. В.** Достоверный и правдоподобный вывод в интеллекту-

- альных системах. 2-е изд., испр. и доп. / Под ред. В. Н. Вагина и Д. А. Поспелова. М.: ФИЗМАТЛИТ, 2008. 712 с.
16. **Беляков С. Л., Белякова М. Л., Самойлов Д. С.** Геоинформационный сервис ситуационного центра // Информационные технологии. 2011. № 8. С. 29–32.
17. **Берлянт А. М.** Теория геоизобразжений. М.: Изд-во Московского гос. ун-та им. М. В. Ломоносова, 2006. 262 с.

S. L. Belyakov, DSc., Professor, e-mail: beliacov@yandex.ru;
M. L. Belyakova, Ph. D., Associate Professor, e-mail: mamitabha@yandex.ru;
A. A. Glushkov, Postgraduate Student, e-mail: andrey@glushkov.net;
 Federal State-Owned Autonomy Educational Establishment of Higher Vocational Education
 "Southern Federal University", Taganrog

Images Meta-Transformation when Searching for Reliable Solutions in Intelligent Geographic Information Systems

The article devoted a problem of decision-making at lack of information. Taking decisions implemented with the support geoinformation system. The reliability is a criterion of quality of decisions that forms a geographic information system. We propose a method of construction of solutions based on an experience. An experience presented by cartographic imagery. Each image consists of a center and transformation. The center describes a specific situation. The transformation set possible changes in the center, which do not alter the meaning of precedent in general. In this paper proposed to use meta-transformation images. The essence of meta-transformation is to create a new situations according to a known precedents and describe them on the map. Meta-transformation is a vector. The elements of the vector are the transformations of the image a precedent. A measure of the reliability of the cartographic representation of a situation has been introduced. It is a number of a changes of a center of a image. A metric is a norm of a vector of a meta-transformation. Using of meta-transformation allows a special way to evaluate a semantic proximity of situations and to adapt earlier decisions. The factors which determine the accuracy of the generated solutions has been analyzed. The results of the article can be used for a supply chain management, in developing a various logistics projects. The authors reviewed the example, showing the formation of reliable solutions.

Keywords: decision making, knowledge, case based resoninng, intelligent systems, cartographic visualization, geographic information systems

References

1. **Rozenberg I. N.** Sputnikovye i geoinformatsionnye tekhnologii v intellektual'nykh sistemakh upravleniya [Satellite and GIS technology in intelligent control systems], *Zheleznodorozhnyy transport*, 2013, no. 3, pp. 28–32.
2. **Fang Y., Dhandas V., Arriaga E.** *Spatial Thinking in Planning Practice*, Portland: Portland State University, 2014.
3. **Belyakov S. L., Rozenberg I. N.** *Programmnye intellektual'nye obolochki geoinformatsionnykh sistem* [Software intelligent shell of Geoinformation Systems], Moscow, Nauchnyy mir, 2010, 132 p.
4. **Pangbourne K., Alvanides S.** Towards intelligent transport geography, *Journal of Transport Geography*, 2015, no. 34, pp. 231–232.
5. **Belyakov S. L., Bozhenyuk A. V., Belykova M. L., Rozenberg I. N.** Model Of Intellectual Visualization Of Geoinformation Service, *Proc. 28th European Conference on Modelling and Simulation ECMS 2014*, 2014, pp. 326–333.
6. **Varshavskiy P. R., Ereemeev A. P.** *Poisk resheniya na osnove strukturnoy analogii dlya intellektual'nykh sistem podderzhki prinyatiya resheniy* [Search for solutions on the basis of structural analogy for intelligent decision support systems], *Izvestiya RAN. Teoriya i sistemy upravleniya*, 2005, no. 1, pp. 97–109.
7. **Vizil'ter Ju. V., Zheltov S. Ju.** Mery shodstva i metriki sravneniya form mozaichnykh izobrazhenij [Similarity measures and comparison metrics for image shapes], *Izvestiya Rossijskoj akademii nauk. Teoriya i sistemy upravleniya*, [Journal Of Computer And Systems Sciences International], 2014, vol. 53, no. 4, pp. 542–555.
8. **Ivanov D. A.** *Upravlenie cepjami postavok* [The supply chain management], Sankt-Peterburg, Izd-vo Politehn. un-ta, 2009.
9. **Cvetkov V. Ja.** Situacionnoe modelirovanie v geoinformatike [Situational modeling in Geoinformatics], *Informacionnye tehnologii*, 2014, no. 6, pp. 64–69.
10. **Malczewski J.** GIS-based land-use suitability analysis: a critical overview, *Progress in Planning*, 2004, vol. 62, pp. 3–65.
11. **Galton A.** Spatial and temporal knowledge representation, *Earth Science Informatics*, 2009, vol. 2, no. 3, pp. 169–187.
12. **Beljakov S. L., Beljakova M. L., Savel'eva M. N.** Precedentnyj analiz obrazov v intellektual'nykh geoinformacionnykh sistemah [Case-analysis of images in Intelligent Information Systems], *Informacionnye tehnologii*, 2013, no. 7, pp. 22–25.
13. **Kuznecov O. P.** O konceptual'noj semantike [On the conceptual semantics], *Iskusstvennyj intellekt i prinjatje reshenij*, 2012, no. 4, pp. 32–42.
14. **Osipov G. S., Panov A. I., Chudova N. V.** Upravlenie povedeniem kak funkciya soznaniya. I. Kartina mira i celepolaganie [Management behavior as a function of consciousness. I. The picture of the world, and goal-setting], *Izvestiya Rossijskoj akademii nauk. Teoriya i sistemy upravleniya*, 2014, no. 4, pp. 49.
15. **Vagin V. N., Golovina E. Ju., Zagorjanskaja A. A., Fomina M. V.** *Dostovernyj i pravdopodobnyj vyvod v intellektual'nykh sistemah* [Reliable and credible conclusion in intelligent systems]. 2-е изд., испр. и доп., ed. Vagina V. N. i Pospelova D. A. M.: FIZMATLIT, 2008.
16. **Beljakov S. L., Beljakova M. L., Samojlov D. S.** Геоинформационный сервис ситуационного центра [Geoinformation service of situational center], *Informacionnye tehnologii*, 2011, no. 8, pp. 29–32.
17. **Berlyant A. M.** *Teoriya geoizobrazhenij* [Theory of geoimages], Moscow, Publishing house of Moscow state University named after M. V. Lomonosov, 2006.

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ CRYPTOSAFETY INFORMATION

УДК 004.056.5

И. В. Машкина, д-р техн. наук, проф., e-mail: mashkina.vtzi@gmail.com,

А. Ю. Сенцова, аспирант, ассистент кафедры ВТиЗИ, e-mail: sentsova.alina@yandex.ru
ФГБОУ ВПО "Уфимский государственный авиационный технический университет", г. Уфа

Обеспечение информационной безопасности системы облачных вычислений*

Работа посвящена разработке модели угроз и частной политики информационной безопасности системы облачных вычислений. Модель угроз разрабатывается в виде нечеткой когнитивной карты, которая позволяет выполнить моделирование процессов распространения угроз информационной системе, построенной с использованием технологии облачных вычислений, через эксплуатируемые уязвимости компонентов ее инфраструктуры. Формирование частной политики безопасности основано на использовании модели ролевого разграничения доступа.

Ключевые слова: облачные вычисления, облако сообщества, система облачных вычислений, поставщик облачных услуг, потребитель облачных услуг, модель угроз, нечеткая когнитивная карта, частная политика безопасности, иерархия ролей

Введение

Сегодня в индустрии информационных технологий можно наблюдать стремительные темпы развития информационных систем, построенных на основе технологии облачных вычислений (ИСОТ) [1], однако при этом недостаточно широко освещены проблемы использования облачных сервисов с точки зрения информационной безопасности (ИБ). Вместе с тем использование средств, обеспечивающих функционирование облачных вычислений, позволяет говорить о новых потенциально возможных угрозах информационной безопасности, которые будут являться специфическими для облачных сред.

В NIST [2] описываются четыре модели развертывания облачных вычислений: частное облако (*private*), облако сообщества (*community*), публичное облако (*public*), и гибридное облако (*hybrid*). Различные модели развертывания облаков подразумевают различное размещение контролируемого потребителем периметра безопасности и, следовательно, разный уровень контроля, который подписчики могут осуществлять в отношении ресурсов, доверяемых облаку.

В работе рассматривается облако сообщества, которое представляет собой вычислительное облако, используемое несколькими организациями, решающими общие задачи. Облако сообщества мо-

жет находиться в совместной собственности сообщества, управляться одним или несколькими членами сообщества, а также быть в юрисдикции третьей стороны (поставщика облачных услуг). С точки зрения информационной безопасности облако сообщества можно разделить на несколько систем облачных вычислений (СОБВ) — систем информационного взаимодействия поставщика с конкретным потребителем облачных услуг. В каждой СОБВ обязательно должна быть введена собственная политика информационной безопасности, установлены соответствующие барьеры в целях защиты критичных информационных активов потребителя облачных услуг не только от злоумышленника, неправомерных действий сотрудников служб поставщика облачных услуг и сотрудников потребителя (внутренних нарушителей), но и от неправомерных действий других членов облака сообщества, являющихся в этом случае внешними нарушителями по отношению к СОБВ. Концепция облака сообщества представлена на рис. 1.

Разработка обобщенной архитектуры системы облачных вычислений

Любое вычислительное облако представляет собой систему, построенную на основе клиент-серверной архитектуры. Модель клиент-серверного взаимодействия характеризуется наличием двух взаимодействующих процессов — клиента и сервера. Архитектура облачных вычислений является

* Работа поддержана грантом РФФИ 14-07-00928-а.

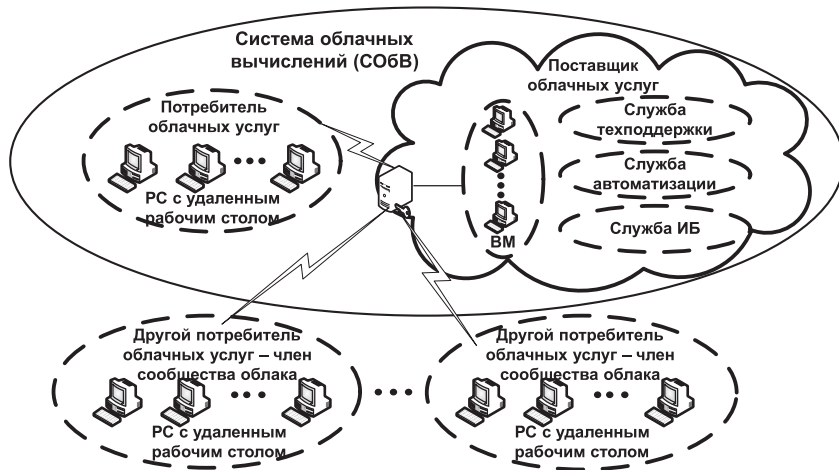


Рис. 1. Концепция облака сообщества и СОБВ

более современной версией известных схем клиент-серверных взаимодействий.

Есть ряд задач, которые не требуют средств виртуализации, однако решение все равно будет считаться облачным. В каждом случае, для каждого

инфраструктуры системы облачных вычислений для облака сообщества, представленная на рис. 2.

При разработке варианта инфраструктуры СОБВ учтены требования для типовой ИСОТ и требования архитектуры безопасности.

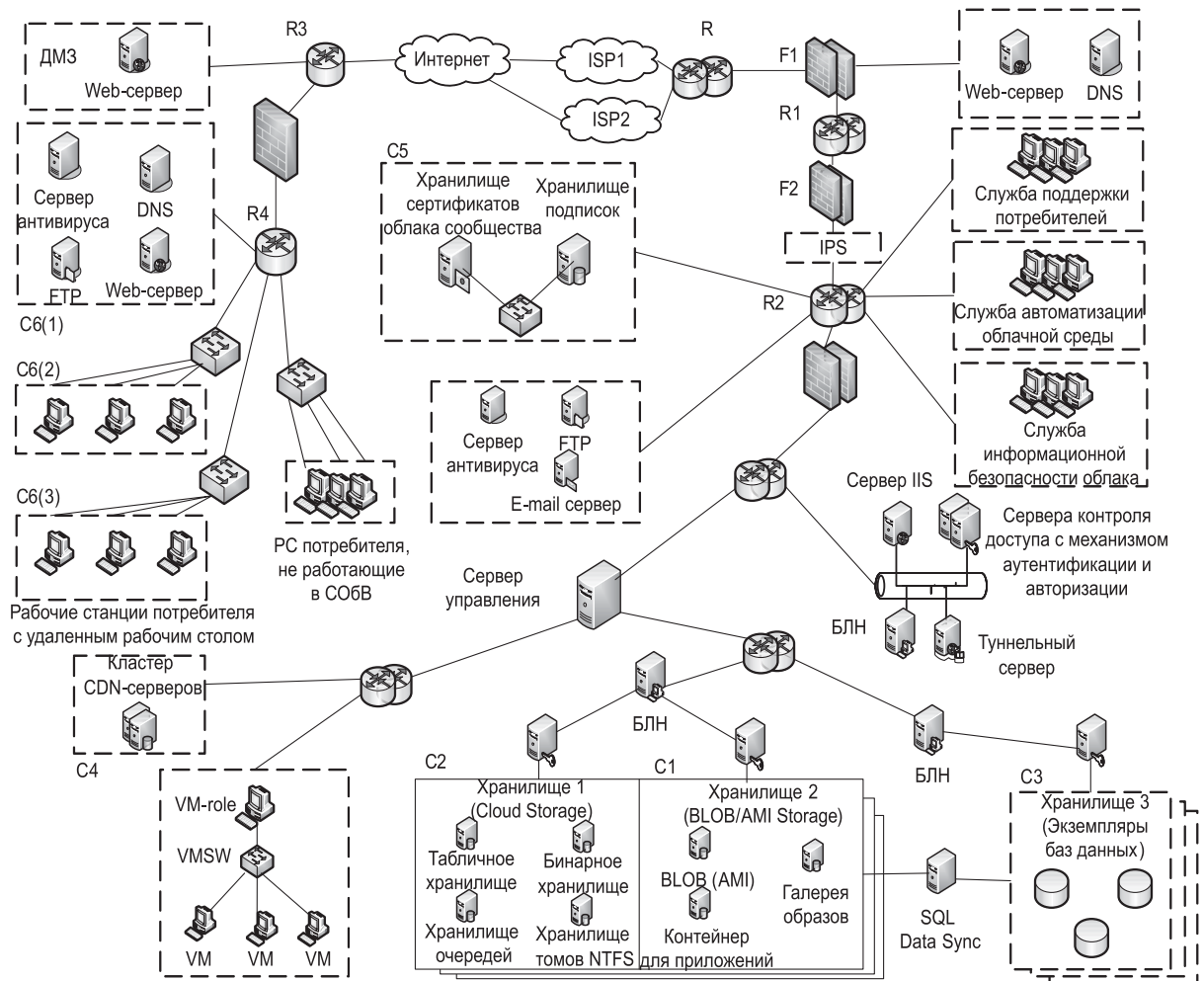


Рис. 2. Модель инфраструктуры системы облачных вычислений

Согласно ГОСТ [4] в состав ИСОТ входят следующие компоненты:

- со стороны потребителя облачных услуг: рабочие станции сотрудников потребителя облачных услуг с удаленным рабочим столом, облачные периферийные устройства и ПО (облачный клиент), сетевое оборудование для осуществления межоблачного взаимодействия;
- со стороны поставщика облачных услуг: пограничный сервер для потребителя (сервер IIS), функциональные серверы для обеспечения бизнес-процессов потребителя облачных услуг, хранилища данных (представлены в архитектуре тремя хранилищами), сетевое оборудование, включая сетевое оборудование для осуществления межоблачного взаимодействия.

Сеть потребителя облачных услуг представлена минимальным набором простейших компонентов инфраструктуры.

Инфраструктура информационной системы поставщика включает в себя *пограничный сервер* для связи с конкретным потребителем, реализованный в виде *сервера IIS*, серверы контроля доступа с механизмом аутентификации и авторизации для входа в рабочую часть вычислительного облака потребителя, *триада службами потребителя облачных услуг*, а также служебными *серверами для выдачи и хранения сертификатов* членов облака сообщества и *данных о подписках* каждого потребителя. Весь трафик аутентификации и авторизации потребителя, а также трафик, содержащий обрабатываемые в облаке критичные активы потребителя облачных услуг, шифруется с помощью туннельного сервера, а доступность сервера IIS гарантируется балансировщиком нагрузки (БЛН).

В рабочей области облака доступ к хранилищам и виртуальным машинам осуществляется с помощью *сервера управления* облаком. *Кластер CDN серверов (Content Delivery Network)* хранит контент сети доставки облака и является отдельно заказываемой функцией облака сообщества для обеспечения более быстрого доступа к наиболее часто используемым данным потребителя.

Администратор потребителя облачных услуг управляет всей сетью виртуальных машин (ВМ) с помощью специализированной ВМ — *VM-role*, к которой через *виртуальный коммутатор* подключены все виртуальные машины потребителя облачных услуг.

Доступ к *хранилищам системы облачных вычислений* осуществляется через БЛН с дополнительной аутентификацией и авторизацией сотрудника потребителя облачных услуг.

Хранилище 1 (Cloud Storage) — хранилище, в котором находятся неструктурированные данные потребителя (документы, видеофайлы, чертежи, схемы) и в том числе тома для подключения к вызываемому сотрудником потребителя образу виртуальной машины. Информация для аутентификации и ав-

торизации пользователя для доступа к данному хранилищу прописана в учетной записи пользователя, под именем которого авторизовался в системе сотрудник потребителя облачных услуг.

Хранилище 2 (BLOB/AMI Storage) — хранилище образов виртуальных машин, которое также включает в себя галерею образов и контейнер для приложений. Информация для аутентификации и авторизации пользователя для доступа к данному хранилищу имеет несколько уровней. При низком уровне доступа сотрудник потребителя имеет доступ только к образу виртуальной машины, с которым ему положено работать по бизнес-процессам. Администратор потребителя имеет полный доступ к галерее образов, содержащей "чистые" образы виртуальных машин, а также к контейнеру приложений, в котором находятся все заказанные потребителем облачных услуг приложения, готовые к установке в образ ВМ. Повышение уровня доступа ко второму хранилищу регламентируется политикой безопасности потребителя облачных услуг и находится в юрисдикции компании потребителя облачных услуг.

Хранилище 3 служит для хранения *экземпляров баз данных* потребителя облачных услуг. Для доступа в хранилище также нужно пройти процедуру аутентификации, которая определит конкретный экземпляр базы/баз данных, они должны быть доступны сотруднику потребителя облачных услуг в соответствии с бизнес-процессами.

Совокупность пограничного сервера, функциональных серверов и хранилищ данных, а также сетевого оборудования для поддержки серверов представляет собой *типовой облачный сервер ИСОТ*.

Уровень оборудования представлен маршрутизаторами, коммутаторами и серверами, *уровень управления* — сервером управления и сервисами безопасности, такими как, например, межсетевые экраны, IPS, туннельные серверы шифрования трафика и серверы контроля доступа с механизмами аутентификации и авторизации потребителя облачных услуг.

Типовая *демилитаризованная зона* поставщика облачных услуг включает в себя: публичный web-сервер, DNS-сервер, сервер приложений для управления информационными ресурсами web-портала. В демилитаризованной зоне должна храниться и обрабатываться только общедоступная открытая информация, что отражено в инфраструктуре СОБВ.

Чтобы минимизировать риски потребителя и поставщика в процессе реализации облачных сервисов, внутренний web-сервер для клиент-серверного взаимодействия (сервер IIS) должен быть размещен в другом сетевом сегменте, как и сервер CDN, обеспечивающий временное хранение ресурсов потребителя облачных услуг и позволяющий уменьшить время доступа пользователей к кэшированным данным. В соответствии с требованиями

архитектуры безопасности следует обеспечить сегрегацию этих серверов от общедоступных, располагая их во внутренней защищенной подсети поставщика облачных услуг (во внутренней рабочей области облака).

Разработка частной политики безопасности системы облачных вычислений

Многими экспертами отмечается, что потребитель облачных услуг имеет тот уровень защищенности в облачной среде, который обеспечивается поставщиком [8,9]. Однако для защищенности информации в СОБВ необходимо обратить внимание на разработку политики безопасности всей системы облачных вычислений и гарантированно защитить пользовательские рабочие места на стороне потребителя.

Политика информационной безопасности организации — совокупность руководящих принципов, правил, процедур и практических приемов в области безопасности, которые регулируют управление, защиту и распределение ценной информации [10].

Политика информационной безопасности разрабатывается на основе сведений о конкретных технологиях обработки информационных активов предприятия, об используемых в данной системе информационных потоках, реализующих деловые процессы. Для устранения угроз, связанных с неопределенностью при распределении ответственности, поставщику облачных услуг необходимо тщательно прорабатывать частную политику безопасности, а именно политику управления доступом, формализованную в виде матрицы доступа. Для составления матрицы доступа необходимо сформировать иерархическую структуру, составленную на основе тщательного анализа возможных ролей поставщика и потребителя облачных услуг.

Как отмечается в работе [11], политика безопасности любой информационной системы (ИС) при разработке информационно-безопасных технологий состоит из множества *частных политик*, направленных на конкретные аспекты безопасности ИС. Частные политики безопасности, детализирующие положения политики ИБ, формируются на основе принципов, требований и задач, определенных в политике информационной безопасности, с учетом дополнительной классификации активов и угроз, определения владельцев критичных активов, анализа, оценки рисков и возможных последствий реализаций угроз в границах области действия регламентируемой области или технологии [12].

Актуальность разработки частных политик информационной безопасности объясняется необходимостью планирования и управления ИБ на всех этапах жизненного цикла ИСОТ и СОБВ. В случае разработки частной политики безопасности ИСОТ необходимо учитывать специфику межоблачных взаимодействий между поставщиком и потребителем облачных услуг. С помощью правильно состав-

ленной политики ИБ можно обеспечить безопасное, доверенное и адекватное управление системой облачных вычислений, поддержку непрерывности межоблачного взаимодействия, повышение уровня доверия потребителя к поставщику облачных услуг и, как следствие, минимизировать риски нарушения информационной безопасности в СОБВ, т. е. повысить защищенность информации при обработке ее в СОБВ.

В ходе исследований разработана модель политики информационной безопасности СОБВ, включающая перечень частных политик ИБ с их детализацией:

- политика идентификации и аутентификации субъектов доступа и объектов доступа;
- *политика управления доступом субъектов доступа к объектам доступа*;
- политика ограничения программной среды;
- политика менеджмента инцидентов ИБ;
- политика антивирусной защиты СОБВ;
- политика средств обнаружения вторжений;
- политика по контролю (анализу) защищенности;
- политика обеспечения целостности ПО СОБВ;
- политика информационной безопасности облачного сервера;
- политика межсетевое экранирование;
- политика централизованного управления СОБВ.

Таким образом, для обеспечения защищенности необходимо создание частных политик, которые должны неукоснительно соблюдаться как поставщиком, так и потребителем облачных услуг.

Приведем методику разработки одной из частных политик ИБ СОБВ — политики управления доступом субъектов доступа к объектам доступа, которая строится на основе ролевой модели и детализирована следующим образом:

- реализация необходимых методов и правил ограничения доступа в ходе межоблачного взаимодействия;
- разделение обязанностей, полномочий (ролей), всех линий администраторов поставщика;
- назначение минимально необходимых правил и привилегий пользователям СОБВ и сотрудникам поставщика облачных услуг;
- разграничение доступа к объектам, расположенным за пределами виртуальных машин.

Использование ролевой модели требует решения задач выявления множества субъектов, специфичных объектов доступа для такого объекта защиты как СОБВ, построения иерархических схем ролей с учетом функций субъектов как со стороны поставщика, так и со стороны потребителя облачных услуг. Далее здесь под частной политикой безопасности СОБВ следует понимать политику управления доступом.

Разработка частной политики безопасности СОБВ позволяет избежать угроз информационной безопасности, связанных с неопределенностью от-

ответственности в системе облачных вычислений, реализация которых способна привести к существенным разногласиям между поставщиком и потребителем облачных услуг по вопросам, связанным с определением их прав и обязанностей.

Роли задаются для различных должностей в облаке, и пользователи соотносятся к ролям, основанным на ответственности и профессионализме. Пользователи могут быть переназначены на другую роль. Роли могут наделяться новыми разрешениями по мере подключения новых приложений и заказе потребителем новых услуг, разрешения могут отбираться у ролей, когда это необходимо потребителю или поставщику услуг.

Установим для информационных объектов в СОБВ следующие уровни конфиденциальности:

Таблица 1
Множество информационных объектов доступа СОБВ

Обозначение	Наименование	Уровень конфиденциальности
o1	Сайт поставщика облачных услуг	ОИ
o2	Множество логинов и паролей личных кабинетов сотрудников потребителя облачных услуг	К
o3 (1)	Виртуальные машины отдела потребителя облачных услуг, осуществляющего работу по проекту 1	СК
o3 (i)	Виртуальные машины отдела потребителя облачных услуг, осуществляющего работу по проекту <i>i</i>	СК
o4 (1)	Информационные ресурсы по проекту 1, хранящиеся в облачном хранилище	СК
o4 (i)	Информационные ресурсы по проекту <i>i</i> , хранящиеся в облачном хранилище	СК
o5	Файлы СОБВ, относящиеся к конфигурированию собственных виртуальных машин конкретным потребителем облачных услуг	СК
o6 (1)	Файлы СОБВ, относящиеся к управлению внутриоблачным пространством поставщиком облачных услуг	СК
o6 (2)	Файлы СОБВ, относящиеся к сервисам безопасности поставщика облачных услуг	СК
o7	Данные о серверном времени, скорости доступа и обработки данных, объеме хранимых в хранилище данных	К
o8	Данные о фактическом распределении доступа в едином пуле облака	СК
o9	Объем предоставленных потребителю услуг	К
o10 (1)	Информационные ресурсы по проекту 1, хранящиеся на стороне потребителя облачных услуг	К
o10 (i)	Информационные ресурсы по проекту <i>i</i> , хранящиеся на стороне потребителя облачных услуг	К
o11 (1)	Экземпляры отдела, работающего по проекту 1, запускаемые в физической операционной среде (физическом кластере поставщика облачных услуг)	СК
o11 (i)	Экземпляры отдела, работающего по проекту <i>i</i> , запускаемые в физической операционной среде (физическом кластере поставщика облачных услуг)	СК

ОИ — открытая информация, К — конфиденциально, СК — строго конфиденциально

В результате исследований выявлено множество информационных объектов для системы облачных вычислений (табл. 1).

Множество ролей пользователей (субъектов доступа) системы облачных вычислений, сформированное в ходе исследований, представлено в табл. 2.

Таблица 2

Множество субъектов доступа в СОБВ

Обозначение	Наименование	Уровень доступа
L1	Технический директор поставщика облачных услуг	СК
LT1	Сотрудник первой линии техподдержки поставщика облачных услуг	К
LT2	Сотрудник второй линии техподдержки поставщика облачных услуг	СК
LT3	Сотрудник третьей линии техподдержки поставщика облачных услуг	К
S1	Руководитель службы автоматизации ИСОТ	СК
S2	Главный специалист по ИСОТ	СК
S3	Администратор инфраструктуры ИСОТ	К
S4	Эксперт по виртуализации в облачных вычислениях	К
AV1	Начальник службы безопасности облачного поставщика	СК
AV2	Специалист по защите программного обеспечения и платформ поставщика услуги SaaS	К
AV3	Специалист по защите облачной инфраструктуры поставщика услуги SaaS	К
AV4	Специалист по защите кластера физических серверов поставщика	К
P1	Технический директор потребителя облачных услуг	СК
P2, P3	Руководители подразделений потребителя облачных услуг, осуществляющих эксплуатацию СОБВ в соответствии с бизнес-процессами	СК
A1	Начальник отдела автоматизации и безопасности потребителя	СК
A2	Администратор безопасности потребителя облачных услуг	СК
A3	Работник, осуществляющий интеграцию и сопровождение SaaS ИСОТ (менеджер ИСОТ)	СК
A4	Администратор штатных средств защиты потребителя	К
P4, P5	Сотрудники потребителя облачных услуг, осуществляющие эксплуатацию СОБВ по проекту 1 в соответствии с бизнес-процессами предприятия	К
P6, P7	Сотрудники потребителя облачных услуг, осуществляющие эксплуатацию СОБВ по проекту 2 в соответствии с бизнес-процессами предприятия	К
P8, P9	Сотрудники подразделений потребителя облачных услуг, работающие по проектам 1 и 2 соответственно, не имеющие права эксплуатировать СОБВ в соответствии с бизнес-процессами	ОИ
P10	Сотрудники подразделения потребителя облачных услуг, не работающие по проектам 1 и 2, и не имеющие права эксплуатировать СОБВ в соответствии с бизнес-процессами	ОИ

Так как система облачных вычислений — это система, в которой взаимодействуют поставщик и потребитель облачных услуг, в работе предложено модифицировать ролевую модель разграничения доступа таким образом, чтобы каждая из представленных сторон (потребитель и поставщик) имела свою максимальную роль в иерархии, в отличие от известной ролевой модели разграничения доступа, где максимальная роль в иерархии может быть только одна. Для поставщика облачных услуг примем, что максимальной ролью является роль технического директора поставщика (L1), для потребителя, соответственно, — технического директора потребителя облачных услуг (P1).

В общем случае иерархия ролей потребителя будет многоуровневой и распределенной. На рис. 3 представлена разработанная иерархическая структура ролей для множества информационных субъектов и объектов при $i = 2$.

В примере, проиллюстрированном рис. 3, потребитель облачных услуг имеет два подразделения, осуществляющих эксплуатацию СОБВ в соответствии с бизнес-процессами. В каждом из подразделений минимальная роль отводится сотрудникам потребителя облачных услуг, не имеющим права эксплуатировать СОБВ в соответствии с бизнес-процессами (P8, P9, P10), а максимальная роль — руководителям подразделений потребителя облачных услуг, осуществляющих эксплуатацию СОБВ в соответствии с бизнес-процессами (P2, P3). Кроме того, в иерархии учтено, что два подразделения потребителя могут выполнять работу в СОБВ над разными проектами (проекты 1 и 2), которые, в соответствии с бизнес-процессами не имеют общих и пересекающихся ресурсов и активов. Таким образом, сотрудники подразделения, работающего по проекту 1, не имеют доступа к информационным ресурсам и активам СОБВ подразделения, работающего по проекту 2, и наоборот.

В иерархии потребителя облачных услуг, помимо двух подразделений, работающих по проектам 1 и 2, есть третье подразделение, отвечающее за автоматизацию и информационную безопасность компа-

нии. Максимальная роль в этом подразделении отводится начальнику отдела автоматизации и безопасности потребителя облачных услуг (A1), а минимальная роль — администратору штатных средств защиты (A4), под которыми понимаются традиционные средства защиты, не входящие в систему безопасности облачной среды потребителя.

Иерархия поставщика облачных услуг, где максимальная роль отведена техническому директору поставщика (L1), состоит из трех служб-отделов: службы поддержки потребителей облачных услуг; службы автоматизации облачной среды; службы информационной безопасности поставщика облачных услуг.

Служба поддержки потребителей состоит из трех линий (LT1, LT2, LT3) поддержки, которые взаимодействуют напрямую с потребителями облачных услуг и помогают конкретному поставщику решать возникающие вопросы и проблемы в реальном масштабе времени. В ходе исследований были выделены три возможные линии технической поддержки облаков [13]:

- сотрудники *первой линии* техподдержки поставщика облачных услуг (LT1), которые при обращении к ним потребителя ликвидируют технические сбои в инфраструктуре, влияющие на предоставляемые пользователям сервисы; данные сотрудники не обладают высокими привилегиями в СОБВ, и не имеют доступа к сервисам безопасности СОБВ;
- сотрудники *второй линии* техподдержки поставщика облачных услуг (LT2) — группа специалистов высокого профиля, которая обладает достаточной компетенцией и способна решать проблемы как с инфраструктурой СОБВ, так и с ее сервисами;
- сотрудники *третьей линии* техподдержки поставщика облачных услуг (LT3) являются сотрудниками разработчика и производителя технологии облачных вычислений (Amazon, Google, Microsoft).

Служба автоматизации облачной среды ответственна за разработку и процесс интеграции в SaaS

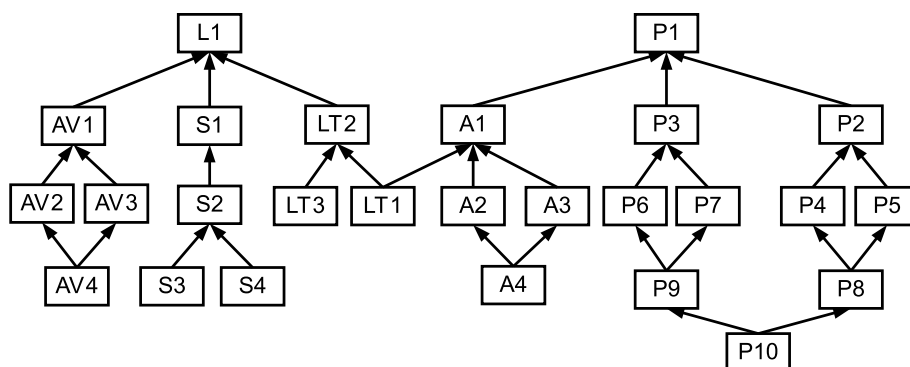


Рис. 3. Иерархическая структура ролей в СОБВ

облачных вычислений со стороны потребителя облачных услуг; сотрудники службы занимаются вопросами оптимального управления облачными сервисами в условиях существующих ограничений сети потребителя облачных услуг. Максимальную роль в данной службе будет иметь руководитель службы автоматизации ИСОТ (S1), а минимальными ролями будут обладать администратор инфраструктуры ИСОТ (S3) и эксперт по виртуализации в облачных вычислениях (S4).

Служба информационной безопасности поставщика облачных услуг отвечает за безопасность облачной среды со стороны поставщика облачных услуг. В данной службе роли распределены на три составляющие защиты облака: защиты программного обеспечения и платформ поставщика услуги SaaS (роль AV2); защиты облачной инфраструктуры поставщика услуги SaaS (роль AV3); защиты кластера физических серверов поставщика облачных услуг (роль AV4). Максимальной в данной службе будет роль начальника службы безопасности облачного поставщика (AV1), а минимальной — специалиста по защите кластера физических серверов поставщика облачных услуг (AV4).

Поставщик облачных услуг ни в коем случае не должен обладать какими-либо правами доступа к информации, которую обрабатывает потребитель в облаке. К такой информации относятся виртуальные машины потребителя, информационные ресурсы, хранящиеся в облачном хранилище, информационные ресурсы, хранящиеся на стороне потребителя облачных услуг, и экземпляры, виртуальные машины, запускаемые в физическом кластере поставщика, конфигурационная и управляющая информация потребителя. Администраторы безопасности потребителя облачных услуг конфигурируют собственные виртуальные машины, и конфигурационные файлы, относящиеся к конкретному потребителю, должны быть скрыты от служб поставщика. Таким образом, в СОБВ используются IP-адреса, каждый из которых ассоциируется с учетной записью клиента, а не с конкретным экземпляром виртуальной машины. Для запуска виртуальной машины ей должен быть присвоен атрибут, указывающий, какие учетные записи облачного web-сервиса имеют право запускать конкретную виртуальную машину.

Одновременно с этим поставщик обладает правами на управление и конфигурирование внутри-облачного пространства и собственных сервисов безопасности, чтобы осуществить защиту данных конкретного потребителя не только от злоумышленников, но и от других потребителей услуг облака.

Таким образом, руководитель подразделения и сотрудники потребителя облачных услуг, осуществляющие эксплуатацию СОБВ по проекту 1 в соответствии с бизнес-процессами предприятия, могут читать и запускать виртуальные машины (o3) по проекту 1 и имеют полный доступ на собственные экземпляры, запускаемые в физической операционной среде (o11). Соответственно, руководитель подразделения и сотрудники потребителя облачных услуг, осуществляющие эксплуатацию СОБВ по проекту 2 в соответствии с бизнес-процессами предприятия, могут читать и запускать виртуальные машины (o3) по проекту 2 и имеют полный доступ на собственные экземпляры, запускаемые в физической операционной среде (o11). Вносить

изменения в настройки конфигурационных файлов образов виртуальных машин и экземпляров, запускаемых в физической операционной среде, по обоим проектам могут администратор безопасности потребителя облачных услуг и менеджер ИСОТ. Начальник отдела автоматизации и безопасности потребителя облачных услуг и технический директор облака обладают полными правами по отношению к образам и экземплярам проектов своей организации.

Информационные ресурсы по проектам, хранящиеся в облачном хранилище (o4), и информационные ресурсы по проектам, хранящиеся на стороне потребителя облачных услуг (o10), доступны с полными правами руководителю соответствующего подразделения и сотрудникам потребителя облачных услуг, осуществляющим эксплуатацию СОБВ по заданному проекту в соответствии с бизнес-процессами предприятия, а также техническому директору потребителя облачных услуг.

К файлам СОБВ, относящимся к конфигурированию собственных виртуальных машин конкретным потребителем облачных услуг (o5), имеют полные права доступа менеджер ИСОТ, начальник отдела автоматизации и безопасности потребителя и технический директор потребителя.

Правами на чтение множества логинов и паролей личных кабинетов сотрудников потребителя облачных услуг (o2) обладают сотрудники потребителя облачных услуг, осуществляющие эксплуатацию СОБВ в соответствии с бизнес-процессами предприятия. Данное множество доступно для специалиста по защите программного обеспечения и платформ и для начальника службы безопасности облачного поставщика с правом вносить правки. Полными правами на доступ ко множеству логинов и паролей обладают руководители подразделений потребителя, осуществляющих эксплуатацию СОБВ в соответствии с бизнес-процессами, начальник отдела автоматизации и безопасности потребителя и технический директор потребителя.

Читать сайт поставщика облачных услуг (o1) могут все сотрудники потребителя облачных услуг без исключения. Кроме того, правами только на чтение сайта обладают следующие сотрудники поставщика облачных услуг: эксперт по виртуализации, специалист по защите кластера физических серверов, специалист по защите облачной инфраструктуры, администратор инфраструктуры ИСОТ, специалисты первой и третьей линии техподдержки потребителя облачных услуг, специалист по защите программного обеспечения и платформ, начальник службы безопасности облачного поставщика. Полные права на доступ к сайту имеют главный специалист ИСОТ, руководитель службы автоматизации ИСОТ, сотрудник второй линии техподдержки потребителя облачных услуг и технический директор поставщика облачных услуг.

Данные об объеме предоставленных потребителю услуг (о9) доступны по чтению руководителям подразделений потребителя, осуществляющих эксплуатацию СОБВ в соответствии с бизнес-процессами, сотрудникам службы автоматизации ИСОТ и техническому директору потребителя. Кроме того, прочесть эти данные могут специалист по защите облачной инфраструктуры, начальник службы безопасности облачного поставщика, сотрудник первой линии техподдержки, сотрудники службы автоматизации ИСОТ. Правами на чтение и на внесение правок в объем предоставленных потребителю услуг имеют только сотрудник второй линии техподдержки, руководитель службы автоматизации ИСОТ и технический директор поставщика облачных услуг.

Данные о серверном времени, скорости доступа и обработки данных, а также объеме хранимых в облачном хранилище данных (о7) доступны по чтению руководителям подразделений потребителя, осуществляющих эксплуатацию СОБВ в соответствии с бизнес-процессами, администратору штатных средств защиты потребителя и администратору безопасности потребителя. Со стороны поставщика эти данные доступны по чтению для специалиста по защите облачной инфраструктуры, начальнику службы безопасности облачного поставщика, главному специалисту ИСОТ, администратору инфраструктуры ИСОТ и эксперту по виртуализации в облачных вычислениях. Полный доступ к объекту о7 (см. табл. 1) имеют менеджер ИСОТ, начальник отдела автоматизации и безопасности потребителя облачных услуг и технический директор потребителя, а также руководитель службы автоматизации облачной среды, сотрудники первой и второй линии техподдержки потребителя облачных услуг и технический директор поставщика облачных услуг.

Данные о фактическом распределении доступа в едином пуле облака (о8) доступны по чтению со стороны поставщика специалисту по защите облачной инфраструктуры, начальнику службы безопасности облачного поставщика и эксперту по виртуализации в облачных вычислениях. Полный доступ к этим данным имеет руководитель и главный специалист службы автоматизации ИСОТ, сотрудники первой и второй линии техподдержки потребителя облачных услуг и технический директор поставщика облачных услуг. Так как начальник отдела автоматизации и безопасности потребителя, а также технический директор потребителя наследуют все права сотрудника первой линии техподдержки, который в свою очередь является сотрудником поставщика облачных услуг, то они также имеют полный доступ к данным о фактическом распределении доступа в едином пуле облака.

Конфигурационные файлы внутриоблачного пространства и файлы поставщика, содержащие данные о конфигурировании собственных средств

безопасности поставщика (об), должны быть закрыты для доступа любому сотруднику потребителя облачных услуг в целях повышения защищенности всей системы облачных вычислений. Вносить правки в файлы СОБВ, относящиеся к управлению внутриоблачным пространством поставщиком облачных услуг, может сотрудник третьей линии техподдержки потребителя, а по чтению они доступны специалисту по защите программного обеспечения и платформ поставщика, а также начальнику отдела безопасности поставщика. Полный доступ к файлам управления внутриоблачным пространством имеют сотрудники службы автоматизации (руководитель, главный специалист, администратор инфраструктуры и эксперт по виртуализации), сотрудники второй линии техподдержки и технический директор поставщика облачных услуг.

К файлам СОБВ, относящимся к сервисам безопасности поставщика облачных услуг, имеют полный доступ все сотрудники службы безопасности поставщика облачных услуг (начальник службы, специалист защиты ПО и платформ, специалист по защите инфраструктуры и специалист по защите кластера физических серверов), а также технический директор поставщика облачных услуг.

Таким образом, с учетом приведенных выше условий и ограничений разработана матрица доступа ролей пользователей (субъектов доступа) к множеству объектов доступа СОБВ (табл. 3).

Разработка модели угроз СОБВ в виде нечеткой когнитивной карты

Обеспечение защищенности СОБВ при разумных вложениях является серьезной актуальной проблемой как для потребителя, так и для поставщика облачных услуг. Как обеспечить защищенность информации, обрабатываемой в облаке? Как оценить уровень риска нарушения безопасности информации, который обеспечивается поставщиком при предоставлении потребителю облачных услуг? Эти вопросы возникают как при проектировании системы обеспечения информационной безопасности инфраструктур поставщика и потребителя, так и в процессе функционирования СОБВ.

Для решения проблемы управления рисками нарушения ИБ в СОБВ необходимо идентифицировать внешние и внутренние факторы, влияющие на риск, и оценить уровень риска в количественном выражении [5].

В рамках данных исследований СОБВ во внимание принимаются угрозы *несанкционированного доступа*, в результате которых происходит получение информации заинтересованным субъектом (злоумышленником или пользователем-нарушителем облака сообщества), с нарушением прав и правил, т. е. угрозы, связанные с нарушением частной политики безопасности.

Матрица прав доступа ролей пользователей СОБВ

	o1	o2	o3 (1)	o3 (2)	o4 (1)	o4 (2)	o5	o6 (1)	o6 (2)	o7	o8	o9	o10 (1)	o10 (2)	o11 (1)	o11 (2)
L1	rw	w	—	—	—	—	—	rw	rw	rw	rw	rw	—	—	—	—
LT2	rw	w	—	—	—	—	—	rw	—	rw	rw	rw	—	—	—	—
LT1	r	—	—	—	—	—	—	—	—	rw	rw	r	—	—	—	—
LT3	r	—	—	—	—	—	—	w	—	—	—	—	—	—	—	—
S1	rw	—	—	—	—	—	—	rw	—	rw	rw	rw	—	—	—	—
S2	rw	—	—	—	—	—	—	rw	—	r	rw	r	—	—	—	—
S3	r	—	—	—	—	—	—	rw	—	r	—	r	—	—	—	—
S4	r	—	—	—	—	—	—	rw	—	r	r	r	—	—	—	—
AV1	r	w	—	—	—	—	—	r	rw	r	r	r	—	—	—	—
AV2	r	w	—	—	—	—	—	r	rw	—	—	—	—	—	—	—
AV3	r	—	—	—	—	—	—	—	rw	r	r	r	—	—	—	—
AV4	r	—	—	—	—	—	—	—	rw	—	—	—	—	—	—	—
P1	r	rw	rwe	rwe	rw	rw	rw	—	—	rw	rw	r	rw	rw	rw	rw
A1	r	rw	rw	rw	—	—	rw	—	—	rw	rw	—	—	—	rw	rw
A3	r	rw	—	w	—	—	rw	—	—	rw	—	—	—	—	w	w
A2	r	rw	—	w	—	—	—	—	—	r	—	—	—	—	w	w
A4	r	rw	—	—	—	—	—	—	—	r	—	—	—	—	—	—
P2	r	rw	re	—	rw	—	—	—	—	r	—	r	rw	—	rw	—
P4,5	r	r	re	—	rw	—	—	—	—	—	—	—	rw	—	rw	—
P8	r	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
P3	r	rw	—	re	—	rw	—	—	—	r	—	r	—	rw	—	rw
P6,7	r	r	—	re	—	rw	—	—	—	—	—	—	—	rw	—	rw
P9	r	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

При решении проблемы оценивания рисков нарушения ИБ в СОБВ предложено осуществить идентификацию угроз, реализуемых за счет использования уязвимостей компонентов инфраструктуры и сервисов безопасности. Любая реализация системы обеспечения информационной безопасности — это схема, содержащая в себе требования к защите объекта и модель угроз, отражающую защищаемую среду и состояние оборонительных барьеров.

В работе приведем результаты построения модели угроз в виде нечетких когнитивных карт (НКК). Нечеткие когнитивные карты были разработаны с учетом всех возможных источников угроз, объектов атак и уязвимостей компонентов инфраструктуры СОБВ, расположенных на путях распространения атак.

Когнитивная карта иллюстрирует возможные пути угроз, реализуемых потенциальным злоумышленником и нарушителем, позволяет провести анализ несанкционированного проникновения к информационным ресурсам: показывает, откуда осуществляется атакующее воздействие, какие информационные объекты могут являться целями проведения атак, какими средствами защищена информационная система, какие уязвимости могут быть использованы во время проведения атаки.

Модель угроз в СОБВ, разработанная на основе построения нечетких когнитивных карт, адекватна объекту защиты и позволяет обеспечить оценку уровней угроз. Преимуществом использования НКК в данном приложении в сравнении с какими-либо другими методами является возможность учесть инфраструктуру СОБВ, организовать знания об особенностях распространения угроз в СОБВ, а также формализовать численно неизмеримые факторы, такие как вероятности угроз.

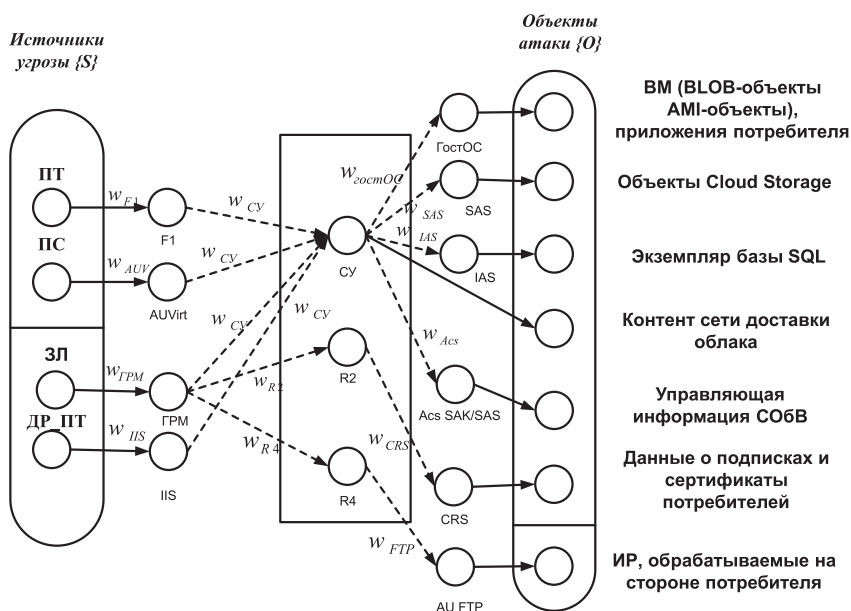


Рис. 4. Обобщенная модель угроз в виде НКК

На рис. 4 представлена обобщенная модель угроз в СОБВ в виде нечеткой когнитивной карты, в которой приведены начальные и конечные уязвимости компонентов инфраструктуры системы облачных вычислений на пути распространения атак.

На рис. 4 использованы следующие обозначения: ПТ — субъект доступа — сотрудник-нарушитель политики управления доступом компании потребителя облачных услуг; ПС — субъект доступа — сотрудник службы поставщика облачных услуг; ДР_ПТ — субъект доступа — сторонний нарушитель — субъект, являющийся другим потребителем облачных услуг облака сообщества (или запущенные им процессы); ЗЛ — субъект доступа — злоумышленник — субъект, не являющийся потребителем облачных услуг (или запущенные им процессы); w — уровни уязвимостей компонентов инфраструктуры СОБВ (веса промежуточных концептов); R — маршрутизатор; ПIS — пограничный сервер; AUVirt, SAK, SAS, IAS — серверы аутентификации облака; FTP — файловый сервер; ГостОС — гостевая операционная система; СУ — сервер управления. Оценки уровней уязвимостей получены с использованием методики CVSS и международной базы данных уязвимостей (NVD) [6].

Использование нечетких когнитивных карт позволяет выполнить моделирование процессов распространения угроз в СОБВ через используемые уязвимости ее компонентов. При этом полученная модель обладает свойством *наглядности* и простотой *понимания и перевода* содержательного знания эксперта на математический язык.

Использование НКК для моделирования угроз позволяет провести численную оценку уровня риска по методу работы [7], т. е. оценить уровень защищенности.

Анализ визуализированных путей реализации угроз служит основой для обоснования выбора средств защиты — барьеров на путях реализации угроз (в том числе в виртуальной среде), снижающих значение риска нарушения ИБ до приемлемого уровня, что позволит обеспечить требуемый уровень защищенности СОБВ.

Полученные результаты

Сформированные в виде матрицы правила разграничения доступа применяются при настройке средств контроля доступа в системе облачных вычислений и для определения полномочий прав пользователей (или запущенных ими процессов) на осуществление тех или иных процедур над защищенными данными. В СОБВ используются IP-адреса, каждый из которых ассоциируется с учетной записью клиента облачных вычислений. Для запуска виртуальной машины каждому IP-адресу должны быть присвоены соответствующие атрибуты, указывающие, какие учетные записи облачного

web-сервиса имеют право запускать ту или иную конкретную виртуальную машину.

Модель угроз, построенная на основе нечетких когнитивных карт, позволяет выполнить моделирование процессов распространения угроз информационной системе через используемые уязвимости ее компонентов. При этом полученная модель обладает свойством *наглядности* и простотой *понимания*. Анализ визуализированных путей реализации угроз служит основой для обоснования выбора средств защиты — барьеров на пути реализации угроз, снижающих значение риска нарушения ИБ до приемлемого уровня.

Заключение

При разработке частной политики управления доступом в СОБВ предлагается ввести в иерархию две максимальные роли: одну со стороны поставщика потребителя облачных услуг (роль технического директора поставщика облачных услуг), другую со стороны потребителя облачных услуг (роль технического директора потребителя облачных услуг), которые имели бы одновременно и *максимально необходимую роль* в собственном подразделении облака сообщества. *Достоинством* предложенного в работе подхода к разработке частной политики информационной безопасности в системе облачных вычислений на основе модели ролевого разграничения доступа является возможность *исключения* прав *суперпользователя*, который может напрямую обращаться к результирующим потокам данных потребителя облачных услуг и управлять всеми конфигурационными файлами системы облачных вычислений.

Соблюдение требований частной политики безопасности СОБВ позволит существенно снизить риски использования облачных вычислений потребителем облачных услуг и, как следствие, увеличить доверие потенциальных потребителей к ИСОТ.

Разработанная с учетом частной политики безопасности модель преднамеренных (целенаправленных) угроз нарушения информационной безопасности в системе облачных вычислений, основанная на построении нечетких когнитивных карт, позволяет учесть угрозы и уязвимости, связанные с динамической масштабируемостью, консолидацией вычислительных ресурсов, возможностью самообслуживания потребителя облачных услуг, а также учесть такой источник угроз, как другой потребитель облачных услуг.

Список литературы

1. Барский А.Б., Желенков Б. В. Средства оптимизации информационного взаимодействия ресурсных процессоров для минимизации времени "облачных" вычислений // Информационные технологии. 2016. Т. 22, № 1. С. 14—21.
2. NIST. Официальный сайт. URL: <http://www.nist.gov/itl/cloud/> (дата обращения: 28.03.2016).

3. **Авраменко А.** Облачные вычисления. Взгляд из IBM // *Jet Info*, 2010, № 10, С. 63–75.
4. **Защита информации.** Требования по защите информации, обрабатываемой с использованием технологий "Облачных вычислений" ГОСТ РХХХХХ-20XX (проект, первая редакция). URL: <http://docs.cntd.ru/document/1200102839> (дата обращения 23.03.2016).
5. **Глушенко С. А., Долженко А. И.** Система поддержки принятия решений нечеткого моделирования рисков информационной безопасности организации // *Информационные технологии*, 2015, Т. 21, № 1, С. 68–74.
6. **Международная база данных уязвимостей.** Официальный сайт. URL: <https://nvd.nist.gov/> (дата обращения: 28.03.2016).
7. **Гузайров М. Б., Машкина И. В., Степанова Е. С.** Построение модели угроз с помощью нечетких когнитивных карт на основе сетевой политики безопасности // *Безопасность информационных технологий*, 2011, № 2, С. 37–49.
8. **Шаньгин В. Ф.** Защита информации в компьютерных сетях и сетях. М.: ДМК Пресс, 2012. 592 с.
9. **Демурчев Н. Г., Ищенко С. О.** Проблемы обеспечения информационной безопасности при переходе на облачные вы-

числения // *Материалы XI Международной научно-практической конференции "Информационная безопасность"*, Ч. 1. Таганрог: Изд-во ТТИ ЮФУ, 2010. С. 147–151.

10. **ГОСТ Р 50922—2006.** Защита информации. Основные термины и определения М.: Стандартинформ, 2007. 12 с.

11. **Варлатая С. К., Шахинова М. В.** Анализ методов описания политики безопасности при разработке информационно-безопасных технологий. // *Доклады ТУСУР*, Ч. 1. Аудит безопасности. Июнь 2010. № 1 (21). С. 10–13.

12. **РС БР ИББС-2.0—2007.** Рекомендации в области стандартизации Банка России. Обеспечение информационной безопасности организаций банковской системы Российской Федерации. Методические рекомендации по документации в области обеспечения информационной безопасности в соответствии с требованиями СТО БР ИББС-1.0. М.: Стандартинформ, 2007. 15 с.

13. **Методология организации технической поддержки.** URL: <http://msbro.ru/index.php/archives/2717> (дата обращения 28.03.2016).

I. V. Mashkina, DSc, Professor, mashkina.vtzi@gmail.com
A. Yu. Sentsova, Postgraduate Student, sentsova.alina@yandex.ru
 Ufa State Aviation Technical University (UGATU), Ufa, Russia

Information Security of Cloud Computing System

The development model of threats and private information security policy cloud computing systems are discussed. The threat model on the based of fuzzy cognitive map is developed, which allows modeling the propagation of threats to the information system, built using cloud computing technologies, through the exploited vulnerability of the components of its infrastructure. The formation of the private security policy based on the use of the model RBAC.

Keywords: cloud computing, community cloud, the cloud computing system, cloud provider, cloud service consumer, the threat model, fuzzy cognitive map, private security policy, a role hierarchy

References

1. **Barskij A. B., Zhelenkov B. V.** Sredstva optimizacii informacionnogo vzaimodejstviya resursnyh processorov dlja minimizacii vremeni "oblachnyh" vychislenij (Means of optimization of informational interaction of resource processors to minimize the time of "cloud" computing), *Informacionnye tehnologii*, 2016, no. 1, vol. 22, pp. 14–21.
2. **NIST.** Oficial'nyj sajt (The official NIST website), URL: <http://www.nist.gov/itl/cloud/> (data of access: 28.03.2016).
3. **Авраменко А.** Oblachnye vychislenija. Vzlgjad iz IBM (Cloud computing. A view from IBM), *Jet Info*, 2010, no. 10, pp. 63–75.
4. **Zashhita informacii. Trebovanija po zashhite informacii, obrabatyvaemoj s ispol'zovaniem tehnologij Oblachnyh vychislenij** ГОСТ РХХХХХ-20XX (proekt, pervaja redakcija) (The protection of information. Requirements for protection of information processed using technologies of Cloud computing), URL: <http://docs.cntd.ru/document/1200102839> (data obrashhenija 28.03.2016).
5. **Glushenko S. A., Dolzhenko A. I.** Sistema podderzhki prinjatija reshenij nechetkogo modelirovanija riskov informacionnoj bezopasnosti organizacii (Support system decision making fuzzy modeling information security risk organization), *Informacionnye tehnologii*, 2015, vol. 21, no. 1, pp. 68–74.
6. **Mezhdunarodnaja baza dannyh uязvimostej** Oficial'nyj sajt (The official website of the national vulnerabilities database), URL: <https://nvd.nist.gov/> (data of access: 28.03.2016).
7. **Guzairov M. B., Mashkina I. V., Stepanova E. S.** Postroenie modeli ugroz s pomoshh'ju nechetkih kogitivnyh kart na osnove setevoj politiki bezopasnosti (The building threat models using fuzzy cognitive maps based on network security policy), *Bezopasnost' informacionnyh tehnologij*, 2011, no. 2, pp. 37–49.

8. **Shan'gin V. F.** *Zashhita informacii v komp'juternyh sistemah i setjah* (Information security in computer systems and networks), Moscow: ДМК Пресс, 2012, 592 p.

9. **Demurchev N. G., Ishhenko S. O.** Problemy obespechenija informacionnoj bezopasnosti pri perehode na oblachnye vychislenija (Problems of information security in the transition to cloud computing), *Materijaly XI Mezhdunarodnoj nauchno-prakticheskoj konferencii "Informacionnaja bezopasnost'"*, Ch. 1. Таганрог: Изд-во ТТИ ЮФУ, 2010, pp. 147–151.

10. **ГОСТ Р 50922—2006.** *Zashhita informacii. Osnovnye terminy i opredelenija* (The protection of information. Key terms and definitions). М.: Стандартинформ, 2007, 12 p.

11. **Varlataja S. K., Shahinova M. V.** Analiz metodov opisaniya politiki bezopasnosti pri razrabotke informacionno-bezopasnyh tehnologij. (The analysis of methods of the description of the security policy by working out of information-safe technologies), *Doklady TUSUR, Ch. 1 "Audit bezopasnosti"*, ijun' 2010, no. 1 (21), pp. 10–13.

12. **РС БР ИББС-2.0—2007.** *Rekomendacii v oblasti standartizacii Banka Rossii.* Obespechenie informacionnoj bezopasnosti organizacij bankovskoj sistemy Rossijskoj Federacii. Metodicheskie rekomendacii po dokumentacii v oblasti obespechenija informacionnoj bezopasnosti v sootvetstvii s trebovanijami SТО БР ИББС-1.0. (Recommendations standardization of the Bank of Russia. Ensuring the information security of organizations of Bank system of the Russian Federation. Guidelines for documentation in the field of information security in accordance with the requirements of SТО БР ИББС-1.0.), Moscow, Standartinform, 2007. 15 p.

13. **Metodologija organizacii tehničeskoj podderzhki** (Methodology technical support). URL: <http://msbro.ru/index.php/archives/2717> (data of access: 28.03.2016).

Вопросы корректности и универсальности подхода к моделированию надежностных параметров и характеристик угроз атак

Исследованы вопросы корректности и универсальности предложенного подхода к моделированию угроз атак, позволяющего определять надежностные параметры и характеристики угрозы атаки. Подход основан на построении марковской модели угрозы атаки — модели с дискретными состояниями и непрерывным временем, с последующим ее преобразованием в модель вероятностного разрежения входных потоков случайных событий. Исследование, проведенное на модели вероятностного разрежения входных потоков случайных событий, позволило сделать вывод о том, что марковская модель с дискретными состояниями и непрерывным временем без потерь должна быть счетной (не конечной), и о том, что применение подобной модели корректно для решения задачи моделирования угроз атак. Предложен подход к формированию допущений для преобразования счетной марковской модели в конечную с использованием закона Пуассона. Исследованы вопросы универсальности предложенного подхода к моделированию угроз атак, достигаемой за счет реализации предложенного подхода к объединению состояний в модели вероятностного разрежения входных потоков случайных событий.

Ключевые слова: угроза атаки, угроза уязвимости, резервирование, нивелирование, математическое моделирование, параметр и характеристика безопасности, надежность информационной безопасности, количественная мера актуальности угрозы атаки, марковская модель, разрежение входных потоков случайных событий

Введение

В работах [1, 2] изложен метод моделирования угрозы атаки марковской моделью с дискретными состояниями и непрерывным временем без потерь, основанный на введенной в [3] интерпретации угрозы атаки¹ схемой параллельного резервирования создающих ее угроз уязвимостей². В результате моделирования могут быть рассчитаны надежностные параметры и характеристики угрозы атаки, которые могут использоваться для количественной оценки уровня ее актуальности при принятии решения о необходимости в отношении нее реализации защиты. Искомые параметры и характеристики угрозы атаки названы нами "надежностными", поскольку при этом моделируется не атака, как процесс последовательного деструктивного воздействия нарушителем на систему [4—8], а именно угроза атаки, как процесс возникновения и устранения в системе отказов информационной безопасности — реальных угроз атак, создаваемых возникающими в системе реальными угрозами уязвимостей. Как следует из ГОСТ 27.002—89 "Надежность в технике. Основные понятия. Термины и определения", надежность — это свойство объекта сохранять во времени в установленных пределах значения всех параметров, характеризую-

щих способность выполнять требуемые функции в заданных режимах и условиях применения, технического обслуживания, хранения и транспортирования. Исходя из данного определения и проводя моделирование в рамках предложенной интерпретации угрозы атаки, можем говорить об определении надежностных параметров и характеристик безопасности информационной системы, а в общем случае — о свойстве надежности информационной безопасности, под которой понимаем свойство информационной системы сохранять во времени в установленных пределах значения всех характеристик безопасности, определяющих способность выполнять требуемые функции в безопасном режиме. А вот построение и включение в модель угрозы атаки математической модели нарушителя [2] уже позволяет получать необходимые при проектировании системы защиты информации эксплуатационные параметры и характеристики безопасности, в данном случае — реализации возникающей в информационной системе реальной угрозы атаки потенциальным нарушителем, поскольку в данной модели уже учитывается сложность реализации угрозы атаки и готовность потенциального нарушителя к реализации реальной угрозы атаки в целях получения несанкционированного доступа к информации, представляющей для него определенную ценность.

В [1] дано обоснование корректности использования марковских процессов (процессов без последствия) для моделирования угрозы атаки при предложенной ее интерпретации, в [2] рассмотрена возможность использования с этой целью конечных (конечное число состояний) марковских моделей с дискретными состояниями и непрерывным временем (конечные непрерывные цепи Маркова) без потерь, однако не исследованы вопросы кор-

¹ Под угрозой атаки понимаем угрозу безопасности, создаваемую совокупностью угроз уязвимостей, реализация которых потенциальным нарушителем необходима для осуществления этой атаки. Под реальной угрозой атаки понимаем условие, при котором все уязвимости, необходимые для реализации этой атаки, одновременно присутствуют в информационной системе — все угрозы уязвимостей, создающие угрозу этой атаки, реальны.

² Под угрозой уязвимостей понимаем угрозу возникновения (присутствия) свойства информационной системы, обуславливающего возможность реализации на нее атаки потенциальным нарушителем.

ректности применения для решаемых задач моделирования данного математического аппарата. Проведем данные исследования в этой работе.

Подход к моделированию угрозы атаки и проблема корректности используемой марковской модели

Прежде всего, уточним, что в [9] была введена классификация угроз уязвимостей, которые в общем случае могут быть отнесены к технологическим — безусловным или условным, это уязвимости, связанные с некорректностью реализации защиты, и к уязвимостям реализации, в первую очередь, это ошибки программирования³. Угрозы технологических уязвимостей, в том числе условных, возникающих при выявлении уязвимости реализации, должны нивелироваться системами защиты информации, что можно позиционировать как постановку задачи защиты информации от несанкционированного доступа в общем виде [9]. А вот угрозы уязвимостей реализации⁴, которые с какими-то интенсивностями возникают и устраняются в системе, должны рассматриваться при моделировании угрозы атаки.

Очень важным, как увидим далее, является то, что, как показано в [1], следует говорить не об угрозах уязвимостей реализации, а об угрозах уязвимостей реализации соответствующего типа, предполагая при этом возможность одновременного (не одномоментного) появления в системе нескольких угроз уязвимостей одного типа. При этом под угрозами уязвимостей реализации одного типа понимаем угрозы реализации, создающие угрозу атаки на одну и ту же условную технологическую уязвимость.

Исходными данными при моделировании выступают интенсивность возникновения уязвимости λ и интенсивность устранения уязвимости μ , которые могут быть объективно, без использования каких-либо экспертных оценок, определены из соответствующей статистики [1]. Задачей же моделирования является определение параметров и характеристики безопасности угрозы атаки, создаваемой соответствующей совокупностью (набором) угроз

³ Под угрозой технологических уязвимостей понимаем технологические недостатки построения информационной системы в части обеспечения безопасности информации, включая отсутствие требуемых функций защиты информации, либо некорректность их реализации, не позволяющие в полном объеме реализовать защиту от несанкционированного доступа к информации. Под угрозой безусловной технологической уязвимости понимаем угрозу технологической уязвимости, присутствующую в информационной системе всегда — без возникновения каких-либо дополнительных условий. Данная угроза всегда реальна. Под угрозой условной технологической уязвимости понимаем угрозу технологической уязвимости, которая создается (становится реальной) в информационной системе при возникновении неких дополнительных условий, без которых соответствующая штатная возможность системы не несет в себе угрозы безопасности информации.

⁴ Под угрозой уязвимостей реализации понимаем ошибки реализации (программирования) используемых в информационной системе средств или некоторые штатные возможности системных средств и/или приложений, создающие условия возникновения в системе реальной угрозы условной технологической уязвимости.

уязвимостей реализации, таких как интенсивность возникновения и устранения в системе реальной угрозы атаки (λ_a и μ_a), вероятности готовности к безопасной эксплуатации информационной системы в отношении угрозы атаки (P_{0a}), среднего времени наработки на отказ безопасности информационной системы (восстанавливаемая система) в отношении угрозы атаки ($T_{0y.a}$), среднего времени восстановления безопасности информационной системы ($T_{B.y.a}$) в отношении угрозы атаки, которые могут рассматриваться в качестве количественных оценок (меры) актуальности угрозы атаки при последующем проектировании системы защиты.

Подход к моделированию угрозы атаки [2] состоит в построении конечной марковской модели с дискретными состояниями и непрерывным временем с последующим приведением построенной марковской модели угрозы атаки к модели вероятностного разрежения входных потоков случайных событий и расчетом на ней требуемых параметров и характеристик угрозы атаки.

Проиллюстрируем данный подход к моделированию на примере. Пусть угроза атаки создается двумя угрозами (типами угроз) уязвимостей реализации (исключили из рассмотрения соответствующие угрозы технологических уязвимостей), с соответствующими их параметрами — интенсивностями выявления и устранения уязвимостей (аналогичным образом можно построить модель для угрозы атаки любой сложности). Состояния системы обозначим через S_{ij} , где i — число выявленных уязвимостей первого типа, j — число уязвимостей второго типа. Размеченный граф системы состояний случайного процесса (марковского процесса) приведен на рис. 1, а.

Построение модели вероятностного разрежения входных потоков основано на следующих соображениях. Входной поток событий, поступающий в систему (на вход марковской модели), вероятностно разрежается — распределяется между состояниями системы в том смысле, что событие может наступить в случайный момент времени, когда система находится в одном из возможных своих состояний (переходы между состояниями в марковской модели осуществляются мгновенно).

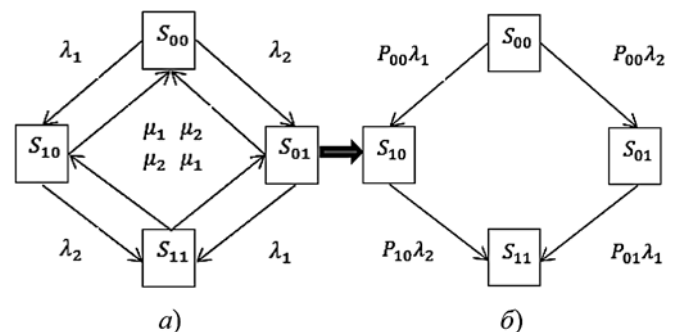


Рис. 1. Иллюстрация преобразования размеченного графа системы состояний случайного процесса марковской модели: а — марковская модель; б — модель вероятностного разрежения входных потоков

Важнейшее свойство входного потока случайных событий, которое требуется учитывать при моделировании угрозы атаки, — это его стационарность, поскольку очевидно, что уязвимости в системе возникают в случайные моменты времени, какая-либо регулярность данного потока отсутствует. С учетом выполнения свойства ординарности и отсутствия последствия [1], примем при моделировании, что входной поток случайных событий простейший (стационарный пуассоновский), что далее будем учитывать при анализе корректности моделей.

Получим из модели, приведенной на рис. 1, а, модель вероятностного разрежения входных потоков (рис. 1, б). При этом будем исходить из того, что вероятностное разрежение простейшего потока событий, при котором любое событие случайным образом с некоторой вероятностью p исключается из потока независимо от того, исключены другие события или нет, приводит к образованию простейшего потока с интенсивностью $\lambda' = p\lambda$, где λ — интенсивность исходного потока. Поток исключенных событий — тоже простейший с интенсивностью $\lambda'' = (1 - p)\lambda$ [10].

Замечание. В данном случае мы рассматриваем не исключение событий из потока, что реализуется в системах с потерями [10], а их распределение между состояниями системы — вероятностное разрежение входных потоков между возможными состояниями системы S_{ij} . В данном случае моделируется система без потерь, так как каждая возникающая в системе уязвимость реализации должна непременно устраняться.

При построении данной модели будем исходить из того, что вероятность нахождения системы в каком-либо состоянии в исходной марковской модели с дискретными состояниями и непрерывным временем (см. рис. 1, а), интерпретируется как доля времени нахождения системы в этом состоянии. Естественно, можно утверждать, что входной поток — поток, поступающий на вход модели с интенсивностью λ , распределяется между состояниями системы S_{ij} пропорционально P_{ij} — вероятностям нахождения системы в состояниях S_{ij} , определяемым для исходной марковской модели. С учетом этого получаем модель, приведенную на рис. 1, б.

Принципиальное отличие данной модели от марковской модели состоит в том, что переходы между состояниями на этой модели "взвешиваются" (размечаются) не интенсивностями возникновения событий в системе, а интенсивностями соответствующих вероятностно разреженных потоков — интенсивностями переходов между состояниями.

Для обоснования корректности данного преобразования достаточно построить модель вероятностного разрежения потоков (всех потоков, не только входных) в системе (рис. 2).

Как видим, обозначив через P_{ij} вероятность нахождения системы в состоянии S_{ij} , можем для обеих

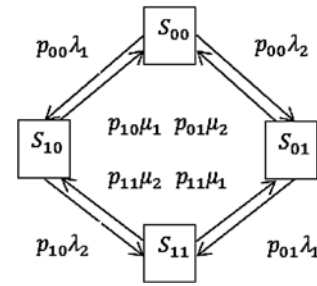


Рис. 2. Модель полного вероятностного разрежения потоков в системе

моделей, представленных на рис. 1, а и на рис. 2, записать одну и ту же систему линейных уравнений, что подтверждает корректность модели вероятностного разрежения входных потоков случайных событий:

$$\begin{cases} \mu_1 P_{10} + \mu_2 P_{01} = (\lambda_1 + \lambda_2) P_{00}; \\ \lambda_1 P_{00} + \mu_2 P_{11} = (\lambda_2 + \mu_1) P_{10}; \\ \lambda_2 P_{00} + \mu_1 P_{11} = (\lambda_1 + \mu_2) P_{01}; \\ \lambda_2 P_{10} + \lambda_1 P_{01} = (\mu_1 + \mu_2) P_{11}, \end{cases}$$

решая которую, с учетом полной группы событий, т.е. используя условие

$$P_{00} + P_{01} + P_{10} + P_{11} = 1,$$

находим искомые предельные (или финальные) вероятности состояний.

С использованием модели вероятностного разрежения входных потоков параметр безопасности угрозы атаки λ_a может быть рассчитан по следующей формуле:

$$\lambda_a = \sum_{S_i \in S_{(R-1)}} P_{S_i} \lambda_{S_i} S_R,$$

где $S_{(R-1)}$ — множество состояний системы, характеризующих отсутствием в системе реальной угрозы атаки (не все создающие угрозу атаки уязвимости выявлены и устранены), в каждом из которых система находится с вероятностью $P_{S_{(R-1)}}$, S_R — состояние возникновения в системе реальной угрозы атаки (все создающие угрозу атаки уязвимости выявлены, но не устранены). В состоянии S_R из состояний $S_{(R-1)}$ в системе осуществляется переход с интенсивностью $\lambda_{S_{(R-1)}, S_R}$.

Например, для модели, представленной на рис. 1, б, λ_a определяется следующим образом:

$$\lambda_a = P_{10} \lambda_2 + P_{01} \lambda_1.$$

Очевидно, что в стационарном (установившемся) режиме функционирования системы за долю времени нахождения системы в состоянии, характеризующем возникновением реальной угрозы атаки, определяемой, как $1 - P_{0a}$, где P_{0a} — это вероятность готовности системы к безопасной эксплуатации в отношении угрозы атаки, из состояния, характеризующего реальную угрозу атаки, исходит поступающий

в него поток событий интенсивностью λ_a (система без потерь, все выявляемые уязвимости устраняются). Сказанное позволяет рассчитывать параметр безопасности μ_a угрозы атаки следующим образом:

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

Для рассматриваемого примера имеем:

$$\mu_a = \frac{P_{10}\lambda_2 + P_{01}\lambda_1}{P_{11}}.$$

Соответствующим образом можем определить и характеристики безопасности угрозы атаки.

Вероятность готовности информационной системы к безопасной эксплуатации в отношении угрозы атаки:

$$P_{0a} = P_{00} + P_{10} + P_{01} = \frac{\mu_1\mu_2 + \lambda_1\mu_2 + \lambda_2\mu_1}{(\lambda_1 + \mu_1)(\lambda_2 + \mu_2)},$$

среднее время наработки на отказ безопасности информационной системы (восстанавливаемая система) в отношении угрозы атаки $T_{0y.a}$, среднее время восстановления безопасности информационной системы $T_{в.у.а}$ в отношении угрозы атаки:

$$T_{в.у.а} = \frac{1}{\mu_a}, T_{0y.a} = \frac{1}{\lambda_a} - T_{в.у.а}. \quad (1)$$

Замечание. Следуя модели, представленной на рис. 1, а, отношением $1/\lambda_a$ определяется характеристика среднего времени наработки системы между отказами безопасности, равного $T_{0y.a} + T_{в.у.а}$.

Отметим, что подобным образом может быть построена математическая модель угрозы атаки, создаваемой любым числом угроз (типов угроз) уязвимостей реализации, например, для случая, когда угроза атаки создается тремя угрозами уязвимостей реализации, модель представлена на рис. 3.

На графе представлены следующие возможные состояния системы: S_0 — исходное состояние системы, S_i — в системе возникла и не устранена одна из уязвимостей реализации (на тип выявленной уязвимости указывает индекс), S_{ij} — в системе возникли и не устранены одновременно две уязвимос-

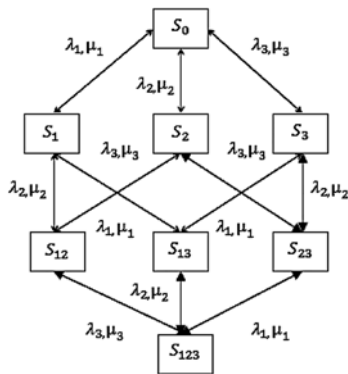


Рис. 3. Размеченный граф системы состояний случайного процесса для угрозы атаки, создаваемой тремя угрозами уязвимостей реализации

ти, S_{ij} — в системе возникли и не устранены одновременно все три уязвимости.

Теперь рассмотрим проблему использования для моделирования угрозы атаки конечной (с конечным числом состояний) марковской модели с дискретными состояниями и непрерывным временем. Именно такие модели приведены на рис. 1, а и рис. 3. Исследовать данную проблему нам позволит модель вероятностного разрежения входных потоков.

Замечание. Если число возможных состояний конечно или счетно (всем возможным состояниям могут быть присвоены порядковые номера), то случайный процесс называется процессом с дискретными состояниями [11].

Модель с конечным числом состояний нам необходима, так как требуется анализировать интенсивности переходов между определенными состояниями, набор которых в модели должен быть конечным.

Определим, например, интенсивность потока событий, циркулирующего в модели вероятностного разрежения входных потоков, создаваемого возникновением и устранением первого типа уязвимостей (см. рис. 1). На вход модели для этой угрозы уязвимостей поступает соответствующий простейший поток событий с интенсивностью λ_1 . Этот поток событий переводит систему из состояния S_{00} , в котором система находится с вероятностью P_{00} , и из состояния S_{10} , в котором система находится с вероятностью P_{10} , т.е. вероятно разрежается между двумя состояниями системы S_{00} и S_{01} . Как следствие, интенсивность рассматриваемого потока событий, циркулирующего в модели, обозначим ее $\lambda_{п1}$ определяется следующим образом:

$$\lambda_{п1} = (P_{00} + P_{01})\lambda_1 < \lambda_1,$$

Вызвано данное противоречие ($\lambda_{п1} < \lambda_1$) тем, что не из всех состояний марковской модели есть переходы, создаваемые потоком событий с интенсивностью λ_1 , поступающим на вход марковской модели — переходы отсутствуют для состояний S_{10} и S_{11} , входной поток разрежается не между всеми состояниями, т.е. в системе присутствуют интервалы времени, в течение которых события в систему не поступают, что в общем случае уже не позволяет говорить о корректности использования для такой модели простейшего (стационарного пуассоновского) входного потока случайных событий. Естественно, используя подобную модель, имеем погрешность моделирования, которая применительно к рассматриваемому примеру тем больше, чем больше значение суммы $P_{10} + P_{11}$ (в общем случае — это сумма значения вероятностей состояний, из которых не выходит анализируемый поток событий). В нашем случае — при моделировании угроз атак — погрешность моделирования, вызванная рассматриваемой проблемой, может быть достаточно большой, так как для определенных уязвимостей (типов уязвимостей) интенсивность их возникновения, как следует из существующей статистики, может быть достаточно велика.

Построение корректных марковских моделей угрозы атаки

Сформулируем и докажем несколько важных утверждений, касающихся рассмотренной проблемы моделирования угрозы атаки.

Аксиома. Марковская модель с дискретными состояниями и непрерывным временем при стационарных потоках входных случайных событий корректна при условии корректного вероятностного разрежения ею всех входных потоков случайных событий.

Утверждение 1. Модель угрозы атаки, как системы без потерь с дискретными состояниями и непрерывным временем, корректна в общем случае (без соответствующих допущений) только при условии, что из каждого состояния на графе системы состояний случайного процесса исходят все I входных потоков событий с интенсивностями $\lambda_i, i = 1, \dots, I$.

Доказательство. Только при выполнении этого условия в общем случае (без каких-либо допущений) для всех I входных потоков событий с интенсивностями $\lambda_i, i = 1, \dots, I$, будет выполняться условие $\lambda_{\Pi} = \lambda_i$, что подтверждает корректность вероятностного разрежения входных потоков случайных событий для этой модели и, в том числе, обуславливает корректность определения на такой модели надежных параметров и характеристик безопасности угрозы атаки.

Утверждение 2. В общем случае (без каких-либо обоснованных допущений) для моделирования угрозы атаки должны использоваться счетные (с бесконечным числом состояний) марковские модели с дискретными состояниями и непрерывным временем.

Доказательство. Условие того, что из каждого состояния на графе системы состояний случайного процесса исходят все I входных потоков случайных событий с интенсивностями $\lambda_i, i = 1, \dots, I$, выполнимо только при бесконечном числе состояний на графе.

Выводы.

1. Число возможных состояний в марковской модели с дискретными состояниями и непрерывным временем должно быть счетным, поскольку при конечном числе состояний в модели некорректно разрежаются входные потоки.

2. Марковская модель с дискретными состояниями и непрерывным временем без потерь может применяться для математического моделирования объектов, характеризующихся возможностью одновременного (не одномоментного) возникновения в системе двух и более событий одного типа.

3. Марковская модель с дискретными состояниями и непрерывным временем может применяться для математического моделирования угроз атак, поскольку в системе возможно одновременное возникновение нескольких реальных угроз уязвимостей реализации одного типа.

Пример графа переходов корректной модели угрозы атаки для случая возникновения в системе двух типов угроз уязвимостей реализации для слу-

чая, ранее проиллюстрированного на рис. 1, а, приведен на рис. 4.

Замечание. Аналогичным образом может быть построена модель для любого числа типов угроз уязвимостей.

На графе представлены следующие обозначения состояний системы: S_{ij} , где i — число возникших и не устраненных уязвимостей реализации первого типа, j — число возникших и не устраненных уязвимостей реализации второго типа.

При существовании в системе только одного потока событий представленный на рис. 3 граф выродается в схему "гибели и размножения" с бесконечной очередью (все выявляемые уязвимости должны устраняться) и с n обслуживающими приборами — каналами обслуживания; предполагаем, что n выявленных в системе уязвимостей (естественно, в данной системе одного типа) могут устраняться одновременно, остальные же будут ожидать в очереди на обслуживание (рис. 5) [11].

Замечание. Все сказанное в полной мере относится и к невозстанавливаемым системам, которые могут использоваться для моделирования угрозы атаки, направленной на нарушение конфиденциальности обрабатываемой информации, поскольку конфиденциальность похищенной информации не восстанавливается. В данном случае при моделировании используется поглощающее (из которого нет переходов) состояние S_a (рис. 6).

Однако, как ранее отмечали, нам необходима именно конечная (с конечным числом состояний) марковская модель, поскольку, в отличие от теории массового обслуживания, в данном случае моделируются не обслуживающие приборы, а интенсивности переходов между состояниями, для чего используется модель вероятностного разрежения входных потоков случайных событий. Понятно, что любое математическое моделирование предпо-

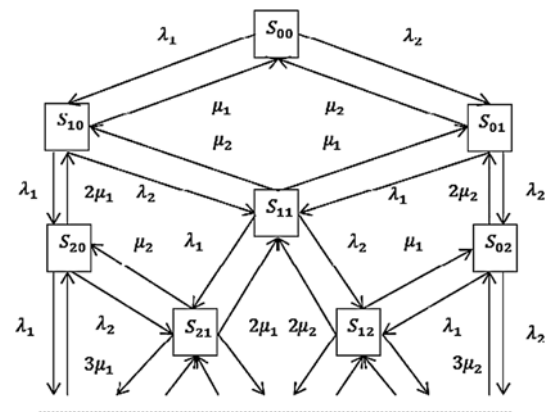


Рис. 4. Размеченный граф системы состояний случайного процесса корректной модели

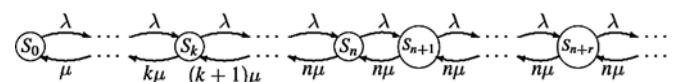


Рис. 5. Схема гибели и размножения

лагает некий набор обоснованно сделанных допущений. Рассмотрим, каким образом можно сделать обоснованные допущения для перехода от корректной счетной к конечной модели угрозы атаки.

В целях обоснованного введения соответствующих допущений для моделирования воспользуемся законом Пуассона (поскольку мы используем простейший — пуассоновский — входной поток при моделировании) [11]. Вероятность того, что на некотором произвольно взятом на временной оси интервале времени t наступит ровно m событий $P_m(t)$, поток которых характеризуется интенсивностью возникновения событий λ , определяется выражением

$$P_m(t) = \frac{(\lambda t)^m}{m!} e^{-\lambda t}.$$

Нас интересует вероятность возникновения в системе одновременно (не одномоментно) нескольких событий — одновременное выявление нескольких уязвимостей реализации одного типа на интервале времени устранения уязвимостей реализации этого типа, средней продолжительности $t = 1/\mu$. Используя параметр нагрузки (или коэффициента нагрузки) $\rho = \lambda/\mu$, можем определить требуемую нам вероятность

$$P_m(\rho) = \frac{\rho^m}{m!} e^{-\rho}.$$

Естественно, что на точность моделирования сказывается то, какие состояния S_{ij} в модели (см. рис. 4), мы можем исключить, сделав тем самым исходную счетную модель конечной.

Результаты расчетов $P_m(\rho)$ — вероятности того, что при значении нагрузки $\rho = \lambda/\mu$ за произвольно взятый на временной оси интервал времени $t = 1/\mu$ (продолжительность устранения возникшей реальной угрозы уязвимостей реализации) в информационной системе одновременно будет выявлено и не устранено m уязвимостей реализации этого типа, приведены в таблице.

Как видим из таблицы, необходимость учета вероятностей $P_{m > 1}(\rho)$ для угроз уязвимостей реализации при моделировании определяется значением характеристики ρ и требованиями к точности модели. При этом подход к построению корректной модели угрозы безопасности состоит в следующем.

Для каждого типа угрозы уязвимостей, с учетом заданных требований к точности моделирования, посредством расчета значений вероятностей $P_m(\rho)$ определяется число \max_i учитываемых при моделировании одновременно возникающих в системе уязвимостей одного типа (событий). Все состояния

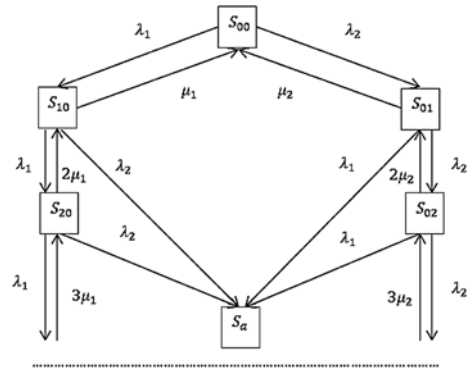


Рис. 6. Размеченный граф системы состояний случайного процесса корректной модели невозстанавливаемой системы

$S_i > \max_{ij}$ и дуги между ними исключаются из размеченного графа системы состояний случайного процесса счетной модели, в результате чего получаем искомую корректную (в части вводимых допущений) конечную модель угрозы атаки, корректность которой обосновывается тем, что вероятность события — появление одновременно $\max(i + 1)$ уязвимостей одного типа, не сказывается на результатах моделирования с требуемой точностью.

Приведем пример построения корректной конечной модели угрозы атаки. Построим конечную модель угрозы атаки, создаваемой двумя типами угроз уязвимостей реализации (см. рис. 1, а) в предположении, что применительно к первой угрозе уязвимостей (первого типа) необходимо учитывать одновременное возникновение в системе двух уязвимостей, для второй же угрозы уязвимостей (второго типа) одновременным появлением в системе нескольких уязвимостей можно пренебречь. В результате вводимых по рассмотренному правилу допущений получим корректную модель угрозы атаки, размеченный граф системы состояний случайного процесса для которой представлен на рис. 7, а, модель вероятностного разрежения входных потоков для расчета надежностных параметров угрозы атаки — на рис. 7, б.

Оценим корректность данной модели, для чего опять же определим интенсивность первого потока событий, циркулирующего в модели, аналогично тому, как это было сделано для марковской модели, приведенной на рис. 1, а. Этот поток событий переводит систему из состояний $S_{00}, S_{10}, S_{01}, S_{11}$ в другие состояния. Как следствие, интенсивность рассматриваемого потока событий в системе $\lambda_{п1}$ определяется уже следующим образом:

$$\lambda_{п1} = (P_{00} + P_{10} + P_{01} + P_{11})\lambda_1.$$

Результаты расчетов

ρ	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
$P_2(\rho)$	0,005	0,016	0,033	0,054	0,076	0,099	0,122	0,144	0,164
$P_3(\rho)$	0	0,001	0,003	0,007	0,013	0,020	0,028	0,038	0,049

В данном случае выполняется условие: $\lambda_{п1} \approx \lambda_1$, за счет введенных допущений при моделировании — из модели исключены состояния, не влияющие на результаты моделирования с требуемой точностью (переходами в которые можно пренебречь).

Универсальность подхода к моделированию угроз атак

Универсальность рассматриваемого подхода к моделированию угроз атак обеспечивается возможностью объединения состояний в модели вероятностного разрежения входных потоков случайных событий, что позволяет формировать состояния, интересующие при проведении соответствующих оценок, и применительно к ним рассчитывать соответствующие параметры и характеристики безопасности. Проиллюстрируем сказанное примером.

Используя построенную модель вероятностного разрежения входных потоков (см. рис. 7, б), можем оценить параметры и характеристики безопасности для состояний S_{11} — в системе присутствует одна реальная угроза атаки, и S_{21} — в системе присутствуют одновременно две реальные угрозы атаки:

$$\lambda_{a1} = P_{10}\lambda_2 + P_{01}\lambda_1.$$

$$\lambda_{a2} = P_{20}\lambda_2 + P_{11}\lambda_1.$$

Построим модель для оценки события — в системе присутствует хотя бы одна реальная угроза ата-

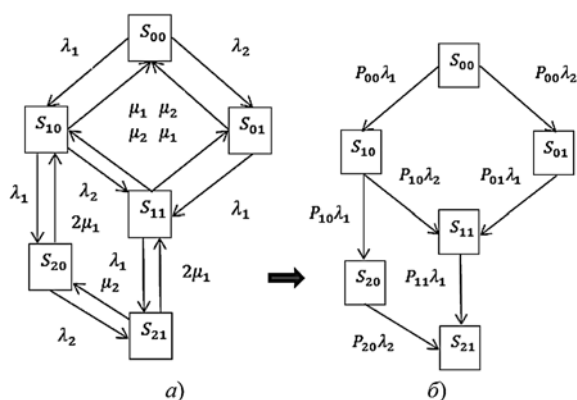


Рис. 7. Иллюстрация преобразования размеченного графа системы состояний случайного процесса корректной марковской модели: а — корректная марковская модель угрозы атаки; б — модель вероятностного разрежения входных потоков

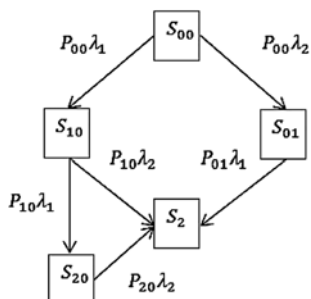


Рис. 8. Модель вероятностного разрежения входных потоков с объединением состояний

ки, т.е. система находится в состоянии отказа безопасности, обозначим это состояние через S_2 . Построение соответствующей модели состоит в объединении состояний S_{11} и S_{21} (при этом $P_2 = P_{11} + P_{21}$), в результате чего получаем модель, приведенную на рис. 8.

Обозначим параметры отказов и восстановления безопасности информационной системы в отношении угрозы атаки, как λ_0 и μ_B . Для модели, приведенной на рис. 8, они могут быть рассчитаны следующим образом:

$$\lambda_0 = \lambda_2(P_{10} + P_{20}) + P_{01}\lambda_1;$$

$$\mu_B = \frac{(P_{10} + P_{20})\lambda_2 + P_{01}\lambda_1}{P_2} = \frac{(P_{10} + P_{20})\lambda_2 + P_{01}\lambda_1}{P_{11} + P_{21}},$$

где параметр μ_B рассчитывается с учетом того, что безопасность системы, нарушаемая с интенсивностью λ_0 , восстанавливается за долю времени $P_2 = P_{11} + P_{21}$ (это доля времени нахождения системы в объединенном состоянии S_2).

Замечание. Несложно показать, что для λ_0 справедливо соотношение

$$\lambda_0 = \lambda_{a1} + \lambda_{a2} - P_{11}\lambda_1.$$

Соответствующим образом, как это было рассмотрено ранее (с использованием λ_0 и μ_B , по формулам (1) рассчитываются и требуемые характеристики безопасности угрозы атаки.

Сформулируем требование к корректности объединения состояний в модели вероятностного разрежения входных потоков случайных событий.

Утверждение 3. Корректное объединение состояний на графе системы состояний случайного процесса реализуется в том случае, если из объединяемых состояний под воздействием одного и того же потока входных случайных событий реализуются переходы в одно и то же, в том числе объединенное, состояние.

Доказательство. В противном случае невозможно построение корректной марковской модели, поскольку под воздействием одного и того же потока входных случайных событий без какого-либо его разрежения должен осуществляться переход из одного состояния сразу в несколько состояний на графе системы состояний случайного процесса.

Как видим, изложенный подход к моделированию универсален в том смысле, что при его использовании, за счет возможности объединения событий в модели вероятностного разрежения входных потоков, могут моделировать различные события применительно к угрозе атаки, причем, как на информационную систему, так и на систему защиты информации.

Заключение

Исследования в данной работе проводились в части рассмотрения подхода к моделированию надежных параметров и характеристик безопас-

ности угроз атак. Рассмотренный в работе подход к моделированию мы далее будем использовать и применительно к моделированию эксплуатационных параметров и характеристик безопасности, но при этом уже моделируется не угроза атаки, а реализация возникающей в информационной системе угрозы атаки потенциальным нарушителем, при котором учитывается заинтересованность и готовность потенциального нарушителя к реализации реальной угрозы атаки соответствующей сложности для получения им несанкционированного доступа к информации, характеризующей определенной ценностью для потенциального нарушителя. В данном случае уже будем строить модели с потерями входных случайных событий. Эти вопросы мы рассмотрим в следующей работе.

Список литературы

1. **Щеглов К. А., Щеглов А. Ю.** Интерпретация и моделирование угрозы атаки на информационную систему. Часть 1. Моделирование угрозы уязвимости и интерпретация угрозы атаки // Информационные технологии. 2015. Т. 21. № 12. С. 930–940.
2. **Щеглов К. А., Щеглов А. Ю.** Интерпретация и моделирование угрозы атаки на информационную систему. Часть 2.

Моделирование угрозы атаки // Информационные технологии. 2016. Т. 22. № 1. С. 54–64.

3. **Щеглов К. А., Щеглов А. Ю.** Математические модели эксплуатационной информационной безопасности // Вопросы защиты информации. 2014. Вып. 106. № 3. С. 52–65.
4. **Котенко Д. И., Котенко И. В., Саенко И. Б.** Методы и средства моделирования атак в больших компьютерных сетях: стоящие проблемы // Труды СПИИРАН. 2012. Вып. 3 (22). С. 5–30.
5. **Климов С. М., Сычев М. П., Астрахов А. В.** Противодействие компьютерным атакам. Технологические основы: Электронное учебное издание. М.: МГТУ им. Н. Э. Баумана, 2013. [Электронный ресурс]. URL: <http://wwwcdl.bmstu.ru/iu10/comp-atak-techno.htm>, свободный (10.02.2016).
6. **Zhang S., Song S.** Novel Attack Graph Posterior Inference Model Based on Bayesian Network // Journal of Information Security, 2011. N. 2. P. 8–27.
7. **Kichkaylo T., Ryutov T., Orosz M. D., Neches R.** Planning to Discover and Counteract Attacks // Informatica (Slovenia), 2010, N. 34 (2). P. 159–168.
8. **Gamal M. M., Hasan D., Hegazy A. F.** A Security Analysis Framework Powered by an Expert System // International Journal of Computer Science and Security. 2011. Vol. 4. P. 505–526.
9. **Щеглов К. А.** Постановка и подходы к решению задачи защиты информации от несанкционированного доступа в общем виде // Вестник компьютерных и информационных технологий. 2016. № 1. С. 32–44.
10. **Алиев Т. И.** Основы моделирования дискретных систем. СПб: Изд. СПбГУ ИТМО, 2009.
11. **Вентцель Е. С.** Исследование операций. Задачи, принципы, методология. М.: Высшая школа, 2007.

K. A. Shcheglov, Graduate Student, **A. Yu. Shcheglov**, Professor, e-mail.ru: info@npp-itb.spb.ru; St. Petersburg university of ITMO, Russia

Correctness and Versatility Problems of Modeling Attack Threats Reliability Parameters and Characteristics Approach

We did research correctness and universalism problems of suggested attack threat modeling approach which allows to determine reliability parameters and characteristics of attack threat security. This approach is based on building attack threat Markov model (a model with discrete states and continuous time) with its following transformation to input streams probability rarefaction model. Input streams probability rarefaction model based research allowed to do the conclusion that discrete states and continues time Markov model must be counting (not finite) and that such model use is correct for solving attack threats modeling problem. We suggest an approach to elaborate assumptions for counting Markov model transformation into finite model with the usage of Poisson law. We researched universalism problem of suggested attack threat modeling approach basing on states unification possibility in input streams rarefaction probability model.

Keywords: attack threat, vulnerability threat, reservation, leveling, mathematical modeling, security parameters and characteristics, informational security reliability, attack threat actuality quantitative measure, Markov model, input streams rarefaction

References

1. **Shcheglov K. A., Shcheglov A. Yu.** Interpretaciya i modelirovanie ugrozy ataki na informacionnyuyu sistemu. CHast' 1. Modelirovanie ugrozy uyazvimosti i interpretaciya ugrozy ataki, *Informacionnye tekhnologii*, 2015, vol. 21, no. 12, pp. 930–940 (in Russian).
2. **Shcheglov K. A., Shcheglov A. Yu.** Interpretaciya i modelirovanie ugrozy ataki na informacionnyuyu sistemu. CHast' 2. Modelirovanie ugrozy ataki, *Informacionnye tekhnologii*, 2016, vol. 22, no. 1, pp. 54–64 (in Russian).
3. **Shcheglov K. A., Shcheglov A. Yu.** Matematicheskie modeli ehkspluatacionnoj informacionnoj bezopasnosti, *Voprosy zashchity informacii*, 2014, vyp. 106, no. 3, pp. 52–65 (in Russian).
4. **Kotenko D. I., Kotenko I. V., Saenko I. B.** Metody i sredstva modelirovaniya atak v bol'shix komp'yuternyx setyah: sostoyanie problemy, *Trudy SPIIRAN*, 2012, vyp. 3 (22), pp. 5–30 (in Russian).
5. **Klimov S. M., Sychev M. P., Astrahov A. V.** *Protivodejstvie komp'yuternym atakam. Tekhnologicheskie osnovy*: EHlektronnoe uchebnoe izdanie, Moscow: MGTU imeni N. EH. Bauman, 2013.

[EHlektronnyj resurs]. URL: <http://wwwcdl.bmstu.ru/iu10/comp-atak-techno.htm>, svobodnyj (10.02.2016) (in Russian).

6. **Zhang S., Song S.** Novel Attack Graph Posterior Inference Model Based on Bayesian Network, *Journal of Information Security*, 2011, no. 2, pp. 8–27.
7. **Kichkaylo T., Ryutov T., Orosz M. D., Neches R.** Planning to Discover and Counteract Attacks, *Informatica* (Slovenia), 2010, no. 34 (2), pp. 159–168.
8. **Gamal M. M., Hasan D., Hegazy A. F.** A Security Analysis Framework Powered by an Expert System, *International Journal of Computer Science and Security*, 2011, vol. 4, pp. 505–526.
9. **Shcheglov K. A.** Postanovka i podhody k resheniyu zadachi zashchity informacii ot nesankcionirovannogo dostupa v obshchem vide, *Vestnik komp'yuternyh i informacionnyh tekhnologij*, 2016, no. 1, pp. 32–44 (in Russian).
10. **Aliev T. I.** *Osnovy modelirovaniya diskretnyh sistem*. SPb: Izd. SPbGU ITMO, 2009 (in Russian).
11. **Ventcel' E. S.** *Issledovanie operacij. Zadachi, principy, metodologiya*, Moscow, Vysshaya shkola, 2007 (in Russian).

УДК 004.5

Р. М. Алыгулиев, д-р техн. наук, зав. отделом, e-mail: r.aliguliyev@gmail.com,
И. Я. Алекперова, зав. сектором, e-mail: airada.09@gmail.com,
Институт информационных технологий НАН Азербайджана, г. Баку

Проблема BIG DATA в нефтегазовой отрасли: состояние и перспективы

Тема больших данных (BIG DATA) в нефтегазовой отрасли многогранна. Крупнейшие нефтегазовые компании мира уже давно имеют дело с большими объемами данных для принятия решения. Чтобы увеличить производство и устоять перед конкурентами, нефтегазовые компании внедряют эффективный анализ больших данных.

Выявлены проблемы, связанные с большими данными, дана классификация источников данных в нефтегазовой промышленности. Также исследованы подходы, используемые крупнейшими нефтегазовыми компаниями для решения проблем, связанных с хранением и обработкой больших данных.

Ключевые слова: нефтегазовые компании, большие данные, информационные технологии, анализ данных, аналитика, сложные аналитические запросы

Введение

В то время как быстро развиваются информационные технологии, возможность доступа, анализа и управления огромными объемами данных приобретает все большее значение для разведки и доставки нефти и газа. Проблемы, связанные с большими данными в нефтегазовой отрасли, сегодня как никогда актуальны. Процессы разведки, разработки и производства, связанные с добычей нефти и природного газа, генерируют большие данные. Компании соперничают друг с другом и используют самые новейшие информационные технологии для хранения и обработки больших данных.

В современном мире нефтегазовая промышленность является высокотехнологичной, капиталоемкой отраслью экономики, и в мировой нефтеперерабатывающей промышленности происходят заметные концептуальные, структурные сдвиги. В больших нефтегазовых компаниях, таких как Saudi Aramco (Саудовская Аравия), ExxonMobil (США), Royal Dutch Shell (Голландия — Великобритания), PetroChina (Китай), Chevron (США), Kuwait Petroleum Company (Кувейт), BP (Великобритания), Total (Франция), Petrobras (Бразилия), Qatar Petroleum (Катар), "Роснефть" (Россия), "Газпром" (Россия), Национальная нефтяная компания Ирана (Исламская Республика Иран) и т. д., среднесуточная добыча нефти составляет несколько миллионов баррелей. Эти компании хорошо осознают ценность анализа больших данных и используют их в новых проектах. Это не только коммерческие проекты. Основные из них сконцентрированы на нефтегазоносных пластах, другие — на бурении скважин.

Чтобы увеличить добычу нефти и газа, крупнейшие нефтегазовые компании все чаще прибегают к исследованиям, связанным с большими данными. Неудивительно, что в погоне за нефтью некоторые из них решили попытаться сократить свои невероятные расходы и риски, связанные с добычей природного углеводородного топлива. Было определено, что если большие данные помогут хоть одной крупной нефтедобывающей компании добыть 1 % дополнительной нефти, этого процента хватит, чтобы три года обеспечивать продуктами, полученными из этого объема, целый мир [1].

Нефтяная промышленность является отраслью экономики, которая занимается добычей, переработкой, транспортировкой, хранением и продажей нефти и сопутствующих нефтепродуктов. Современные компании, работающие в сфере нефтегазового бизнеса, являются либо вертикально интегрированными, либо специализированными. Интегрированные компании включают в свой состав всю производственную цепочку — от скважины до продажи бензина, добычу и переработку газа, а также нефтехимическую отрасль и производство альтернативных топлив. Специализированные компании работают в более узкой сфере — это геологоразведочные работы, добыча, переработка, транспортировка нефти и газа. Но эти компании тесно взаимодействуют с интегрированными компаниями. И интегрированные, и специализированные компании связывает одна цель — дополнительная прибыль и конкурентоустойчивость.

Погоня за дополнительной нефтью подталкивает компании всерьез заняться пересмотром производственных процессов, обращая основное внимание на повышение эффективности и внедрение инноваций,

так как сейчас одинаково важно снизить затраты и найти новые направления роста производительности. В этих условиях большое значение приобретают большие данные. В нефтегазовых компаниях используются огромные объемы различной информации. Речь идет не только об анализе больших данных геологоразведки, но также важен анализ всех данных, полученных при управлении технологическими процессами, начиная от перекачки нефти и газа и заканчивая механизмами контроля всего производственного процесса — от добычи до доставки потребителю.

Проблемы больших данных

В современном мире нефтегазовые компании работают в быстро меняющемся рынке. Новые методы извлечения сырья и появление альтернативных видов энергии создают новые проблемы. Разные виды энергии выходят на рынок и при этом создают избыточное предложение. Политические события в мире также создают дефицит информации о нефтегазовых компаниях. Понимание направления рыночного ценообразования и спроса имеет решающее значение. Например, когда местная поставка начнет обгонять спрос, нефтеперерабатывающим предприятиям, возможно, потребуется изучать другие рынки и политические отношения соседних стран.

Крупнейшие нефтегазовые компании стремятся повысить эффективность бизнеса и производительности, но при этом сталкиваются с целым рядом проблем, основные из них перечислены ниже:

- нефть трудно найти. Для обнаружения и описания нефтяных залежей требуются новейшая технология и дорогое оборудование;
- производство нефти стоит дорого. Для производства одного барреля нефти необходимо большое количество техники и трудовых ресурсов;
- добыча нефти представляет потенциальные проблемы для охраны окружающей среды и безопасности человека, которые необходимо безотлагательно решать.

Неопределенность и изменчивость цен на нефть и газ, изменение политики в области энергетики, экологические проблемы (например, от глобального потепления, цунами и борьбы с последствиями добычи сланцевой нефти), возникающие вследствие конкуренции со стороны новых источников энергии, а также постоянные операционные затраты на управление и т.д. создают большие проблемы, которые требуют срочного и эффективного решения [2].

Аналитики утверждают, что начало новой эры нефти и газа будет играть огромную роль в будущем энергетических компаний. Несмотря на развитие возобновляемых технологий мировой спрос на нефть по-прежнему будет расти до 2035 г., поскольку все больше людей используют автомобили и другие средства передвижения, которые работают на бензине. А так как добыча становится все более трудным делом, необходимы для удовлетворения будущих потребностей [3] инвестиции на \$20 трлн.

Многие компании сталкиваются с проблемами и неопределенностью, но продолжают делать упор на более эффективное управление активами и контроль.

Также остается необходимость сосредоточиться на эффективной и действенной разведке нефти и газа. Не только поддержание благоприятной экологической репутации, общественные и государственные отношения являются приоритетом для хорошего бизнеса, но и получение одобрения правительства для новой разведки и добычи, когда это необходимо. Типы данных, используемых аналитиками для этих целей, могут варьироваться в широких пределах. Одной из главных задач остается осмысление этих данных способами, которые окажут влияние на бизнес.

Источники больших данных (BIG Data)

Столкнувшись с разными проблемами, многие нефтегазовые компании видят решение проблем в исследовании больших данных, непрерывно растущих и хранящихся в различных базах данных компаний. Они могут предоставить такие данные в качестве важных источников информации, необходимых для оптимизации геологоразведочных работ, бурения, добычи и доставки. Нельзя не отметить новые источники данных, такие как данные из социальных медиа, мобильных устройств и личных архивов. Они могут обеспечить понимание настроения местного населения, недовольного работой промышленных объектов и трубопроводов или пострадавшего в авариях или при несчастных случаях. Хорошо управляемые нефтяные объекты имеют важное значение для поддержания положительной репутации в оценке финансовых активов.

В своем стремлении узнать, что лежит ниже поверхности земли и как это добыть, энергетические компании на протяжении многих лет инвестируют в сейсмические исследования программное обеспечение, средства визуализации и другие цифровые технологии. На сегодняшний день нефтяные и газовые компании анализируют данные из различных источников. Эти источники данных могут включать (рис. 1, см. четвертую сторону обложки):

- данные, полученные во время разведки нефти и газа, бурения, добычи, транспортировки и переработки нефти;
- традиционные корпоративные данные операционных систем;
- социальные медиа;
- информационные веб-проекты;
- данные из объявлений, в том числе о работе;
- демографические данные;
- история разведки нефти и газа;
- данные о доставке, хранении, а также о цене нефти и газа.

Основываясь на опыте ведущих нефтегазовых компаний, можно проследить несколько основных тенденций использования данных. Одна из них — разведка новых месторождений нефти и газа. Для решения этой важной задачи необходимо проводить сбор и анализ данных картирования океанского и морского дна и на его основе выявлять особенности и признаки, которые могут указывать на наличие месторождений. Однако огромные размеры файлов сканирования и резкий ежегодный рост совокупных объемов хранимых данных требуют наличия горизон-

тально масштабируемой платформы больших данных для быстрого добавления емкости.

Если говорить о сейсморазведке, информация при сканировании может дать отдельные файлы размером в несколько терабайтов, а это значит, что нефтегазовые компании должны каждый год увеличивать емкость на несколько петабайт, чтобы справиться с объемами поступающих данных. Кроме того, по мере совершенствования алгоритмов анализа и поиска новых решений особое значение приобретает ретроспективный анализ архивированных данных. В таких данных могут быть скрыты непредсказуемая информация, геологические тайны. Для эффективного анализа этих данных важно размещение всех архивных данных на более доступном уровне хранения. Некоторые нефтяные компании уже работают в этом направлении, развертывая файловые системы для хранения и извлечения огромных объемов данных, относящихся к разведке, добыче и переработке нефти и газа [3, 4].

Нефтегазовые компании все чаще осознают, что их огромные объемы данных хранятся в разных репозиториях. Это сильно усложняет аналитику и снижает производительность. Объединение этих ресурсов в единой системе станет стимулом для разработки нового поколения приложений для аналитики. Разведка новых месторождений нефти и газа и объединение данных формируют информационную основу для третьей важной задачи, которая заключается в использовании аналитики как стимула для других драйверов эффективности. Это, прежде всего, относится к использованию более сложных алгоритмов поиска при анализе данных сейсморазведки, но нефтегазовые компании хранят огромные объемы других полезных данных. К ним относятся:

- данные видеонаблюдения промышленных объектов, которые можно использовать для повышения безопасности или оптимизации процессов;
- данные работы оборудования, входящего в сложные цепочки взаимосвязанных процессов обработки, добычи и поставки, которые можно использовать для прогнозирования техобслуживания и снижения убытков в случае непредвиденного отключения или выхода из строя оборудования.

Правильное и полное использование этих данных становится приоритетной задачей для компаний. Скорость, с которой генерируются эти данные, быстро растет и приводит к более высоким темпам потребления аналитики, которую жаждут получить производители. Такие увеличение скорости передачи данных и рост источников, естественно, приводят к увеличению совокупных объемов данных. Аналитики хотят получать больше данных при более высоких скоростях, хранить их дольше и анализировать быстрее. Но при решении этих задач возникают проблемы, связанные с большими данными. Технологии BIG Data помогают нефтегазовым компаниям решать эти задачи. BIG Data — это подход, используемый для описания теории и практики применения передовых компьютерных технологий при анализе постоянно растущей информации [4]. Аналитика осуществляет открытие и связь значимых закономерностей в данных (см. рис. 1 на четвертой стороне обложки), как правило, с использованием сложных методов моде-

лирования. Существуют различные аналитические процессы, основные среди них — приобретение, очистка, визуализация, моделирование и оценка данных. Примерно 80 % времени анализа тратится на приобретение и очистку данных, поэтому аналитика всегда остается мишенью для улучшения процесса анализа.

Нефтяные и газовые компании уже давно ориентированы на повышение эффективности геологоразведочных работ за счет использования передовых аналитических средств, применяемых к разнообразным данным [5]. Сейсмическое отображение дает подсказки о том, где бурить. Эти данные играют большую роль при выборе систем наблюдений, изучении условий возбуждения сейсмических колебаний и способов их регистрации [6].

В сложных современных условиях нефтяным компаниям важно, чтобы стоимость новых исследований не росла, а показатели добычи улучшались, чтобы поддерживать рентабельность. Аналогичным образом повышение пропускной способности пласта без превышения затрат является основным фактором сохранения прибыльности. Такие данные традиционно анализируются в хранилищах данных, состоящих из реляционных баз данных. Сегодня некоторые из этих действий перемещаются в кластеры Hadoop — отчасти из-за относительно низкой цены кластеров Hadoop, отчасти из-за его "schema-less" файловой системы, которая является идеальной для прогнозного анализа рабочих нагрузок [7]. Необходимо хорошо понимать важность повышения эксплуатационной эффективности на буровых площадках, трубопроводах и нефтеперерабатывающих заводах. Это поможет лучше понять, когда обслуживание должно произойти раньше, и путем мониторинга изменять состояние ключевых компонентов, предоставляя еще большую эффективность и экономию.

По большей части информация идет от датчиков и других потоковых источников данных с внедрением системы управления BIG Data, включающей традиционные хранилища данных и новые резервуары данных (с участием Hadoop и/или NoSQL базы данных). Более широкие типы данных могут быть проанализированы, чтобы гарантировать то, что бизнес станет более гибким. Прогнозные аналитические решения, развернутые в системах управления BIG Data (включая Hadoop), становятся обычной практикой для повышения общей безопасности, надежности и снижения стоимости. Подобный анализ также указывает на потенциальные проблемы безопасности и выявляет экологические факторы риска до возникновения дорогостоящих аварий. Прогностическая аналитика широко используется при создании хранилищ данных, оптимизации маршрутизации транспортных средств и расходных материалов для разведки и добычи нефти и газа [7]. Такая маршрутизация имеет основополагающее значение для обеспечения минимально возможной стоимости доставки при сохранении рентабельности и решении основных вопросов технического обслуживания и безопасности.

Цены на газ и нефть могут быть весьма неустойчивыми, но их стоимость в течение срока реализации проекта может определить, сколько финансов требуют разведка, добыча и переработка. Здесь прогностическая аналитика может сыграть решающую

роль в понимании вероятного направления ценообразования и определения правильности разведки и выбора уровня производства. Такие данные могут также указывать на относительность вложения правильных инвестиций в недвижимость.

Возможности больших данных

Многие существующие бизнес-возможности расширяются, когда большое количество данных и их разнообразие становятся частью информационной архитектуры. Информационные технологии нефтегазовых компаний, как правило, используются для решения задач, которые обеспечивают следующее:

- улучшение разведки рентабельности инвестиций. Разведка нефтегазовых месторождений требует обширных сейсмического и экологического анализов, а также анализа стоимости производства, с тем чтобы определить, можно ли получить прибыль. Также надо учитывать меняющиеся рыночные условия;
- повышение эффективности производства. Эффективность производства обеспечивается за счет постоянного анализа буровых работ и своевременного технического обслуживания оборудования. Дистанционный мониторинг и анализ данных также имеют особое значение для определения состояний трубопроводов и нефтеперерабатывающих заводов;
- решение вопросов, связанных с окружающей средой и безопасностью. Анализ больших данных играет важную роль в защите окружающей среды, прогнозировании природных явлений и решении вопросов, связанных с безопасностью, которые могут быть побочным продуктом неоптимального управления операциями;
- улучшение отношений с общественностью и укрепление связи с государственными органами. Понимание общественного мнения и оперативное реагирование на запросы со стороны государственных и правительственных учреждений чрезвычайно важны при установлении доверительных отношений с более широкими сообществами;
- повышение рентабельности и своевременной цепочки поставок при материально-техническом обеспечении. Своевременная поставка запчастей, расходных материалов, оборудования и персонала является ключевым фактором в работе предприятия. Он обеспечивает оптимальное сохранение производственных мощностей и минимальный простой времени. Эффективное управление логистики также может обеспечить более экономически эффективную цепочку поставок;
- анализ рынка текущих цен, являющийся решающим фактором в определении того, когда надо проводить исследования, расширять производство, покупать и продавать недвижимость и менять инвестиционную стратегию;
- эффективное использование современных информационных технологий. Эффективность информационных технологий при определении больших объемов данных часто бывает трудно доказуемой, но иногда ИТ-организации склонны к первоначальному обоснованию при развертывании этих типов решений.

Технологии BIG Data дают возможность оптимизировать отображение геологической среды, указывать, каким образом и где бурить, определять производительность труда, а также обеспечивать точность в работе грузовиков и железнодорожных операций. Используя аналитику больших данных, можно увеличить время бесперебойной работы, сократить стоимость обслуживания техники, повысить производительность труда, уменьшить число ошибок и переделок. Области функционирования, трудности и возможности решения вышеупомянутых задач с помощью технологий BIG Data показаны ниже.

Исследование. В рыночных условиях трудности являются одной из наиболее "жизнеспособных" областей. Анализ больших данных при этом дает возможности:

- улучшения геологического анализа с помощью моделирования неисправностей и сейсмического анализа, что приводит к большим находкам, более предсказуемым результатам, а также к лучшему планированию скважин;
- улучшение прогнозирования рынка, чтобы определять, когда возможны разведка и добыча.

Бурение и добыча. Основной проблемой является повышение эффективности работы при одновременном соблюдении экологических стандартов. Анализ больших данных дает возможности:

- лучшей диагностики обслуживания, способствующей более высокому уровню доступности оборудования;
- предсказывать и избегать риска катастрофического сбоя датчиков на сверлах на основании показателей;
- улучшения управления энергопотреблением, что позволяет снижать стоимость бурения;
- эффективного управления активами и персоналом, снижающего стоимость производства;
- дистанционного управления операциями и мониторинга трубопроводов и оборудования для более низкой стоимости, снижения риска для окружающей среды, а также повышения безопасности.

Цепи поставок и материально-техническое обеспечение. Оптимальное время для доставки частей, расходных материалов и персонала, необходимых при разведке и добыче на участках, является основной проблемой. Анализ больших данных дает возможности:

- прогнозировать время поставки важнейших частей, оборудования и квалифицированного персонала;
- минимализировать время простоя;
- оптимально использовать топливо и стоимость доставки.

Управление недвижимостью. Продажа и вывод из эксплуатации недвижимости не считается финансово привлекательным. Анализ больших данных дает возможности:

- продажи недвижимости, находящейся в надлежащем состоянии, за приемлемую цену;
- продажи недвижимости для финансирования новых исследований на более прибыльных местах;
- слежения за соблюдением экологических норм на местных участках и несением потенциальных расходов по очистке.

Отношения с общественностью и государственными органами. Основой этой задачи является поддержание хороших рабочих отношений с обществом и

правительственными чиновниками. Анализ больших данных дает возможности:

- быстрого определения общественных настроений по отношению к разведке, бурению и добыче;
- более активного реагирования в случае внезапного изменения настроений персонала или руководства;
- быстрого реагирования на требования по воздействию на окружающую среду и другие производственные процессы бурения.

Прогнозирование цен на рынке. Необходимо для понимания направления рыночного ценообразования и будущего влияния на рентабельность участков, которые в настоящее время изучаются и в производстве. Анализ больших данных дает возможности:

- принятия более взвешенного решения о том, где и когда начать разведку нефти и газа;
- лучшего распределения кадров, оборудования и расходных материалов по требованиям;
- более эффективного управления недвижимостью.

Опыт нефтегазовых компаний в области больших данных

В большинстве нефтегазовых компаний исследования вводятся по основным компонентам. Менеджмент, безопасность и руководство имеют решающее значение и всегда стоят на первом месте. Типичная информационная архитектура для обработки больших данных показана на рис. 2 (см. четвертую сторону обложки).

Метод используется многими нефтегазовыми компаниями, где данные сохраняются и структурируются в зависимости от обстоятельств, а затем анализируются для принятия значимых решений [7]. Разнообразие базовых платформ обеспечивает решающую роль в нефтегазовых компаниях.

Некоторые компании используют методологии на основе поэтапных подходов. Эти подходы почти всегда наиболее успешны там, где для обработки больших данных необходимо определить текущее состояние и пробелы, чтобы лучше понять, в каком направлении двигаться дальше.

Компания **Royal Dutch Shell** уже некоторое время разрабатывает идею "нефтяного месторождения с управляемыми данными" в попытке снизить стоимость бурения, так как это главный расход в отрасли. **Royal Dutch Shell** является одной из крупнейших нефтегазовых компаний и четвертой в мире по объему выручки после **BP**, **Chevron**, **Total** и **ExxonMobil** [8].

Shell использует волоконно-оптические кабели, созданные в партнерстве с **Hewlett-Packard** для этих датчиков, и данные, передаваемые на свои частные серверы, поддерживаемые **Amazon Web Services** [8, 9]. Это дает гораздо более точную картину больших данных. Чтобы справиться с циклопическим массивом данных, **Shell** начала работать с **Nadoor** в **Amazon Virtual Private Cloud** [9]. Новые датчики могут сохранить большее количество данных, которые затем анализируются с использованием искусственного интеллекта и создаются 3D- и 4D-карты нефтяных скважин и резервуаров.

Использование интеллектуального анализа данных также помогает обнаружить нефть в казалось бы уже иссякших скважинах. Что именно там происхо-

дит и как проводится анализ данных — секрет фирмы и делиться опытом **Shell** категорически отказывается. Стало известно, что с помощью больших данных ожидается построение подробнейшей модели структуры Земли до глубин от 1 до 12 км. Если учесть, что в будущих планах компании предполагается развертывание датчиков примерно на 10 тысячах скважин, то возможно, что скоро одна только **Shell** будет владеть десятью эксабайтами данных.

По словам президента энергетической технологии компании **Chevron** Пола Сигла (**Paul Siegele**) данные, участвующие в разведке нефти и газа, колеблются. Тем не менее, без технологии, которая включает в себя распределенные датчики, высокоскоростные средства связи, массовые операции передачи данных по добыче и управлению данными дистанционных буровых работ, решение задач поиска и использования новых природных ресурсов было бы почти невозможными. Сегодня внутренний ИТ-трафик компании **Chevron** превышает 1,5 тбайт в день [10].

Chevron инвестирует значительные средства в технологии визуализации, которые позволяют создавать карты энергетических залежей, располагающихся на десятки тысяч футов ниже поверхности дна океана. Инновации в области информационных технологий, в том числе возможности визуализации больших данных и прогнозного анализа, дают возможность таким компаниям, как **Shell**, **BP**, **PLC** и **Chevron**, используя поток данных, выявить новые месторождения нефти и газа.

BP объявила в декабре прошлого года, что она строит самый большой в мире коммерческих исследований суперкомпьютер в Хьюстоне. Используемые компанией аналитические системы значительно продлили жизнь крупнейших нефтяных месторождений, принадлежащих компании [11].

Компания **Schlumberger** подписала с американской компанией **Chevron** контракт по обеспечению доступа к программной платформе **Petrel E & P**. Программа предназначена для сбора и обработки больших данных. По данным **Schlumberger**, программная платформа **Petrel E & P** позволяет нефтяным и газовым компаниям стандартизировать рабочие процессы от разведки до производства и принимать более обоснованные решения с четким пониманием возможностей и рисков [12].

Компания **Halliburton** имеет в своем арсенале пакеты, помогающие предсказывать простои и незаметные потери времени за счет анализа в режиме реального времени на платформе **Decision Space**. В сфере сейсмических исследований внедрены системы автоматического обнаружения отказов на базе машинного обучения. Кроме того, для анализа вращения при каротаже во время бурения применяются такие алгоритмы, как метод опорных векторов и метод регрессии. В течение многих лет некоторые специалисты выступают за использование данных алгоритмов в нефтегазовой отрасли. Однако никогда ранее нефтегазовая отрасль не генерировала такие большие объемы данных, как сегодня. На сегодняшний день алгоритмы машинного обучения и визуализация данных реализуются в виде моделей предсказания или в режиме реального времени. Из проектов однократ-

ного применения анализ превратился в непрерывный процесс совершенствования [13].

Компания Halliburton предлагает пакет решений для анализа данных в сфере моделирования скорости проходки при бурении, например, с использованием обучения анализу данных на базе нескольких машин. На основании полученных в системе неявных знаний модуль анализа незаметных потерь времени (Decision Space Invisible Lost Time, DSILT) определяет возможные неполадки в процессе бурения и дает рекомендации, какие действия следует предпринять. Для построения моделей DSILT требуются компьютеры высокой производительности. Системы реального времени для моделирования DSILT могут работать на настольных компьютерах и малых устройствах непосредственно на месторождении. В соответствии с другими подходами также широко используется кластерный анализ больших массивов данных [7, 13].

Харольд Уолтерс (Harold Walters), главный технический консультант компании Halliburton, заявил, что буквально через десять лет машинное обучение и анализ больших массивов данных получат повсеместное распространение и будут внедряться всеми специалистами, отвечающими за принятие решений [14].

Для компании **Saudi Aramco** анализ больших данных проводится несколькими организациями. Данные собираются в хранилище данных, которое обеспечивает объединение, очистку и представление данных. После того как данные поступили в хранилище данных, становится намного легче обнаруживать аномалии. Для этого используются сложные аналитические запросы. Преимущество подхода в том, что в то время, как можно сканировать в режиме реального времени хранилища данных, базы являются гораздо более управляемыми, и это облегчает сканирование недостающих данных и шаблонов для поиска, где данные не вяжутся с остальной частью показаний [15].

Один из подходов для обработки больших данных ориентирован на кластерную систему хранения данных (Isilon). Российская компания Sky Technics Group (STG) ввела в эксплуатацию Isilon для одной из крупнейших нефтяных компаний России. Главными достоинствами подхода являются не только производительность и масштабируемость, но и простота управления, связанная с автоматизацией большинства процессов [16].

В компании "**Сургутнефтегаз**" данные из ERP (Enterprise Resource Planning) в реальном времени передаются в СУБД HANA (High-Performance Analytic Appliance). HANA — высокопроизводительная NewSQL платформа для хранения и обработки данных, которая обеспечивает работу со сложными аналитическими запросами. Этот подход стал реальной инновацией, благодаря которой удалось на порядок сократить основные параметры обработки информации и попутно упростить архитектуру информационных систем. В компании функционируют 29 приложений SAP, и появление резидентных технологий позволило обслуживать их с помощью единой СУБД. Нужно отметить и то, что одним из крупнейших внедрений SAP HANA в мире является инсталляция серверов SAP HANA для компании Shell. Серверы SAP HANA сокращают цикл бизнес-процессов Shell от нескольких дней до нескольких часов [17].

Норвежская специальная служба по информационным технологиям Ziebel объявила, что она завершила сбор данных по нефтегазовой отрасли по заказу компании, желающей остаться неизвестной. В течение восьми месяцев были собраны 1708 Тбайт данных о скважинах в Северном море. Данные были получены с помощью Z-System Ziebel, которая использует волоконно-оптические технологии композитного стержня для доступа и визуализации стволов скважин в режиме реального времени. Z-System также может получить информацию, относящуюся к оптимизации потока скважин, контролю рисков целостности, моделированию пласта и интенсификации добычи нефти с минимальным производством. Z-System обеспечивает хранение и представление этих данных в эффективной и доступной форме. Данные хранятся в пяти серверах Ziebel, каждый из которых включает в себя 36 жестких дисков и имеет емкость 120 Тбайт [18].

Компания **ExxonMobil** считается инициатором многих технологических достижений по сбору и анализу сейсмических данных для моделирования сложной геологической среды, выявления вероятности сбоя во время бурения при одновременном снижении воздействия на окружающую среду [19]. В данный момент в компании ExxonMobil исследования в основном включают в себя следующие направления [20]:

- разработка методов статистического вывода данных с использованием сложных моделей для выявления недостающих данных;
- разработка алгоритмов распознавания образов для многомерных наборов данных;
- построение моделей для временных рядов и пространственно-временных данных;
- разработка высокопроизводительных кодов в режиме реального времени.

PetroChina по величине производства нефти и газа и в плане доходов является одной из передовых компаний в мире. PetroChina стала использовать EMC Isilon для анализа больших данных, полученных из северо-западного филиала своего научно-исследовательского института нефтегазовой разведки и разработки (PetroChina NWGI).

Система EMC Isilon была разработана корпорацией EMC (Electro-Motive Diesel). EMC — один из мировых лидеров в области информационных технологий. PetroChina NWGI консолидировала ряд месторождений нефти и газа, используя модель Isilon X-Series, предложенную EMC и снабжаемую электроэнергией от операционной системы OneFS для сбора и распределения приложений на один общий ресурс хранения, что упрощает управление большими данными для своих критически важных операций. С помощью системы Isilon PetroChina NWGI может масштабировать производительность и емкость линейно или независимо друг от друга по мере изменения потребностей бизнеса, обеспечивая максимальную гибкость в решении массовости и пропускной способности потребностей инфраструктуры ИТ для больших данных.

PetroChina NWGI выполняет ряд геолого-геофизических исследований и разработок по всему Китаю, сосредоточенных на выявлении, добыче, производстве и распределению ресурсов нефти и газа по всему миру. До Isilon традиционные системы хранения

PetroChina NWGI не могли идти в ногу с экспоненциальным ростом данных организаций и интенсивных требований производительности, снижая производительность рабочих процессов и наращивание эксплуатационных расходов.

Используя Isilon, PetroChina NWGI может предоставить своим ученым высокий уровень быстрого одновременного доступа к сейсмическим данным и приложениям, улучшая взаимодействие и операционную эффективность. С 2015 г. PetroChina начала сотрудничать с компанией Alipay в области мобильных платежей, больших объемов данных и других областях. Сотрудники компании признались, что сотрудничество с Alibaba положило начало облачным вычислениям для всестороннего повышения добычи нефти и оказывает значительное влияние на форму обслуживания нефтяных станций [21].

Согласно отчету компании Industrial Internet Insights за 2015 г., большинство компаний нефтегазовой промышленности (65 %) используют большие аналитические данные только для выявления эксплуатационных проблем. Тем не менее, одна треть (29 %) из 250 опрошенных руководителей компании призналась, что для исследования, прогнозного анализа и оптимизации своего бизнеса использует большие данные [22]. Из отчетов компаний [22, 23] ясно, что применение технологий больших данных в нефтегазовой промышленности находится все еще в экспериментальной стадии. Лишь немногие компании используют большие данные для принятия решений. Важность анализа больших данных в нефтегазовой отрасли по-прежнему определяется коммерческими задачами — использования имеющихся данных для повышения эффективности и экономичности и по возможности исключения потенциальных и фактических потерь при операционной деятельности.

Заключение

Разведка и добыча нефти являются сложным и критически важным заданием для поставщиков энергии. Чтобы эффективно и точно находить новые ресурсы и оптимизировать производство из существующих резервуаров, приходится каждый год увеличивать расходы предприятий. Исследования показали, что анализ больших данных является ключевой задачей для увеличения производства. С учетом высокого уровня производительности, который необходимо обеспечивать в нефтегазовой отрасли, первоочередными задачами являются также снижение затрат и безопасность рабочих и окружающей среды.

Исследования показали, что как крупные корпорации, так и независимые компании стремятся эффективно использовать большие данные, чтобы снизить расходы в цепочке поставок, а также оптимизировать стратегии закачивания скважин и уменьшения простоев. В распоряжении компаний оказываются огромные объемы данных из самых различных сфер, и для решения проблем, связанных с большими данными, существуют разные подходы. В ближайшем будущем лидерами будут только те компании, которые смогут эффективно использовать новые возможности анализа больших данных.

Подводя итог, нужно отметить, что значительного снижения затрат можно добиться в первую очередь за счет эффективного использования больших данных. И, наконец, помимо огромных прибылей для нефтяных компаний и выгоды для потребителей есть немалые шансы на то, что большие данные в базах данных нефтяных компаний когда-нибудь пригодятся и в других областях, на первый взгляд совершенно не связанных с нефтегазовой добычей.

Список литературы

1. **Suthaharan S.** Big Data classification: problems and challenges in network intrusion prediction with machine learning // ACM SIGMETRICS Performance Evaluation Review. 2014. Vol. 41. № 4. P. 70—73.
2. **Farris A.** How big data is changing the oil & gas industry // Analytics Informs. 2012, P. 20—47. URL: <http://viewer.zmags.com/publication/90ffcc6b#/90ffcc6b/21>
3. **Chevron** Drills Down Big Data Assets, 2012. URL: http://www.datanami.com/2012/05/10/chevron_drills_down_big_data_assets/
4. **Парошина И.** Большая нефть, "большие данные", большие деньги // Компьютерра. 2013. URL: <http://www.computerra.ru/88216/oil-bigdata/>
5. **Holdaway K.** Harness oil and gas Big Data with analytics: optimize exploration and production with data-driven models. John Wiley Sons Inc. 2014. 384 p.
6. **Кузнецов В. М., Шехтман Г. А., Череповский А. В.** Методика наблюдений в многоволновой сейсморазведке // Технологии сейсморазведки. 2013. № 2. С. 37—59. URL: http://ts.sbras.ru/ru/articles/2013-02_037.pdf
7. **Oracle** enterprise architecture white paper. Improving oil & gas performance with big data. URL: <http://www.oracle.com/us/technologies/big-data/big-data-oil-gas-2515144.pdf>
8. **Global** Onshore Oil and Gas Market 2015-2019 with BP, Chevron, ExxonMobil, Royal Dutch Shell & Total Dominating. URL: <http://www.prnewswire.co.uk/>
9. **Bernard M.** Big Data In Big Oil: How Shell Uses Analytics To Drive Business Success // Forbes Tech. Forbes, Inc. 26 May 2015. URL: <http://www.forbes.com/sites/bernardmarr/2015/05/26/big-data-in-big-oil-how-shell-uses-analytics-to-drive-business-success>.
10. **Baaziz A., Quoniam L.** How to use Big Data technologies to optimize operations in upstream petroleum industry // International Journal of Innovation, 2013, vol. 1, no. 1, pp. 19—25.
11. **Big Oil's** Big Data Push Changing the Future of Energy, 2013. URL: <http://blogs.wsj.com/cio/2013/01/02/big-oils-big-data-push-changing-the-future-of-energy/>
12. **Chevron** and Schlumberger sign data analytics contract, 2015. URL: <http://www.oilreviewmiddleeast.com/information-technology/chevron-and-schlumberger-sign-data-analytics-contract>
13. **Salas A.** Decision Space Geosciences (DSG) Software — End-to-End Solution // 14th International Congress of the Brazilian Geophysical Society and EXPOGEF, Rio de Janeiro. August 3—6, 2015. P. 3.
14. **Walters H.** From the digital oilfield to big data analytics, 2015. URL: <http://www.adjacentoilandgas.co.uk/it-cyber-security/from-the-digital-oilfield-to-big-data-analytics/22100/>
15. **Khakwani M. Sh.** Enterprise Data Solution Leveraging Data Warehousing for Big Data Veracity at Saudi Aramco // Proceedings of the 6th International Conference on Advances in Databases, Knowledge and Data Applications. Chamonix, France. April 20—24, 2014. P. 74—79.
16. **Смирнов Н.** Big Data для нефтяников // Computerworld Россия. 2012. №18. URL: <http://www.osp.ru/cw/2012/18/13016696/>
17. **Мельник О.** Больше данных — больше денег? // CRN. 2013. № 3 (77). URL: <http://www.i-teco.ru/press/publications/579/>

18. **Venables M.** Big oil, big data // Oil & Gas Technology. 2015. URL: <http://www.oilandgastechology.net/upstream-news/big-oil-big-data>

19. **Cohen K.** Big data in the oil patch, ExxonMobil's Perspectives Blog, 2013. URL: <http://www.exxonmobilperspectives.com/2013/07/31/big-data-in-the-oil-patch/>

20. **Data Analytics and Optimization.** URL: <http://www.kdnuggets.com/jobs/13/09-25-exxonmobil-data-analytics-optimization.html>

21. **PetroChina** Extracts Maximum Value from Big Data Using EMC Isilon, 2011. URL: <http://luxembourg.emc.com/about/news/press/2011/20110607-01.htm>

22. **Industrial Internet Insights Report for 2015.** URL: https://www.accenture.com/us-en/_acnmedia/Accenture/next-gen/reassembling-industry/pdf/Accenture-Industrial-Internet-Changing-Competitive-Landscape-Industries.pdf

23. **URL:** http://influencemap.org/site/data/000/089/Influence-Map_Oil_Sector_October_2015.pdf

R. M. Aliguliyev, Head of Department, r.aliguliyev@gmail.com,

I. Ja. Alakperova, Head of Sector, airada.09@gmail.com

Institute of Information Technology of Azerbaijan National Academy of Sciences, Baku

Big DATA Problem in Oil and Gas Industry: Current State and Prospects

Topic big data in oil and gas industry is multifaceted. The biggest oil and gas companies of the world have long had to deal with big data for decision-making. To increase production and resist the competitors, oil and gas companies implement effective big data analytics. The article reveals the problems associated with big data, the classification of the data sources in the oil and gas industry. Also have been explored approaches used by biggest oil and gas companies to solve the problems associated with the storage and processing of big data.

It was confirmed that the volume and variety of data play a key role in increasing business opportunities giant oil and gas companies. Proposals for effective analysis of big data to predict the market prices in the investigation, drilling and materilano logistics and other tasks. Howled found that, taking into account the high level of performance, which should be provided in the oil and gas industry, as a priority are reducing costs and security of workers and the environment. The future prospects of using complete, accurate and fast processing of large data collected in the oil and gas companies have been identified that have become part of the architecture of information.

Keywords: oil & gas companies, big data, information technology, data analysis, analytics, complex analytical queries

References

1. **Suthaharan S.** Big Data classification: problems and challenges in network intrusion prediction with machine learning, *ACM SIGMETRICS Performance Evaluation Review*, 2014, vol. 41, no. 4, pp. 70–73.

2. **Farris A.** How big data is changing the oil & gas industry, *Analytics Inform*, 2012, pp. 20–47. <http://viewer.zmags.com/publication/90ffcc6b#90ffcc6b/21>

3. **Chevron** Drills Down Big Data Assets, 2012. URL: http://www.datanami.com/2012/05/10/chevron_drills_down_big_data_assets/

4. **Paroshina I.** Bolshaya neft, "bolshie dannie", bolshie denqi, *Kompyuterra*, 2013. URL: <http://www.computerra.ru/88216/oil-big-data/> (in Russian).

5. **Holdaway K.** *Harness oil and gas Big Data with analytics: optimize exploration and production with data-driven models.* John Wiley Sons Inc. 2014. 384 p.

6. **Kuznetsov B. M., Shextman Q. A., Cherepovckiy A. B.** Metodika nablyudeniya b mnogoovolnovikh seysmorazvedke, *Tekhnologii seysmorazvedki*, 2013, no. 2, pp. 37–59. URL: http://ts.sbras.ru/ru/articles/2013-02_037.pdf (in Russian).

7. **Oracle** enterprise architecture white paper. *Improving oil & gas performance with big data.* URL: <http://www.oracle.com/us/technologies/big-data/big-data-oil-gas-2515144.pdf>

8. **Global Onshore Oil and Gas Market 2015-2019 with BP, Chevron, ExxonMobil, Royal Dutch Shell & Total Dominating.** URL: <http://www.prnewswire.co.uk/>

9. **Bernard M.** Big Data In Big Oil: How Shell Uses Analytics To Drive Business Success. *Forbes Tech., Forbes, Inc. 26 May 2015.* URL: <http://www.forbes.com/sites/bernardmarr/2015/05/26/big-data-in-big-oil-how-shell-uses-analytics-to-drive-business-success>.

10. **Baaziz A., Quoniam L.** How to use Big Data technologies to optimize operations in upstream petroleum industry, *International Journal of Innovation*, 2013, vol. 1, no. 1, pp. 19–25.

11. **Big Oil's Big Data Push Changing the Future of Energy**, 2013. URL: <http://blogs.wsj.com/cio/2013/01/02/big-oils-big-data-push-changing-the-future-of-energy/>

12. **Chevron and Schlumberger sign data analytics contract**, 2015. URL: <http://www.oilreviewmiddleeast.com/information-technology/chevron-and-schlumberger-sign-data-analytics-contract>

13. **Salas A.** Decision Space Geosciences (DSG) Software — End-to-End Solution, *14th International Congress of the Brazilian Geophysical Society and EXPOGEF*, Rio de Janeiro, August 3–6, 2015, p. 3.

14. **Walters H.** *From the digital oilfield to big data analytics*, 2015. URL: <http://www.adjacentoilandgas.co.uk/it-cyber-security/from-the-digital-oilfield-to-big-data-analytics/22100/>

15. **Khakwani M. Sh.** Enterprise Data Solution Leveraging Data Warehousing for Big Data Veracity at Saudi Aramco, *Proceedings of the 6th International Conference on Advances in Databases, Knowledge and Data Applications*. ChamoniX, France. April 20–24, 2014. pp.74–79.

16. **Smirnov N.** Big Data dlya neftnyanikov, *Computerworld Russian*. 2012, no. 18. URL: <http://www.osp.ru/cw/2012/18/13016696/> (in Russian).

17. **Melnik O.** Bolshie dannie — bolshe deneq? CRN, 2013, no. 3 (77). URL: <http://www.i-teco.ru/press/publications/579/> (in Russian).

18. **Venables M.** Big oil, big data, *Oil & Gas Technology*. 2015. URL: <http://www.oilandgastechology.net/upstream-news/big-oil-big-data>

19. **Cohen K.** Big data in the oil patch, *ExxonMobil's Perspectives Blog*, 2013. URL: <http://www.exxonmobilperspectives.com/2013/07/31/big-data-in-the-oil-patch/>

20. **Data Analytics and Optimization.** URL: <http://www.kdnuggets.com/jobs/13/09-25-exxonmobil-data-analytics-optimization.html>

21. **PetroChina** Extracts Maximum Value from Big Data Using EMC Isilon, 2011. URL: <http://luxembourg.emc.com/about/news/press/2011/20110607-01.htm>

22. **Industrial Internet Insights Report for 2015.** URL: https://www.accenture.com/us-en/_acnmedia/Accenture/next-gen/reassembling-industry/pdf/Accenture-Industrial-Internet-Changing-Competitive-Landscape-Industries.pdf

23. **URL:** http://influencemap.org/site/data/000/089/Influence-Map_Oil_Sector_Octobe

ЖУРНАЛ В ЖУРНАЛЕ



**НЕЙРОСЕТЕВЫЕ
ТЕХНОЛОГИИ**

№ 11

НОЯБРЬ

2016

Главный редактор:

ГАЛУШКИН А.И.

Редакционная коллегия:

АВЕДЬЯН Э.Д.
БАЗИЯН Б.Х.
БЕНЕВОЛЕНСКИЙ С.Б.
БОРИСОВ В.В.
ГОРБАЧЕНКО В.И.
ЖДАНОВ А.А.
ЗЕФИРОВ Н.С.
ЗОЗУЛЯ Ю.И.
КРИЖИЖАНОВСКИЙ Б.В.
КУДРЯВЦЕВ В.Б.
КУЛИК С.Д.
КУРАВСКИЙ Л.С.
РЕДЬКО В.Г.
РУДИНСКИЙ А.В.
СИМОРОВ С.Н.
ФЕДУЛОВ А.С.
ЧЕРВЯКОВ Н.И.

**Иностранные
члены редколлегии:**

БОЯНОВ К.
ВЕЛИЧКОВСКИЙ Б.М.
ГРАБАРЧУК В.
РУТКОВСКИЙ Л.

Редакция:

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

Пекунов В. В.

Компактное описание вариативных полей физических величин в повторяющихся задачах моделирования атмосферного переноса загрязнителей 871

**Кургалин С. Д., Туровский Я. А.,
Борзунов С. В., Адаменко А. А.**

Теоретические аспекты оптимизации эволюционного обучения нейронных сетей с использованием "изоляторов" 875

В. В. Пекунов, д-р техн. наук, инженер, e-mail: pekunov@mail.ru
ОАО "Информатика", г. Иваново

Компактное описание вариативных полей физических величин в повторяющихся задачах моделирования атмосферного переноса загрязнителей

Рассматривается задача сжатия полей данных, полученных при математическом моделировании воздушных потоков. На первом этапе поля кластеризуются по схожим значениям с применением алгоритма k -средних. На втором этапе каждый из полученных метакластеров в свою очередь кластеризуется по пространственным переменным. На третьем этапе каждый из полученных кластеров описывается набором нейронных сетей (в этом и состоит суть предлагаемой компрессии). Предлагается специальный алгоритм выбора оптимального набора нейронных сетей. Алгоритм использует технику "прогноз—коррекция". Прогноз реализуется путем экстраполяции с помощью специально построенной интерполирующей функции. Коррекция реализуется с использованием результатов приближения пробной нейронной сетью. Данная схема дает сжатие полей данных в 45...70 раз. Задача имеет существенную вычислительную сложность и параллельно решается на многоядерной системе.

Ключевые слова: сжатие данных с потерями, кластеризация, нейронные сети, численное моделирование, адаптивный алгоритм, распараллеливание

Введение

Рассмотрим задачу моделирования переноса загрязнителей в определенной сложной расчетной области. В практике расчета экологической безопасности такая задача может решаться неоднократно, с варьированием как направления и скорости ветра, так и расположения и интенсивности источников загрязнителей. При этом целесообразно прибегнуть к однократному решению повторяющейся подзадачи расчета картины воздушных потоков для каждого конкретного направления/скорости ветра с сохранением результатов расчета в файлах. Отметим, что в некоторых случаях указанная подзадача может решаться путем интерполяции по данным из сохраненных результатов расчета для иных, близких к текущему, направлений/скоростей ветра. При больших размерах расчетной сетки (от нескольких сотен до миллионов узлов для областей сложной формы, включающих, например, несколько улиц со строениями разной высоты) объем сохраняемых конечных данных (полученного стационарного распределения воздушных потоков — трех полевых компонент вектора скорости и поля турбулентной вязкости) даже для одной подзадачи при использовании вещественных значений одинарной точности (4-байтных) может достигать нескольких десятков мегабайт. С учетом того, что число подзадач определяется произведением числа возможных направлений ветра (15...30) на число градаций скорости ветра (10...15), его величина может составлять от 150 до 450 подзадач. В результате общий объем сохраняемых данных может достичь одного гигабайта и выше.

Поэтому актуальна задача генерации компактных описаний получаемых вариативных полей физических величин (компонент скорости и турбулентной вязкости). Как показывает практика, для этих данных максимальный коэффициент сжатия без потерь такими достаточно качественными алгоритмами, как RРМD и LZMA, не превышает 0,50...0,65. Для получения большей степени сжатия (например, с коэффициентом 0,01...0,02) необходимо воспользоваться прибли-

женными методами (сжатие с потерями), например, интерполяцией данных нейронными сетями с сохранением их весовых коэффициентов. Применение нейронных сетей также решит проблему получения интерполированных значений физических величин для случая таких направлений/скоростей ветра, для которых нет сохраненных (сжатых или несжатых) данных.

С вычислительной точки зрения задача сжатия может сводиться к кластеризации и обучению динамически определяемого набора нейронных сетей в каждом кластере. Это достаточно затратные процедуры, поэтому для их решения целесообразно использовать параллельные вычисления.

Кластеризация

Поля физических величин обычно характеризуются наличием множества достаточно компактных групп схожих (принадлежащих некоему сравнительно небольшому интервалу) значений, взятых в узловых точках расчетной сетки. Это приводит к идее, что первый и второй этапы описания вариативного поля $V(x, y, z, a)$, где a — некий варьируемый параметр (например, направление ветра), должны состоять в двухслойной кластеризации: сначала метакластеры по значениям $Q(x_i, y_j, z_k) = f(V(x_i, y_j, z_k, a))$ в узловых точках (i, j, k) , затем в пределах каждого из полученных метакластеров — кластеры по координатам. Можно попытаться описать поле в пределах каждого из полученных кластеров нейронной сетью прямого распространения. На входы сети подаются как значение параметра a , так и значения координат x_i, y_j, z_k , а выходом сети будет $V(x_i, y_j, z_k, a)$ в узловой точке (i, j, k) при указанном значении a ¹.

¹ Заметим, что мы сознательно отказались от варианта, при котором сеть сразу же пытается предсказать все N_a значения: в таком случае вход для a отсутствует, но появляется N_a выходов, каждый p -й из которых выдает значение $V(x, y, z, a_p)$. Такой вариант, по нашему мнению, не обеспечивает должного качества интерполяции при небольшом числе нейронов и требует чрезвычайно большого времени на обучение.

Отметим, что поскольку первичная кластеризация должна выделять блоки с близкими значениями $V(x, y, z, a)$ по всем возможным значениям a , то $Q(x, y, z)$ должна являться некоей сверткой значений:

$$Q(x, y, z) = \sum_{p=1}^{N_a} w_p V(x, y, z, a_p), \text{ где } N_a \text{ — число воз-}$$

можных различных значений a , w_p — относительная значимость p -го значения a в общем спектре значений этого параметра. Целесообразно, например, давать большие значения w_p для таких a_p , при которых вероятны повышенные ошибки интерполяции нейронной сетью. Такие значения могут быть назначены исходя из соображений исследователя или определены, например, путем простого эксперимента по попытке обучения N_a нейронных сетей, каждая p -я из которых ($p = \overline{1, N_a}$) приближает все поле $V(x, y, z, a_p)$ при фиксированном a_p , принимая на входы все узловые точки области [1]. По полученным средним погрешностям приближения ε_p исходных данных p -й сетью и выставляются веса:

$$w_p = w_{\min} + \frac{w_{\max} - w_{\min}}{\sum_{r=1}^{N_a} \varepsilon_r} \varepsilon_p$$

где w_{\min} и w_{\max} — соответственно минимальная и максимальная величины значимости.

Построение нейросетевых описаний в кластерах

Третьим этапом описания вариативного поля $V(x, y, z, a)$ в каждом кластере, как и указано выше, является его приближение нейронной сетью. При этом достаточно интересным является вопрос, следует ли воспользоваться одной более сложной сетью для всех значений a_p или же разбить множество $\{a_p\}$ на подмножества, в каждом из которых применить более простую нейронную сеть. Как показали эксперименты по описанию полей скорости воздушных потоков, возможны оба варианта, причем второй требует меньше времени на обучение и дает лучшие результаты приблизительно в 1,5...2 раза чаще, что будет продемонстрировано ниже в следующем разделе. Однако в каждом конкретном случае необходим контекстный выбор разбиения $\{a_p\}$, что приводит к идее организации *адаптивного алгоритма, который выбирает вариант интерполяции* в пределах кластера. При этом, чтобы избежать полного перебора, целесообразно ограничиться лишь несколькими (3...4) основными вариантами, а для повышения скорости отбора вариантов воспользоваться предикцией погрешности для некоторых из них (по уже полученным данным), с ее частичной коррекцией по данным, полученным путем ограниченного расчета нескольких компонент новых вариантов.

Предлагается следующий адаптивный эвристический алгоритм.

1. Воспользуемся трехслойными нейронными сетями с экспоненциальными сигмодами в первых двух слоях и с линейной передаточной функцией в третьем выходном слое.

2. Определим начальное число нейронов при интерполяции единственной нейронной сетью. Для каждого значения a_p , $p = \overline{1, N_a}$, найдем среднее по кластеру значение q_p поля $V(x, y, z, a_p)$, где $(x, y, z) \in S$, причем S — множество точек текущего рассматриваемого кластера. Получаем вектор значений $q = (q_1, q_2, \dots, q_{N_a})$, по которому можно определить набор точек (p, q_p) . По этому набору проводим интерполяцию полиномом достаточно высокого порядка r , где $N_a/4 \leq r < N_a$. Находим число экстремумов N_e полученного полинома (достаточно разбить область определения полинома на малые интервалы и подсчитать суммарное число интервалов, на краях которых производная полинома меняет знак или на одном из краев имеет нулевое значение).

Тогда начальное число нейронов первого слоя

$$N_1 := \lfloor \max(N_{1\min}, \min(N_{1\max}, \alpha N_e)) \rfloor,$$

где N_e — число экстремумов, а второго слоя

$$N_2 := \lfloor \max(N_{2\min}, \min(N_{2\max}, \beta N_e)) \rfloor,$$

где $N_{1\min}$, $N_{1\max}$ — минимально и максимально допустимые количества нейронов первого слоя, α и β — эмпирические коэффициенты, $N_{2\min}$, $N_{2\max}$ — минимально и максимально допустимые количества нейронов второго слоя. В наших экспериментах использовались следующие значения:

$$N_{1\min} = 7, N_{2\min} = 3, N_{1\max} = 18, \\ N_{2\max} = 10, \alpha = 3,3, \beta = 1,4.$$

3. Задаем начальное число сетей $A_1 := 1$. Пусть $r := 1$.

4. Инициализируем массив ошибок $\forall s: E_s := 0$.

5. Если $r > 2$, то экстраполируем погрешность:

5.1. Пусть $g(s) = \left(\sum_{m=0}^{r-1} a_m s^m \right)^{-1}$. Задаем $r - 1$ точек

вида $(s; E_s)$, $s = \overline{1, r-1}$. Проводим по этим точкам интерполяцию функцией $g(s)$ методом наименьших квадратов, в результате определяем коэффициенты a_0, \dots, a_{r-1} .

5.2. Определяем $E_r := g(r)$. Если $E_r < 0$, то $E_r := 0$.

5.3. Оцениваем производную $D := (E_r - E_{r-1}) / (A_r - A_{r-1})$.

5.4. Если $D > 0$ или $|D| < D_{\text{доп}}$ (здесь $D_{\text{доп}}$ — минимально допустимое значение производной), то:

5.4.1. Предыдущее $(r - 1)$ -е приближение с $A_r/2$ сетями признается наиболее успешным.

5.4.2. Выход из алгоритма.

6. Равномерно распределяем значения a_p , $p = \overline{1, N_a}$, по A_r нейронным сетям.

7. Цикл по $q = \overline{1, A_r}$:

7.1. Обучаем q -ю нейронную сеть по связанным с ней значениям (x, y, z, a_p) во всех узлах текущего кластера. Определяем среднюю погрешность E_{avr}

7.2. Если $q = 1$ и $r > 2$, то:

7.2.1. Корректируем предсказанную погрешность: $\tilde{E} := 0,8E_{avr} + 0,2E_r$

7.2.2. Оцениваем скорректированное значение производной $D := (\tilde{E} - E_{r-1}) / (A_r - A_{r-1})$.

7.2.3. Если $D > 0$ или $|D| < D_{\text{доп}}$, то:

7.2.3.1. $E_r := E_{avr}$

7.2.3.2. Предыдущее ($r - 1$)-е приближение с $A_r/2$ нейронными сетями является самым успешным.

7.2.3.3. Выход из алгоритма.

7.3. $E_r := E_r + E_{avr}$

7.4. Конец цикла по q .

8. Определяем среднюю погрешность: $E_r := E_r/A_r$.

9. Если $E_r \leq E_{\text{треб}}$ (здесь $E_{\text{треб}}$ — требуемое минимально допустимое значение погрешности), то переходим к п. 15.

10. $A_{r+1} := 2A_r$

11. $r := r + 1$.

12. $N_1 := \max(N_{1\min}, \lfloor \gamma \cdot N_1 \rfloor)$.

13. $N_2 := \max(N_{2\min}, \lfloor \delta \cdot N_1 \rfloor)$.

14. Если $A_r \leq A_{\max}$, то переходим к п. 5.

15. Алгоритм завершен, разбиение с $A_r/2$ сетями является лучшим.

В наших экспериментах использовались следующие значения:

$$D_{\text{доп}} = \frac{0,05}{4}, E_{\text{треб}} = 0,075, \gamma = 0,6, \delta = 0,4, A_{\max} = 8.$$

Экспериментальная часть

В качестве примера рассмотрим построение компактных кластерно-нейросетевых описаний полей скорости воздушных течений в трехмерной расчетной области сложной формы, представляющей участок, включающий улицу Göttinger Straße города Ганновер (Hannover, Германия, 52°20' северной широты, 9°40' восточной долготы) и несколько примыкающих улиц. На рис. 1 представлен соответствующий фрагмент карты Ганновера. Данные о конфигурации расчетной

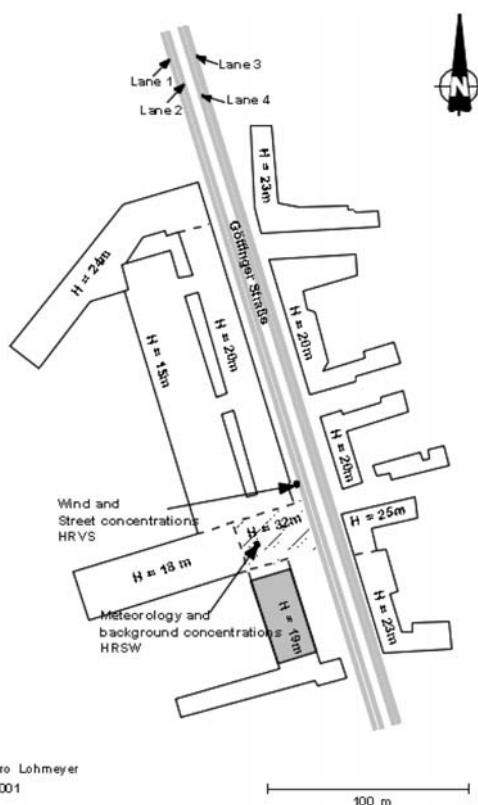


Рис. 1. Фрагмент карты города Ганновер, представляющий расчетную область

области и о различных условиях моделирования были получены из работ [2—4].

Имеющиеся данные (исходные для построения описаний) представляли результаты моделирования [5, 6] (получены на 128-ядерном кластере ИГЭУ на процессорах Opteron 2 ГГц, сеть передачи Gigabit Ethernet) с установившимися значениями полей скорости и турбулентной вязкости, полученные для 18 различных направлений ветра (от 20 до 360° с шагом 20°). Число узлов расчетной сетки составляло $72 \times 86 \times 45 = 278\,640$. Для хранения каждой из трех компонент полученных полей скорости использовались четырехбайтные значения одинарной точности. Соответственно, для хранения данных по каждой компоненте задействовались $18 \times 72 \times 86 \times 45 \times 4 = 20\,062\,080$ байт, т. е. около 19 Мбайт.

Для построения компактных описаний была разработана параллельная программа (М-скрипт) для Octave 3.8.4 (Linux-версия), исполняющаяся на стандартной 16-ядерной машине платформы Google's Compute Engine.

На первом этапе использовался алгоритм k -средних (функция kmeans) для разбиения на 20 метакластеров. На втором этапе использовалось распараллеливание по метакластерам в соответствии с парадигмой "портфель задач" (с той же функцией pararrayfun): каждый метакластер на отдельном ядре тем же алгоритмом (k -средних) делился на 15 кластеров, таким образом, общее число кластеров достигло 300. На третьем этапе при обучении нейронных сетей методом обратного распространения ошибки также использовалось распараллеливание по принципу "портфель задач": один метакластер — одна задача (одно ядро). При этом задавалось максимальное число прогонов по обучающей выборке, равное 100. Кроме того, обучение прекращалось при достижении среднеквадратичной ошибки, равной 0,075.

Использовались все 16 ядер, время обработки каждой компоненты скорости занимало в среднем около 28 мин. Для сравнения: на одном ядре процессора Intel Atom 1,86 ГГц (Windows) время обработки одной компоненты занимало около 16,5 ч. С поправкой на различие тактовых частот процессоров (в Google's Compute Engine используются процессоры Xeon 2,6 ГГц, Linux) можно подсчитать реальные (объясняющиеся преимущественно эффектом распараллеливания) ускорение и эффективность распараллеливания. Они составляют соответственно $\frac{16,5 \cdot 60 \cdot (1,86/2,6)}{28} = 25,3$ и $\frac{16,5 \cdot 60 \cdot (1,86/2,6)}{28 \cdot 16} = 158\%$.

Такой "суперлинейный" эффект объясняется, видимо, двумя факторами:

- а) большей эффективностью скомпилированного для Linux кода Octave по сравнению с кодом, полученным для Windows;
- б) более высокой производительностью процессора Intel Xeon в сравнении с Intel Atom.

В таблице представлены результаты построения компактных описаний по каждой из трех проекций вектора скорости.

Из таблицы очевидно, что при весьма высокой степени сжатия (в 43...71 раз) средняя абсолютная по-

Результаты построения компактных описаний

Параметр	Проекция скорости на ось x	Проекция скорости на ось y	Проекция скорости на ось z
Число весов сетей	68 483	90 505	55 394
Число сетей	965	1371	522
Число нейронов	11 372	15 494	7833
Среднее число сетей на кластер	3,22	4,57	1,74
Объем памяти (байт), требуемый для хранения весовых коэффициентов, сжатой информации о распределении узлов по кластерам, о конфигурации сетей	351 306	454 010	283 002
Коэффициент сжатия	0,018	0,023	0,014
Число кластеров с одной сетью	27	15	138
Число кластеров с двумя сетями	159	92	136
Число кластеров с четырьмя сетями	73	93	24
Число кластеров с восемью сетями	41	100	2
Средняя абсолютная погрешность, м/с	0,218	0,217	0,081

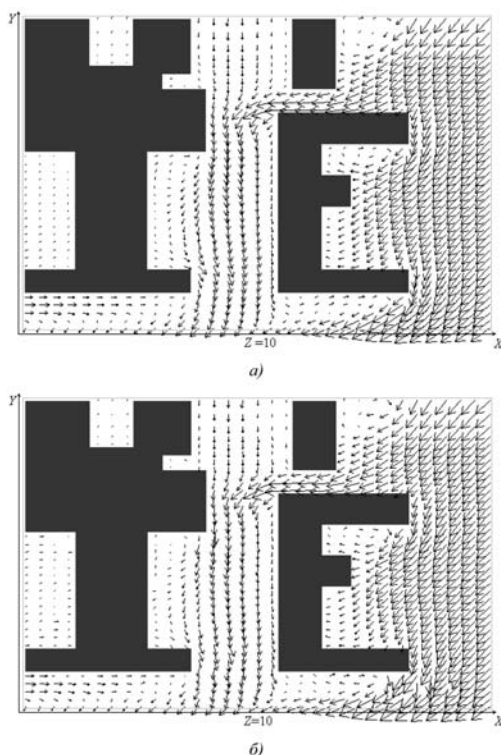


Рис. 2. Распределения воздушных потоков по результатам моделирования (а) и по описаниям (б)

грешность достаточно невелика с учетом того, что скорость воздушного течения в расчетной области принимала значения в диапазоне 0,1...9,5 м/с. Это подтверждается также визуально — на рис. 2 приведены картины скорости воздушных течений во фрагменте области на высоте 10 м, полученные в расчете (а) при направлении ветра в 20° и по компактным описаниям (б) расчетных данных. Очень показательны распределения чисел кластеров по числу задейство-

ванных сетей. По проекциям на оси Ox и Oy преобладают кластеры с более чем одной сетью, что объясняется достаточно сложным характером интерполируемых данных (подтверждается более высокой средней абсолютной погрешностью интерполяции). В то же время по проекции на ось Oz , напротив, доминируют кластеры с одной-двумя сетями, которые при сравнительно несложном характере исходных данных (погрешность интерполяции существенно ниже) уже дают требуемый результат.

На рис. 2, а, б отмечаются лишь небольшие искажения, что, вообще говоря, свидетельствует об адекватности построенных компактных кластерно-нейросетевых описаний. Соответственно, моделирование переноса загрязнителей при использовании восстановленных из описаний полей скорости также должно давать результаты, близкие к тем, которые можно получить при использовании исходных полей скорости.

Выводы

В данной работе предложены основные принципы построения компактных кластерно-нейросетевых описаний, варьирующихся в зависимости от некоего параметра физических величин. Подробно описан адаптивный эвристический алгоритм построения описаний. Алгоритм отличается наличием процедуры предикции-коррекции погрешности для предполагаемого к использованию количества нейронных сетей.

На данных, полученных в результате моделирования воздушных течений на трехмерном участке, включающем несколько улиц города Ганновер, показано, что получаемые таким образом приближенные описания в 43...71 раз компактнее исходных данных, при этом погрешность восстановленных данных имеет вполне приемлемую величину. С качественной точки зрения восстановленные данные также весьма схожи с исходными данными. Такие данные вполне могут применяться для расчета атмосферного переноса загрязнителей в задачах, подобных описанным в [5, 6].

Показаны высокие степень ускорения и эффективность распараллеливания на стандартной 16-ядерной машине платформы Google's Compute Engine.

Список литературы

1. Пекунов В. В. Локальные нейросетевые модели турбулентности // Мат. Междунар. науч.-техн. конф. "XVIII Бенардосовские чтения". Иваново, 2015. Т. 2. С. 331–334.
2. Ketzell M., Louka P., Sahn P., Guilloteau E., Sini J.-F. The use of computational fluid dynamics in modelling air quality in street canyons // Final report of the TMR research network TRAPOS "Optimisation of Modelling Methods for Traffic Pollution in Streets", submitted to the European Commission, July 2001. P. 86–93.
3. Ketzell M., Louka P., Sahn P., Guilloteau E., Sini J.-F., Moussiopoulos N. Intercomparison of Numerical Urban Dispersion Models — Part II: Street Canyon in Hannover, Germany // Proceedings of the 3rd Urban Air Quality conference — Loutraki 19–23 March 2000.
4. Schatzmann M., Leitl B., Liedtke J. Ausbreitung von Kfz-Abgasen in Straßenschluchten // Forschungsbericht FZKA-BWPLUS (Final report of the research project PEF 2 96 001). 1999. P. 1–132.
5. Пекунов В. В. Модель образования и распространения твердых, жидких и газообразных загрязнителей. Оптимальное распараллеливание // Математическое моделирование. 2009. Т. 21. № 3. С. 69–82.
6. Пекунов В. В. Новые методы параллельного моделирования распространения загрязнений в окрестности промышленных и муниципальных объектов // Дис. докт. тех. наук. Иваново, 2009. 274 с.

The Compact Description of Variative Fields of Physical Variables in the Repetitive Tasks of Simulation of Atmospheric Pollution's Transfer

This work deals with the problem of compactization of the data fields resulting from mathematical simulation of the atmospheric flows. In the first stage such fields are clustered (using k-means) by the similar values. In the second stage each such metacluster is clustered by space variables. In the third stage each resulting cluster is described by a sequence of neural networks (this is a proposed compression). A special algorithm of the selection of an optimal set of networks is proposed. This algorithm uses a prediction-correction technique. The prediction is realized by an extrapolation using a specially builded (interpolating) function. The correction is realized using a probe network results. The proposed scheme can give a compression of the data fields into 45–70 times. The task has a big calculating cost and is solved in parallel on the multicore machine.

Keywords: lossy data compression, clustering, artificial neural networks, numerical simulation, adaptive algorithm, parallelization

References

1. **Pekunov V. V.** Lokal'nye neyrosetevye modeli turbulentsnosti (Local neural network models of turbulence), *Mat. Mezhdunar. nauch.-tekh. konf. "XVIII Benardosovskie chteniya"* (Proceedings of the International Scientific-Technical Conference "XVIII Benardos readings"). Ivanovo, 2015, vol. 2, pp. 331–334 (in Russian).

2. **Ketzel M., Louka P., Sahm P., Guilloteau E., Sini J.-F.** The use of computational fluid dynamics in modelling air quality in street canyons, *Included in the final report of the TMR research network TRAPOS "Optimisation of Modelling Methods for Traffic Pollution in Streets"*, submitted to the European Commission, July 2001.

3. **Ketzel M., Louka P., Sahm P., Guilloteau E., Sini J.-F., Moussiopoulos N.** Intercomparison of Numerical Urban Dispersion Models. Part II: Street Canyon in Hannover, Germany, *The 3rd Urban Air Quality conference — Loutraki 19–23 March 2000*.

4. **Schatzmann M., Leitl B., Liedtke J.** *Ausbreitung von Kfz-Abgasen in Straßenschluchten / Forschungsbericht FZKA-BWPLUS* (Final report of the research project PEF 2 96 001). 1999.

5. **Pekunov V. V.** Model' obrazovaniya i rasprostraneniya tverdykh, zhidkikh i gazoobraznykh zagryazniteley. Optimal'noe rasparrallelivanie (Model of creation and propagation of solid, liquid and gaseous pollutants. Optimal Parallelization), *Matematicheskoe modelirovanie* [Mathematical Modelling], 2009, vol. 21, no. 3, pp. 69–82 (in Russian).

6. **Pekunov V. V.** Novye metody parallel'nogo modelirovaniya rasprostraneniya zagryazneniy v okrestnosti promyshlennykh i munitsipal'nykh ob"ektov (New methods of parallel modeling of pollution propagation near the plants and city objects), *Dis. dokt. tekhn. nauk* [Thesis of doctor of engineering]. Ivanovo, 2009 (in Russian).

УДК 004.5; 004.94

С. Д. Кургалин, д-р физ.-мат. наук., профессор, kurgalin@bk.ru,

Я. А. Туровский, канд. мед. наук, доц., yaroslav_turovsk@mail.ru,

С. В. Борзунов, канд. физ.-мат. наук, доц., borzunov@cs.vsu.ru

Воронежский государственный университет, Воронеж

А. А. Адаменко, аспирант, adamenko.artem@gmail.com

Воронежский государственный университет инженерных технологий, Воронеж

Теоретические аспекты оптимизации эволюционного обучения нейрочипов с использованием "изолятов"*

Для эволюционного подхода к обучению нейрочипов, применяемых для восстановления поврежденных нервной ткани, разработан алгоритм оптимизации, в основе которого лежит метод биологической изоляции в популяции. Предложено использовать понятие "изолята" — множества дочерних искусственных нейронных сетей (ИНС) одного поколения, имеющих одну и ту же родительскую ИНС. Представлены теоретические обоснования применения "изолятов" на основе модели линейного города. Они могут быть использованы в большинстве алгоритмов обучения ИНС без учителя, основывающихся на эволюции ИНС путем скрещивания их весовых коэффициентов. Разработанный алгоритм может быть применен при конструировании нейрочипов, используемых для восстановления поврежденной нервной ткани. Он позволяет в большинстве случаев определить правильное направление изменения весовых коэффициентов связей ИНС для уменьшения числа ошибок классификации и достигнуть лучших показателей обучения нейрочипа по сравнению с применением эволюционного алгоритма без "изолятов", что в конечном итоге приводит к улучшению результата восстановления поврежденной нервной ткани.

Ключевые слова: нейрочип, нервная ткань, искусственная нейронная сеть, эволюционный алгоритм, моделирование обучения

* Работа выполнена при частичной финансовой поддержке РФФИ (грант 16-29-08342-офи_м).

Введение

Одним из перспективных направлений исследований, которые проводят в целях восстановления работоспособности поврежденной нервной ткани, является создание нейрочипов (НЧ), которые при необходимости могут взять на себя часть функций нервной ткани, утраченных в результате ее повреждения. Нейрочипы, при их интеграции в нервную ткань, для выполнения своих функций должны быть "подстроены" под нее путем проведения этапа предварительного обучения. Один из вариантов обучения НЧ базируется на использовании искусственных нейронных сетей (ИНС). При этом реализованная тем или иным способом архитектура НЧ должна будет соответствовать функциям и структуре ИНС. В настоящее время существует большое число алгоритмов обучения ИНС "с учителем", в том числе и для ИНС прямого распространения. Однако одним из главных недостатков этих алгоритмов является необходимость наличия обучающих выходных паттернов для соответствующих входных векторов обучающей выборки (ОВ). Выходные паттерны представляют собой ожидаемый вектор классов, соответствующих обучающим входным векторам ОВ. Такой подход к обучению ИНС не представляется возможным для случая, когда ИНС нужно использовать в качестве инструмента для моделирования автоматизированного обучения НЧ для восстановления поврежденной нервной ткани, потому что неизвестны выходные векторы, в роли которых выступают паттерны нейронной активности.

Одной из основных проблем при создании НЧ является то, что существующие в настоящее время возможности не позволяют внедрять с требуемой степенью точности ($\ll 10$ мкм) в заданные участки нервной ткани необходимые для ее изучения приспособления: микроэлектроды, оптические волокна, микропипетки и т.п. Вследствие этого отсутствует возможность сформировать у исследователей и врачей требуемое для проведения реабилитации представление о функционировании той области нервной ткани, в которую вживлено устройство ввода/вывода информации для НЧ.

В случаях, когда не представляется возможным использование алгоритмов обучения ИНС с учителем для моделирования обучения НЧ, наиболее подходящим является применение эволюционного алгоритма [1]. Его удобно использовать, когда в силу особенностей работы НЧ неизвестны его выходные параметры, например, такие, как электрические паттерны активности нейронных микроансамблей, и, следовательно, неясно, куда следует "двигаться" весовым коэффициентам (ВК) межнейронных связей в их n -мерном пространстве для минимизации числа ошибочных распознаваний ИНС. Исходя из того, что конкретные параметры направления движения в n -мерном пространстве ВК отсутствуют, с целью минимизации ошибки распознавания ИНС направление этого движения следует выбирать путем наблюдения за реакцией того целостного организма, в который был вживлен НЧ. Однако у такого алгоритма обучения ИНС есть недостаток — в результате его работы значения функции ошибки в большинстве случаев ока-

зываются расположенными в локальном минимуме функции ошибки ИНС. Причиной этого является то, что "подстройка" ВК ИНС для достижения большего числа правильно распознанных классов ОВ происходит без определения направления изменения ВК в n -мерном пространстве ВК, в отличие от обучения с учителем, например, при использовании метода градиентного спуска. Поэтому для выхода функции ошибки ИНС из локального минимума предусмотрен механизм "встряхивания" ВК, который используется также и в других алгоритмах обучения [2,3]. "Встряхивание" заключается в том, что значения ВК изменяют случайным образом в целях вывода функции ошибки ИНС из локального минимума. Такое "встряхивание", реализованное в созданном в рамках настоящей работы программном обеспечении [4] для моделирования обучения НЧ, основано на применении коэффициента вариации позиций ВК в топологии ИНС определенного поколения. Но и применение этого метода часто бывает недостаточно для вывода функции ошибки ИНС из локальных минимумов и достижения наилучшего распознавания классов ОВ. Таким образом, представляется актуальным совершенствование метода эволюционного алгоритма обучения НЧ, используемого для восстановления поврежденной нервной ткани.

"Изолятом" будем называть множество дочерних ИНС одного поколения, имеющих одну и ту же родительскую ИНС. При создании дочерних ИНС от родительской происходит клонирование последней и изменение значений ВК у дочерней ИНС в пределах заданного диапазона уже имеющихся значений этих же ВК, хранящихся в векторе ВК родительской ИНС на момент клонирования. В итоге, в n -мерном пространстве ВК в окрестностях функции ошибки родительской ИНС формируется пул дочерних ИНС, а исходя из того, что значение функции ошибки родительской ИНС находится вблизи минимума функции ошибки, наличие дочерних ИНС позволят определить направление, в котором нужно двигаться для достижения этого минимума.

Цель настоящей работы — создание подхода на основе использования "изоляции" популяции ИНС применительно к эволюционному алгоритму, его теоретическое обоснование и демонстрация возможностей его применения для моделирования обучения НЧ, восстанавливающего поврежденную нервную ткань.

1. Подход на основе использования "изоляции" популяции ИНС

В настоящей работе предлагается подход на основе применения "изоляции" популяции ИНС и эволюционного алгоритма, использующий аналогию с методом биологической или социальной изоляции в популяции.

Сущность биологического "изолята" заключается в том, что объекты скрещиваются только внутри изолированной географической или социальной группы. Применительно к задачам эволюционного обучения это означает, что на определенном этапе эволюции ИНС скрещивание ВК будет проходить не во всей популяции ИНС, а лишь внутри отобранных субпопуляций, изолированных друг от друга в пространстве ВК. Скрещивание ИНС во всей популяции носит ве-

роятностный характер, следовательно, функция ошибки ИНС может попадать в области локальных минимумов функции ошибки.

Однако трудно определить, находятся ли результаты функции ошибки ИНС вблизи локальных минимумов функции ошибки, поэтому можно использовать лучшие по распознаванию классы ИНС из популяции в качестве родительских для формирования "изолятов".

На рис. 1 изображена блок-схема реализованного в программном пакете *ANNBuilder* алгоритма моделирования обучения НЧ с использованием эволюционного алгоритма. Выделен модуль, отвечающий за процесс "изоляции" ИНС. Каждый i -й ВК w_i^d в дочерней ИНС изменяется по отношению к родительскому ВК w_i^p следующим образом:

$$w_i^d = w_i^p \left(1 + \frac{r}{q} \right), \quad (1)$$

где $r \in [-1; 1]$ является случайной величиной; q — заданный процент изменения ВК дочерней ИНС; $i = 1, \dots, n$, где n — число всех ВК ИНС.

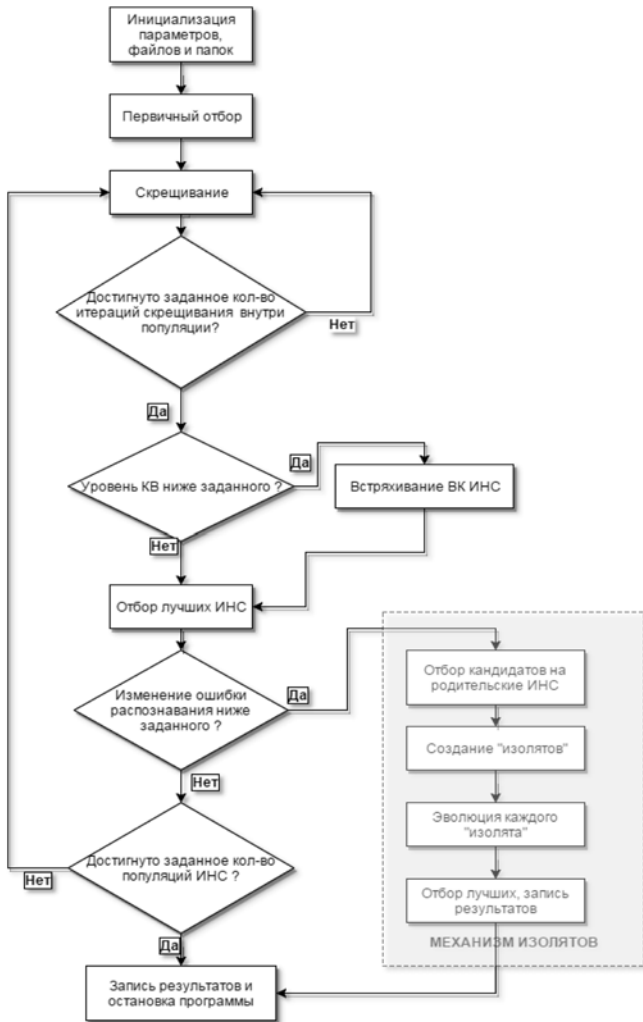


Рис. 1. Блок-схема алгоритма моделирования обучения НЧ с применением метода "изоляции" и использованием эволюционного алгоритма

Вектор \mathbf{W} ВК всей ИНС можно представить в виде совокупности ее ВК:

$$\mathbf{W} = (w_1, w_2, \dots, w_n), \quad (2)$$

Путем изменения родительских векторов создается k дочерних ИНС, содержащих векторы ВК вида (2) от родительской ИНС:

$$\mathbf{W}^{dk} = (w_1^{dk}, w_2^{dk}, \dots, w_n^{dk}), \quad k = 1, \dots, m, \quad (3)$$

где m — заданное число ИНС в поколении.

В свою очередь, в конкретной сессии работы алгоритма может быть t штук "изолятов". Каждый из них имеет по одной родительской ИНС от основного процесса работы эволюционного алгоритма. Только ИНС, получившие наилучшие результаты классификации, становятся родительскими, и дочерние ИНС порождаются именно от них при использовании каждой своего "изолята" t :

$$\mathbf{W}^{dkt} = (w_1^{dkt}, w_2^{dkt}, \dots, w_n^{dkt}), \quad t = 1, \dots, l, \quad (4)$$

где l — максимальное число родительских ИНС, формирующих совокупность "изолятов".

В итоге клонирования создается множество множеств ИНС. Каждый "изолят", в свою очередь, также проходит обучение эволюционным алгоритмом с заданными опциями [5, 6], тем самым формируя более высокую дисперсию значений ВК в какой-либо окрестности их значений, что дает возможность функции ошибки ИНС выйти из локального минимума.

Итак, по окончании эволюции каждой популяции ИНС сравнивается текущий результат распознавания ИНС с результатом, полученным для предыдущего поколения. Если модуль их разности оказывается меньше заданного порогового значения, то это указывает на отсутствие улучшения результата распознавания классов при эволюции ИНС. В этот момент и запускается процесс "изоляции" ИНС. После завершения процесса эволюции в каждом "изоляте" отбирается ИНС с наилучшим показателем распознавания классов, она в дальнейшем используется для выбора лучшей ИНС из всех "изолятов".

2. Основные исходы "изоляции" пулов ИНС

Покажем, что если в ряду итераций обучения "изолята" ИНС не достигла числа ошибок меньше, чем число ошибок для лучшей ИНС первого поколения клонирования "изолята", то значения функции ошибки "изолятов" находятся на "склоне" одного из локальных минимумов функции ошибки ИНС.

Данное утверждение будет применяться при "изоляции" с использованием эволюционного алгоритма в рамках моделирования обучения НЧ, восстанавливающего поврежденную нервную ткань.

Для доказательства утверждения удобно воспользоваться моделью линейного города [7, 8]. Пусть: а) используется один ВК w_0 ; б) существует функция ошибки $E(w_0)$; в) в ходе клонирования получено множество дочерних ИНС, имеющих один ВК w_0 такой, что $w_{\min} < w_0 < w_{\max}$; где w_{\min} и w_{\max} — наименьшее и наибольшее значение ВК среди ВК всего "изолята",

соответственно; г) ВК в первом поколении определяются как

$$w^1 = \frac{w_x^0 + w_y^0}{2}, \quad (5)$$

где w_x^0 и w_y^0 — случайно выбранные ВК для скрещивания первой популяции "изоляция". Тогда, если функция ошибки $E(w_0)$ монотонно убывает или монотонно возрастает в диапазоне изменения w от w_{\min} до w_{\max} , то ВК w^{z+1} поколения, следующего за z -м поколением, полученный как половина от суммы двух предыдущих ВК, не может быть большим максимального из слагаемых этой суммы.

Отметим, что для того, чтобы ускорить изменение в нужном направлении результата классификации с использованием ИНС, возможно получение ВК ИНС следующего поколения с использованием поправочных коэффициентов для w_x^0 и w_y^0 (в случае (5) они равны единице). Но так как значения функции $E(w)$ в окрестностях w_x^0 и w_y^0 неизвестны, то выбор их поправочных коэффициентов для w_x^0 и w_y^0 , в отличие от равных и единичных, весьма затруднен.

Рассмотрим иллюстрацию к доказательству указанного выше утверждения, представленную на рис. 2. Для иллюстрации было выбрано пять ИНС.

На рис. 2 указано пять точек, обозначающих нахождение лучших ИНС поколения в одномерном простран-

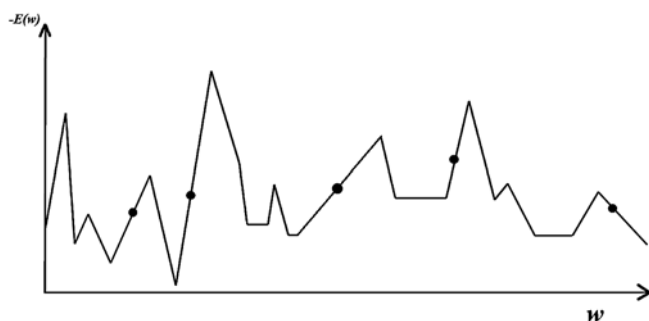


Рис. 2. Демонстрация расположения на графике функции $E(w)$ результатов расчета значений $E(w)$ ИНС для лучших ИНС первого поколения; w — ВК ИНС. Точками на графике обозначены значения функции ошибки родительских ИНС для "изолятов"

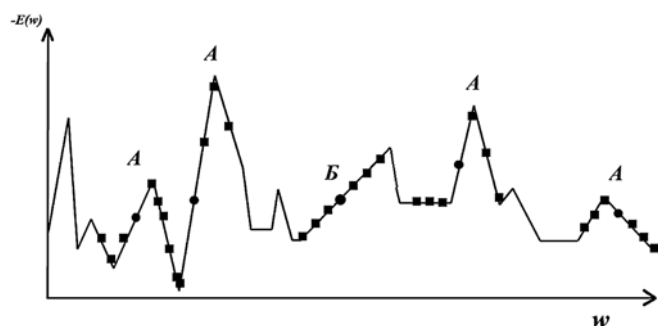


Рис. 3. Значения функции ошибки $E(w)$ для ИНС для распознавания дочерних и родительской ИНС ("изолятов") в одномерном пространстве ВК; w — ВК ИНС. Точками обозначены значения функции ошибки родительской ИНС, а квадратами — значения функции ошибки дочерних ИНС. Символом А на графике отмечено попадание дочерних ИНС относительно родительской ИНС оба склона функции ошибки $E(w)$; символом Б — попадание всех дочерних ИНС на один склон функции ошибки $E(w)$

стве ВК. Эти пять ИНС при запуске процесса "изоляции" станут родительскими. Для простоты иллюстрации использования утверждения было выбрано по шесть дочерних ИНС для каждого "изолята". После порождения дочерних ИНС получим следующее состояние поколения ИНС относительно ошибок классификации этого поколения. На рис. 3 квадратами обозначены ошибки распознавания теперь уже дочерних ИНС, полученных путем изменения ВК пяти родительских ИНС.

С учетом того, что результат скрещивания в разработанном программном обеспечении вычисляется как среднее значение между двумя ВК двух ИНС, из рис. 3 видно, что "изоляты", находящиеся на "склоне" функции ошибки, не дадут лучшего результата, чем результат для "изолята" при первичном клонировании. Если значения функции ошибки дочерних ИНС "изолята" находились на "пиках" или на других "склонах" пространства ошибок при первичном клонировании, то, исходя из средних значений ВК при скрещивании, будет получен лучший результат по сравнению с результатом использования "изолята" при первичном клонировании.

Таким образом, если применение "изолятов" не дает лучшего результата распознавания ИНС по сравнению с лучшим результатом при первичном клонировании, то из этого следует, что значения ВК ИНС "изолятов" находились в тех областях одномерного пространства ВК для текущей ОВ, которые дают результаты распознавания ИНС, попадающие на "склон" одного из локальных минимумов функции ошибки ИНС. Учтем, что если разброс ВК "изолятов" был слишком малым, т.е. значения функции ошибки дочерних ИНС расположены достаточно близко друг к другу и к значению ошибки родительской ИНС, то они ложатся на "склон" локального минимума функции $E(w)$, и мы не получим улучшения результатов распознавания ИНС. Если же ВК для дочерних ИНС достаточно сильно различаются между собой и отличаются от родительской ИНС, то значения функции ошибки дочерних ИНС могут "проваливаться" в локальные минимумы функции ошибки ИНС, и положительный эффект от применения "изолятов", заключающийся в увеличении числа правильных распознаваний классов, уменьшается. В этом случае необходимо заново выполнить клонирование ИНС внутри всего "изолята", взяв в качестве родительской ИНС какую-либо другую ИНС из "изолята" с наилучшим результатом правильного распознавания классов. При этом через несколько итераций клонирования будет достигнут минимум значений функции ошибки ИНС, и таким образом "изоляция" выполнит свою функцию, после чего работа с ним, согласно созданному алгоритму, будет завершена.

На рис. 4 представлен случай, когда все значения функции ошибки всех дочерних ИНС лежат только на одном из ее "склонов", а не на двух, вблизи локальных минимумов. Видно, что согласно утверждению, результат не может быть лучшим, чем значение функции ошибки в точке А, однако функция ошибки $E(w)$ позволяет определить направление изменения ВК, что обеспечивает перемещение значений функции ошибки всего "изолята" в сторону минимума функции ошибки ИНС и, таким образом, дает возможность обнаружить этот минимум.

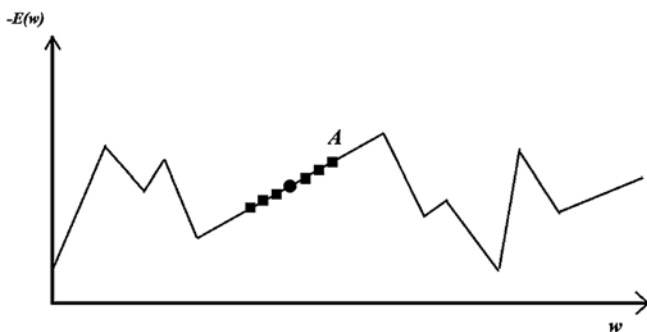


Рис. 4. Демонстрация попадания значений функции ошибки дочерних ИНС на склон функции ошибки $-E(w)$; w — ВК ИНС; точкой обозначено значение функции ошибки родительской ИНС, а квадратами — значения функции ошибки дочерних ИНС

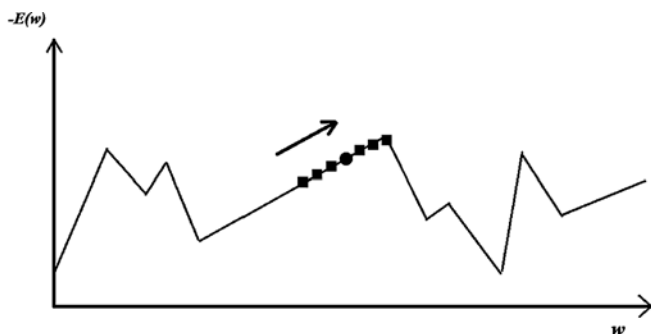


Рис. 5. Сдвиг значений функции $-E(w)$ для "изолята" вниз по "склону" в сторону минимума функции ошибки $-E(w)$ после создания вторичного изолята; w — ВК ИНС; точкой обозначено значение функции ошибки родительской ИНС, а квадратами — значения функции ошибки дочерних ИНС

Если после n -й итерации скрещивания ИНС внутри "изолята" значение функции ошибки ИНС изменится меньше, чем на заданный процент текущего значения ошибки относительно значения ошибки на предыдущей итерации скрещивания, то снова для клонирования берется ИНС, наилучшая по результату распознавания классов. Теперь ИНС берется уже из самого "изолята", и на основе нее делается повторное клонирование ИНС. Тем самым значение функции ошибки $E(w)$ сдвигается вниз по "склону" этой функции и "движется" в сторону ее локального минимума.

На рис. 5 изображен сдвиг значений функции $E(w)$ для "изолята" по "склону" функции ошибки в сторону локального минимума.

Таким образом, повторное создание "изолята" в случае, когда его действие перестает улучшать распознавание классов, позволяет значению ошибки функции ошибки ИНС переместиться вниз по "склону" этой функции в пространстве ошибок и улучшить результат распознавания классов ИНС для всего "изолята".

Представленные выше результаты позволяют сделать вывод, что применение "изолятов" дает возможность оптимизировать обучение НЧ, направленное на восстановление нервной ткани.

Отметим, что подходы, основанные на параллельном поиске нескольких локальных оптимумов в пространстве ВК ИНС с применением "изолятов", не имеют ограничений в области применения и могут быть использованы для решения широкого класса задач, связанных с разработкой классификаторов. Учи-

таявая быстрое развитие систем параллельных вычислений, использование представленного в настоящей работе подхода может в ряде случаев дать существенный выигрыш во времени по сравнению с подходами, применяющими последовательный поиск оптимумов классификатора.

Заключение

Для эволюционного алгоритма обучения ИНС, который используется для моделирования обучения НЧ, предложен алгоритм оптимизации обучения ИНС, в основе которого лежит способ биологической "изоляции". Предложены и обоснованы теоретические аспекты применения "изолятов" на основе модели линейного города. Данный подход оптимизирует поиск глобального минимума функции ошибки ИНС в пространстве ошибок путем изолирования лучших ИНС из всей популяции ИНС при их обучении с применением эволюционного алгоритма. Он может быть применен к большинству алгоритмов обучения ИНС без учителя, основывающихся на эволюции ИНС, путем скрещивания ВК.

Разработанный подход может быть применен для конструирования НЧ, используемого для восстановления поврежденной нервной ткани. Так как при обучении ИНС эволюционным алгоритмом отсутствуют конкретные параметры направления движения в n -мерном пространстве ВК и направление этого движения выбирается только путем наблюдения за реакцией ИНС, моделируя обучение НЧ, то разработанный подход позволяет в большинстве случаев определить правильное направление движения ВК для уменьшения числа ошибок классификации и достигнуть лучших результатов обучения НЧ по сравнению с использованием эволюционного алгоритма без "изолятов", и соответственно, улучшить результаты восстановления поврежденной нервной ткани.

Рассмотрение теоретических аспектов "изоляции" и ее использование при обучении ИНС с применением эволюционного алгоритма позволяет повысить эффективность обучения НЧ, внедренного в поврежденную нервную ткань в целях ее восстановления.

Список литературы

1. Туровский Я. А., Кургалин С. Д., Адаменко А. А. Автоматизированное обучение нейрочипов // Актуальные направления научных исследований века: теория и практика: сб. статей. Воронеж, 2015. Вып. 5—2 (16-2). С. 191—196.
2. Сараев П. В. Численные методы интервального анализа в обучении нейронных сетей // Автоматика и телемеханика. 2012. Вып. 11. С. 129—143.
3. Клюкин В. И., Николаенков Ю. К. Нейросетевые структуры и технологии. Часть 1. Электрические и математические модели нейронов. НС прямого распространения: учеб. пособие для вузов. Воронеж: Изд-во ВГУ, 2008. С. 32—37.
4. Туровский Я. А., Кургалин С. Д., Адаменко А. А. Моделирование обучения нейрочипов, внедренных в нервную ткань // Цифровая обработка сигналов. Воронеж, 2016. Вып. 1. С. 50—56.
5. Свидетельство о государственной регистрации программы для ЭВМ № 2015619800.
6. Свидетельство о государственной регистрации программы для ЭВМ № 2016614262.
7. Горбенко А. М. Модель линейного города с экзогенной конкуренцией по Штакельбергу // МТИП. 2013. Вып. 5, ч. 2. С. 64—81.
8. Лимонов Л. Э. Региональная экономика и пространственное развитие. М.: Изд-во Юрайт, 2015. С. 82—83.

S. D. Kurgalin, DSc, Professor, kurgalin@bk.ru,
Ya. A. Turovsky, PhD, Associate Professor, yaroslav_turovsk@mail.ru,
S. V. Borzunov, PhD, Associate Professor, borzunov@cs.vsu.ru
Voronezh State University, Voronezh
A. A. Adamenko, Postgraduate Student, adamenko.artem@gmail.com
Voronezh State University of Engineering Technologies, Voronezh

Theoretical Aspects of Optimization of Evolutionary Learning of Neurochips Using "Isolates"

The evolutionary learning algorithm neurochips, used to restore the damaged nervous tissue, proposed and developed an optimization algorithm based on the method of biological isolation. Proposed and justified the use of theoretical statements "isolates" based on the model of linear city. "Isolated" can be applied to most artificial neural network learning algorithm (INS) without a teacher, based on the evolution of the INS, by crossing the weighting coefficients. Developed approach can be used to construct neurochip used to restore the damaged nervous tissue. The proposed approach makes it possible in most cases to define the proper direction of weight change coefficient to reduce classification errors and achieve the best results neurochip training, compared with using an evolutionary algorithm without "isolates" and, consequently, improve the recovery of damaged nervous tissue.

Keywords: neurochip, nervous tissue, artificial neural networks, evolutionary algorithms, simulation training

References

1. **Turovskij Ja. A., Kurgalin S. D., Adamenko A. A.** Avtomatizirovanie obuchenija nejrochipov [Automate learning neurochips], *Aktual'nye napravlenija nauchnyh issledovanij XXI veka: teorija i praktika: sb. statej*, Voronezh, 2015, no. 5-2 (16-2), pp. 191–196.
2. **Saraev P. V.** Chislennye metody interval'nogo analiza v obuchenii nejronnyh setej [Numerical methods of interval analysis in training neural networks]. *Avtomat. i telemeh.* Voronezh, 2012, no. 11, pp. 129–143.
3. **Kljukin V. I., Nikolaenkov Ju. K.** Nejrosetevye struktury i tehnologii. Chast' 1. Jelektricheskie i matematicheskie modeli nejronov. NS prjamogo rasprostraneniya. [Neural structures and technologies. Part 1: Electric and mathematical models of neurons] Uchebnoe posobie dlja vuzov / Voronezh, VSU, 2008, pp. 32–37.
4. **Turovskij Ja. A., Kurgalin S. D., Adamenko A. A.** Modelirovanie obuchenija nejrochipov, vnedrennyh v nervnuju tkan' [Simulation training neurochips implanted in nervous tissue], *Cifrovaja obrabotka signalov*, Voronezh, 2016, no. 1, pp. 50–56.
5. **Svidetel'stvo o gosudarstvennoj registracii programmy dlja PC** [Certificate for state registration of the computer program] № 2015619800.
6. **Svidetel'stvo o gosudarstvennoj registracii programmy dlja PC** [Certificate for state registration of the computer program] № 2016614262.
7. **Torbenko A. M.,** Model' linejnogo goroda s jekzogennoj konkurenciej po Shtakel'bergu [Model of linear city with exogenous competition Stackelberg], *MTIP*, 2013, no. 5, vol. 2, pp. 64–81.
8. **Limonov L. Je.** Regional'naja jekonomika i prostranstvennoe razvitie [Regional economy and spatial development], Moscow, Izdatel'stvo Jurajt, 2015, pp. 82–83.

Адрес редакции:

107076, Москва, Стромьинский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Технический редактор *Е. В. Конова*.

Корректор *Т. В. Пчелкина*.

Сдано в набор 28.10.2016. Подписано в печать 20.10.2016. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ IT1116. 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.